PUBLICATION OF Little Lion Scientific R&D, Islamabad PAKISTAN
Journal of Theoretical and Applied Information Technology
15th May 2011. Vol. 27 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

MINING CUSTOMER DATA FOR DECISION MAKING USING NEW HYBRID CLASSIFICATION ALGORITHM

AURANGZEB KHAN, BAHARUM BAHARUDIN, KHAIRULLAH KHAN

Department of Computer and Information Sciences, Universiti Teknologi PETRONAS, Malaysia.

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. Sales patterns from inventory data indicate market trends and can be used in forecasting which has great potential for decision making, strategic planning and market competition. The objectives in this paper are 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 is useful mechanism for business support, investment and surveillance. In this paper we proposed an algorithm for mining patterns of huge 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 to find frequencies of property values of the corresponding items. MFP provides frequent patterns of item attributes in each category of products and also gives sales trend in a compact form. The experimental result shows that the proposed hybrid k-mean plus MFP algorithm can generate more useful pattern from large stock data.

Keywords: *Stock data, Most Frequent Patterns, Clustering, Decision Making,*

1. INTRODUCTION

Data is very important for every organization and business. Data that was measured in gigabytes until recently, is now being measured in terabytes, and will soon approach the peta byte 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. Segment by- segment sales forecasting can produce very useful information. The forecasting can be short term, mid term and long term. Long term forecasting may not produce accurate predictions. However it is very useful in understanding market trends [2]. It is easy to turn cash into inventory, but the challenge is to turn inventory into cash. Effective inventory management enables an organization to meet or exceed customer's expectations of product availability while maximizing net profits and minimizing costs [1, 3]. Only through data mining techniques, it is possible to extract useful pattern and association from the stock data [4]. Data mining techniques like clustering and associations can be used to find meaningful patterns for future predictions [5, 6]. Clustering is used to generate groups of related patterns, while association provides a way to get generalized rules of dependent variables. Patterns from a huge stock data on the basis of these rules can be obtained. The behavior in terms of sales transaction is significant. The general term used for such type of analysis is called Market Basket Analysis [7]. Typically there is lot of different items, placed in a market for selling, in which some of the product will be fast selling items, some will be slow selling items and some will be dead stock i.e. rarely selling items. We consider a scenario of super store or super market in distributed

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

environment, or internet based data processing environment. Decision making in business sector is considered as one of the critical tasks. There is study for data mining for inventory item selection with cross selling considerations which is used for maximal-profit selling items [8, 9]. But our problem is finding out the selling power of the products in the market. This is a useful approach to distinguish the selling frequency of items on the basis of the known attributes, e.g. we can examine that a "black coat of imperial company in winter season has high ratio of sale", here we have basic property related to this example, i.e. color, type, company, season, and location. Similarly we can predict that certain products of certain properties have what type of sale trends in different locations. Thus on the basis of this scenario we can predict the reason of dead-stock, slow moving and fast moving items. Data mining techniques are best suited for the analysis of such type of classification, useful patterns extraction and predictions.

The rest of the paper is organized as in section-2 present the related research and motivation for the proposed work. Section-3 descried the proposed method with pre-processing steps, in section-4 we elevates the results and finally in section-5 we conclude our proposed work.

2. BACKGROND AND RELATED WORK

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. L. K. Soon et al [25] compare the performance term of similarity search. 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 fined 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. Also such type of works includes [28] [29] [10] [11] [12].

In recent years, it has been recognized that the partitioned clustering technique is well suited for clustering a large dataset due to their relatively low computational requirements. The time complexity of the partitioning technique is almost linear, which makes it widely used. The best known partitioning clustering algorithm is the K-means algorithm and its variants [14]. This algorithm is simple, straightforward and is based on the firm foundation of analysis of variances. In addition to the K-means algorithm, several algorithms, such as Particle Swarm Optimization (PSO) [15] is another computational intelligence method that has already been applied to image clustering and other low dimensional datasets [16, 17]. In this work we have used clustering algorithm for clusters and MFP for pattern association among the cluster. The classification of similar objects into different groups, or the partition of data into subsets called clusters. The data in each subset share some common trait-often proximity accordingly to some defined distance measure [18].

