International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 03 Issue: 02 | Feb-2016 www.irjet.net p-ISSN: 2395-0072

Data Mining Based Store Layout Architecture for Supermarket

Aishwarya Madan Mirajkar [1], Aishwarya Prafulla Sankpal [2], Priyanka Shashikant Koli [3], Rupali Anandrao Patil [4], Ajit Ratnakar Pradnyavant[5].

[1] *Department of Computer Science and Engineering, Dr. J. J. Magdum College of Engineering, Jaysingpur, India.*
aishwaryamirajkar123@gmail.com

[2] *Department of Computer Science and Engineering, Dr. J. J. Magdum College of Engineering, Jaysingpur, India.*
aishwaryasankpal123@gmail.com

[3] *Department of Computer Science and Engineering, Dr. J. J. Magdum College of Engineering, Jaysingpur, India.*
priyankakoli664@gmail.com

[4] *Department of Computer Science and Engineering, Dr. J. J. Magdum College of Engineering, Jaysingpur, India.*
rupalipatil2907@gmail.com

[5] *Asst. Professor*
Department of Computer Science and Engineering, Dr. J. J. Magdum College of Engineering, Jaysingpur, India.
icon.ajit77@gmail.com

Abstract: *The mentioned system is designed to find the most frequent combinations of items. It is based on developing an efficient algorithm that outperforms the best available frequent pattern algorithms on a number of typical data sets. This will help in marketing and sales. The technique can be used to uncover interesting cross-sells and related products. The algorithms from association mining have been implemented and then best combination method is utilized to find more interesting results. The analyst then can perform the data mining and extraction and finally conclude the result and make appropriate decision.*

Market basket analysis is an important component of analytical system in retail organizations to determine the placement of goods, designing sales promotions for different segments of customers to improve customer satisfaction and hence the profit of the supermarket. These issues for a leading supermarket are addressed here using frequent itemset mining. The frequent itemsets are mined from the market basket database using the efficient Apriori algorithm and then the association rules are generated.

Key-Words: - Data Mining, Decision Support Systems, Association Rules, Market Basket Analysis, Apriori Algorithm , Store Layout .

1. INTRODUCTION

Data mining is an exciting and challenging field with the ability to solve many complex scientific and business problems. In recent years, the field of data mining has seen an explosion of interest from both academia and industry. Increasing volume of data, increasing awareness of inadequacy of human brain to process data and increasing affordability of machine learning are reasons of growing popularity of data mining. Data mining is a set of automated techniques used to extract buried or previously unknown pieces of information from large databases, using different criteria, which makes it possible to discover patterns and relationships. This new derived information can be utilized in the areas such as decision support, prediction, forecasting and estimation to make important business decisions, which can help in giving a particular business the competitive edge .Nowadays, it has been becoming critically important to make a decision that based on evidences than convictions of the authorities. Decision support system is a computer based information system designed to facilitate the decision making process of semi structured tasks. Central issue in DSS support is improvement of

decision making. Different technologies are invented to meet different decision making goals.

1.1 Data Mining and Decision Support

The extraction of hidden predictive information from large databases is a powerful tool with great potential to help organizations to define the information market needs of tomorrow. Data mining tools predict future trends and behaviors, allowing businesses to make knowledge-driven decisions that will affect the company, both short term and long term. The automated prospective analysis offered by data mining tools of today is much more effective than the analysis provided by tools in the past. Data mining answers business questions that traditionally were too time-consuming to resolve. Data mining tools search databases for hidden patterns, finding predictive information that experts may miss because it was outside their expectations. Data mining is not new. Although data mining is a relatively new