In this work we have proposed an algorithm for mining patterns of huge stock data to show factors affecting the sale of products. In first phase, we divide the stock data in three different clusters categorically using k-means [20] algorithm on the basis of sold quantities of each category of items i.e. Dead-Stock (DS), Slow-Moving (SM) and Fast-Moving (FM). In the second phase we have proposed Most Frequent Pattern (MFP) algorithm to find frequencies of property values of the corresponding items. MFP provides frequent patterns of item attributes in each category of products. This work is similar to Apriori algorithm [26] for strong association among the patterns but it provides most visible patterns of associated objects. From the experimental results from sample data, it illustrated that the proposed algorithm of k-mean and association of MFP can generate more useful pattern from large stock data.

3. PROPOSED METHOD

In this study we proposed an algorithm for mining patterns of huge stock data to predict factors affecting the sale of products. In the first phase, we divide the stock data in three different clusters on

PUBLICATION OF Little Lion Scientific R&D, Islamabad PAKISTAN
Journal of Theoretical and Applied Information Technology
15th May 2011. Vol. 27 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

the basis of sold quantities i.e. Dead-Stock (DS), Slow-Moving (SM) and Fast-Moving (FM) using K-means algorithm. In the second phase we proposed Most Frequent Pattern (MFP) algorithm to find frequencies of property values of the corresponding items. MFP provides frequent patterns of item attributes in each category of products. Cluster analysis is widely used in market research when working with multivariate data from surveys and test panels. Market researchers use cluster analysis to partition the general population of consumers into market segmentation, intra and inter stock patterns and to understand better association between them [19].

3.1 Proposed Architecture

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.

3.2 K-MEANS

K-means [20] is a typical clustering algorithm and has used for classification of data for decades. Proximity is usually measured by some sort of distance; the most commonly being used is the Euclidean distance [21].

$$dist(i,j) = \sqrt{\sum_{k=1}^{l} \left(x_{ik} - x_{jk}\right)^2} \qquad (1)$$

The main idea is to define k centroids, one for each cluster. These centroids should be placed in a cunning way because of different location causes different result. So, the better choice is to place them as much as possible far away from each other.

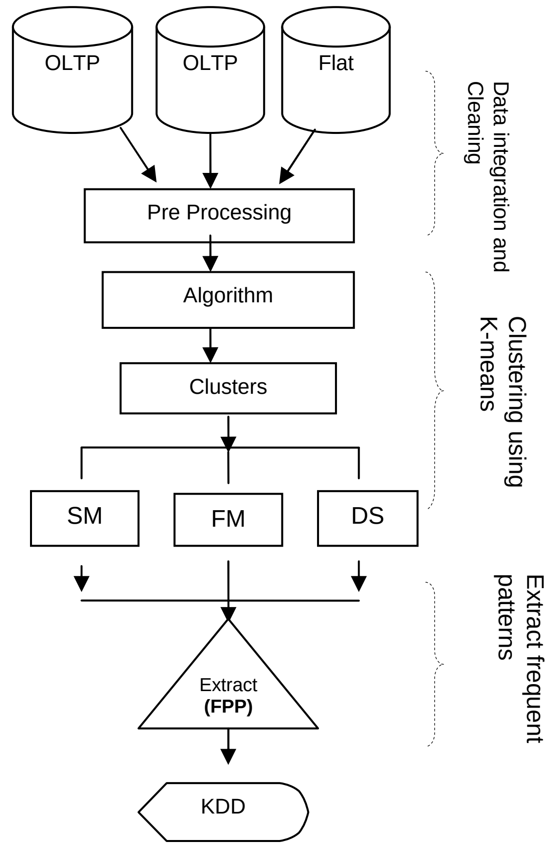

Figure 1: Block Diagram of Proposed Architecture

This algorithm aims at minimizing an objective function, in this case a squared error function. The objective function is

$$J = \sum_{j=1}^{k} \sum_{i=1}^{k} \left\| x_i^{(j)} - c_j \right\|^2 \qquad (2)$$

Where $\left\| x_i^{(j)} - c_j \right\|^2$ is a chosen distance measure between a data point (j) i x and the cluster centre j c , is an indicator of the distance of the n data points from their respective cluster centers. The steps of the K-mean [7] algorithm is as described below:

1. Place K points into the space represented by the objects that are being clustered; these points represent initial group centroids.
2. Assign each object to the group that has the closest centroid.

PUBLICATION OF Little Lion Scientific R&D, Islamabad PAKISTAN
Journal of Theoretical and Applied Information Technology
15th May 2011. Vol. 27 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3. When all objects have been assigned, recalculate the positions of the K centroids.

Repeat Steps 2 and 3 until the centroids no longer move. This produces a separation of the objects into groups from which the metric to be minimized can be calculated.

3.3 Most Frequent Pattern (MFP)

Association rule mining is one of the most important and well defines technique for extract correlations, frequent patterns, associations or causal structures among sets of items in the transaction databases or other repositories. Association rules are widely used in various areas such as risk management, telecomm, market analysis, inventory control, and stock data [27]. Apriori algorithm [26] for strong association among the patterns is highly recommended. In this work we proposed a new algorithm MFP that is more efficiently generates frequent patterns and strong association between them. For this purpose a property matrix containing counted values of corresponding properties of each product has been used as shown in Figure 2.

Let we have set X of N items in a Dataset having set Y of attributes. This algorithm counts maximum of each attribute values for each item in the dataset.

MFP Algorithm: *Let we have set X of N items in a Dataset having set Y of attributes. This algorithm counts maximum of each attribute values y_{ij} for each item in the dataset.*

Input: Datasets (DS)
Output: Matrix
Frequent Property Pattern (FPP):

```
FPP (DS)
Begin
        for each item Xi in DS
            a.  for each attribute
                    i.  count occurrences
                        for Xi
                      C=Count (Xi)
                    ii. Find attribute name
                        of C
                      Mi=Attribute (Ci)
            next [End of inner loop]
            b.  Find Most Frequent Pattern
                    i.  MFP=Combine(Mi)
        next [End of outer loop]
End
```

Figure 2: Proposed Algorithm for Frequents Pattern extraction

3.4 Data Pre-processing

Data in its original format never confirm to the required shape for data mining. It needs to be transformed, integrated, and aggregated so that the mining process can effectively perform on it. There is a need to process the data before it used in the knowledge discovery (KDD) process

```
cket,Blue,Large,Store1,Large,New York,June,410,700
Jacket,Blue,XLarge,Store3,Super,Los Angeles,July,20,850
Jacket,Green,Medium,Store6,Medium,Dallas,May,340,-95
Jacket,Green,Medium,Store9,Large,Las Vegas,June,425,-90
Shirt,Blue,Large,Store8,Small,DC,July,345,610
Shirt,Blue,Large,Store10,Super,Atlanta,August,350,-85
Shirt,Blue,Large,Store2,Medium,Boston,September,355,300
Shirt,Blue,Large,Store2,Medium,Boston,October,360,400
Coat,Black,Medium,Store1,Large,New York,February,430,150
Shirt,White,Small,Store1,Large,New York,March,305,160
Jacket,White,Medium,Store1,Large,New York,April,435,170
Jacket,Black,Large,Store1,Large,New York,May,300,180
Shirt,Black,XLarge,Store1,Large,New York,June,440,190
Coat,Brown,Small,Store5,Small,New York,July,300,5
Glove,White,Medium,Store5,Small,New York,August,445,15
Shirt,Blue,Large,Store5,Small,New York,September,305,25
Jacket,Blue,XLarge,Store5,Small,New York,October,450,35
Hat,Black,Small,Store5,Small,New York,November,310,45
Glove,Blue,Medium,Store5,Small,New York,December,455,55
Coat,Red,Large,Store10,Super,Atlanta,January,190,840
```

Figure 3: Initial Data

Being data quality a key issue with data mining as 50% to 80% of mining experts often spend their time on data quality, the pre-processing in data mining have a key importance[21]. In this case the collected data in Fig-3 was cleaned by using SQL Server Data Transformation Services, and then removed noise from the transformed data. The processed data are shown in the Table-1.

4. EXPERIMENTS AND RESULTS

For experimental analysis we use SPSS and XLMiner, and apply the clustering K-mean for the initial grouping of the whole data. It is clear from the result, that it gives three clusters shown in Tab-3, 4 and 5 of DS, SM and FM stock. So the process has two phases as given below.

4.1 Phase one

As discussed earlier in the first phase of our process, we have used K-Mean [13] clustering technique to classify the products in three groups, which gives accurate clusters of three categories according to the products quantity sold, show in Tab-2.

- Cluster 1 (Dead-Stock): This cluster contains record of those products which have small selling quantity. Every company has dead inventory (or at least sick and dying inventory). It's a natural outgrowth of being in business for any length of time as shown in Tab-3.

PUBLICATION OF Little Lion Scientific R&D, Islamabad PAKISTAN
Journal of Theoretical and Applied Information Technology
15th May 2011. Vol. 27 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1. Processed data

Itemid	color	season	size	company	gender	qty_sold	Dated
1	red	w	4	imp	m	25	1/1/2008
2	green	s	5	jack	f	2	2/2/2008
3	white	w	3	wikkey	m	4	4/3/2008
4	green	s	2	kips	f	6	1/1/2008
5	black	a	3	wikkey	m	7	3/3/2008
6	white	s	4	jack	f	8	1/1/2008
7	red	w	5	imp	m	9	2/2/2008
8	red	s	6	jack	m	0	4/3/2008
9	green	s	7	wikkey	m	0	1/1/2008
10	white	w	1	kips	m	2	3/3/2008
11	green	s	2	jack	m	4	1/1/2008
12	black	s	4	imp	f	6	2/2/2008
13	white	s	5	imp	f	8	4/3/2008
14	red	s	6	jack	f	9	1/1/2008
15	red	w	10	wikkey	m	5	3/3/2008
16	green	s	9	wikkey	m	4	1/1/2008
17	white	w	8	imp	m	6	2/2/2008
18	green	s	7	kips	m	7	4/3/2008
19	black	w	6	jack	m	0	1/1/2008
20	white	s	7	wikkey	m	2	3/3/2008
21	red	a	8	jack	m	33	1/1/2008
22	red	s	6	kips	f	5	2/2/2008
23	green	w	7	jack	f	6	4/3/2008
24	white	s	8	jack	f	7	1/1/2008
25	green	s	5	imp	f	8	3/3/2008
26	black	w	4	kips	f	9	1/1/2008
27	white	s	2	wikkey	f	0	2/2/2008
28	red	s	4	jack	f	2	4/3/2008

Table2: Cluster Results

Itemid	Color	company	gender	Cluster-ID	qty_sold
1	red	imp	m	3	25
2	green	jack	f	1	2
3	white	wikkey	m	1	4
4	green	kips	f	2	6
5	black	wikkey	m	2	7
6	white	jack	f	2	8
7	red	imp	m	2	9
8	red	jack	m	1	0
9	green	wikkey	m	1	0
10	white	kips	m	1	2
11	green	jack	m	1	4
12	black	imp	f	2	6
13	white	imp	f	2	8
14	red	jack	f	2	9
15	red	wikkey	m	2	5
16	green	wikkey	m	1	4
17	white	imp	m	2	6
18	green	kips	m	2	7
19	black	jack	m	1	0
20	white	wikkey	m	1	2
21	red	jack	m	3	33
22	red	kips	f	2	5
23	green	jack	f	2	6
24	white	jack	f	2	7
25	green	imp	f	2	8
26	black	kips	f	2	9
27	white	wikkey	f	1	0
28	red	jack	f	1	2

- Cluster 2 (Slow-Moving): This cluster contains records of those products which have medium selling quantity as shown in Tab-4.

- Cluster 3 (Fast-Moving): This cluster contains records of those products which have large selling quantity as shown in Table-5

4.2 Phase Two

In this phase our proposed algorithm MFP has been used to generate a property matrix shown in Table-6, containing counted values of corresponding properties of each product. This procedure receives data sets from clusters. The first loop scans all the records of the data set. The inner loop counts occurrences of the attribute for a given item and placed in the MFP matrix. Finally maximum occurrences of attributes values within a row give a single pattern. On the basis of these patterns, we can say that why a certain product falls in particular cluster.

PUBLICATION OF Little Lion Scientific R&D, Islamabad PAKISTAN
Journal of Theoretical and Applied Information Technology
15th May 2011. Vol. 27 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

From the first row of Table-6, it is clear that white coat of a female in winter has highest sale in terms of quantity sold. Similarly in the second row it can be seen that black shoes of female in summer has highest sale and so on.

Table3: Cluster1. Dead Stock

Itemid	color	seaso	size	company	gender	qty_sold	Dated
8	red	s	6	jack	m	0	4/3/2008
9	green	s	7	wikkey	m	0	1/1/2008
19	black	w	6	jack	m	0	1/1/2008
27	white	s	2	wikkey	f	0	2/2/2008
2	green	s	5	jack	f	0	2/2/2008
10	white	w	1	kips	m	0	3/3/2008
20	white	s	7	wikkey	m	0	3/3/2008
28	red	s	4	jack	f	0	4/3/2008

Table5: Cluster3. Fast-Moving

Itemid	color	season	size	company	gender	qty_sold	Dated
1	red	w	4	imp	m	20	1/1/2008
21	red	a	8	jack	m	33	1/1/2008

From the Fig-3 it is clear that there are three clusters for the sale patterns provided by using K-mean algorithms on the basis of the sale behavior. In Fig-4 the cluster numbers with data sets are shown, which clearly shows the number of items fall in the category.

Table4: Cluster2. Slow-Moving

Itemid	color	season	size	company	gender	q-sold	Dated
2	green	s	5	jack	f	2	2/2/2008
10	white	w	1	kips	m	2	3/3/2008
20	white	s	7	wikkey	m	2	3/3/2008
28	red	s	4	jack	f	2	4/3/2008
3	white	w	3	wikkey	m	4	4/3/2008
11	green	s	2	jack	m	4	1/1/2008
16	green	s	9	wikkey	m	4	1/1/2008
15	red	w	10	wikkey	m	5	3/3/2008
22	red	s	6	kips	f	5	2/2/2008
4	green	s	2	kips	f	6	1/1/2008
12	black	s	4	imp	f	6	2/2/2008
17	white	w	8	imp	m	6	2/2/2008
23	green	w	7	jack	f	6	4/3/2008
5	black	a	3	wikkey	m	7	3/3/2008
18	green	s	7	kips	m	7	4/3/2008
24	white	s	8	jack	f	7	1/1/2008
6	white	s	4	jack	f	8	1/1/2008
13	white	s	5	imp	f	8	4/3/2008
25	green	s	5	imp	f	8	3/3/2008
7	red	w	5	imp	m	9	2/2/2008
14	red	s	6	jack	f	9	1/1/2008
26	black	w	4	kips	f	9	1/1/2008

PUBLICATION OF Little Lion Scientific R&D, Islamabad PAKISTAN
Journal of Theoretical and Applied Information Technology
15th May 2011. Vol. 27 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 6 Results of Frequents Patterns (MFP Matrix)

MFP Matrix													
	Colors				Gender			Season					
Item	Red	White	Black	Max	Male	Female	Max	Winter	Summer	Spring	Autumn	Max	MFP
Coat	2	5	3	White	3	7	F	5	0	2	3	Winter	Coat-White-F-Winter
Shoes	3	4	5	Black	4	8	F	3	4	3	2	Summer	Shoes-Black-F-Summer
sweeter	5	6	7	white	10	8	M	5	2	6	5	Spring	Sweeter-White-M-Spring

5. CONCLUSION AND FUTURE WORK

In this paper, 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. From the experimental results it is clear that proposed approach is very efficient for mining patterns of huge stock data and predicting the factors affecting the sale of products.

We formulate most frequent pattern of products using their known properties in inventory system. We identified the trends of selling products through their known attributes. Our technique is simple by using matrix and counting of attribute values. The limitation of study is, that it requires proper data format with specific attributes. In future we will extend our work to implement in sentiment analysis process and decision making from online customer reviews and blogs data.

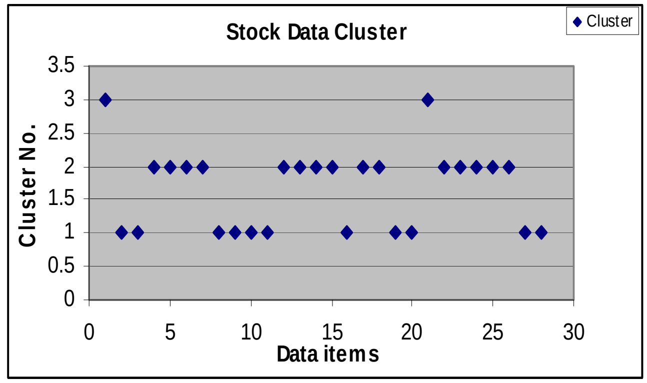

Figure3: Cluster Results

Figure 4: Cluster results with data set

REFRENCES:

[1]. Abubakar, Felix "Customer satisfaction with supermarket retail shopping", 2002.

[2]. http://www.roselladb.com/sales-trend-forecast.htm, visited January, 2010.

[3]. M. Braglia, A. Grassi, R. Montanari "Multi-attribute classification method for spare parts inventory management" 2004.

[4]. Association Analysis of Customer Services from the Enterprise Customer Management System- ICDM-2006.

[5]. Terry Harris,"Optimization creates lean green supply chains", 2008.

[6]. Matt Hartely "Using Data Mining to predict inventory levels" IEEE, 2005

[7]. Jiawan Han, Micheline Kamber "Data Mining Concepts and Techniques" 2nd edition 2004

PUBLICATION OF Little Lion Scientific R&D, Islamabad PAKISTAN
Journal of Theoretical and Applied Information Technology
15th May 2011. Vol. 27 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[8]. L. Frans, Wei, Paul, "Towards an agent based framework for online after sales services" 2006.

[9]. Gebouw D, Diepenbeek, Belgium "Building an Association Rules Framework to Improve Product Assortment Decisions" B-3590, 2004.

[10]. Brijs, Bart, Gilbert, Koen, Geert "A Data Mining Framework for Optimal Product Selection in Retail Supermarket Data: The Generalized PROFSET Model" 2000.

[11]. R. C. Wong, A. W. Fu, K. Wang"Data Mining for Inventory Item Selection with Cross-Selling Considerations" 2005.

[12]. L. Cao, C. Luo, J. Ni, DanLuo, C. Zhang "Stock Data Mining through Fuzzy Genetic Algorithm" Proceeding of JCIS s , 2008.

[13]. P.Thomas, Macredie "Knowledge Discovery and Data Mining" 1999 .

[14]. Artigan, J. A. Clustering Algorithms. Ohn Wiley and Sons, Inc., New York, NY. 1975.

[15]. Kennedy J., Eberhart R. C. and Shi Y., 2001. Swarm Intelligence, Morgan Kaufmann, NewYork.

[16]. Merwe V. D. and Engelbrecht, A. P., 2003. Data clustering using particle swarm optimization. Proceedings of IEEE Congress on Evolutionary Computation 2003(CEC 2003), Canbella, Australia.

[17]. Omran, M., Salman, A. and Engelbrecht, A. P., 2002. Image classification using particle swarm optimization. 2002.

[18]. http://en.wikipedia.org/wiki/cluster_analysis, visited 2009.

[19]. Jo Ting, "Mining of stock data: inter- and inter-stock patternassociative classification" proceddings of 2006 international conference on data mining Las Vegas,USA, June 2006.

[20]. Darken, C. Moody, J. Yale Comput. Sci., New Haven, " Fast adaptive k-means clustering" IEEE- 2002

[21]. http://maya.cs.depual.edu/~mobasher/ webminer/survey/node25.html visited 2008.

[22]. 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.

[23]. M.Cheng Lo, " Decission support system for the integrated inventory model with general distribution demand. Information technology journal 6(7) PP.1069-1074, 2007.

[24]. L.K. Soon and Sang Ho Lee, "Explorative Data Mining on Stock Data Experimental Results and Findings" pringer- ADMA 2007, LNAI 4632, pp. 562–569, 2007.

[25]. M. Al-Noukari, and W. Al-Hussan, "Using Data Mining Techniques for Predicting Future Car market Demand" IEEE, 2008

[26]. Han and Kamber "Data minin concepts and techniques" 2nd Edition , p-234 .

[27]. S. Kotsiantis, Kanellopoulos "Association Rules Mining : A Recent Overview" GESTS International Transactions on Computer Science and Engineering, Vol.32(1), 2006, pp 71- 82 – 2006.

[28]. Richard Ellis, Tony Allen and Miltos Petridis Application of Data Mining for Supply Chain Inventory Forecasting Proceedings of AI, the Twenty-seventh SGAI International Conference on Innovative Techniques and Applications of Artifcial Intelligence, pp.175-188, 2008.

Shu-Hsien Liao, Hsu-hui Ho, Hui-wen Lin, "Mining stock category, association and cluster on Taiwan stock market", Expert Systems with Applications Volume 35 , Issue 1-2 July 2008.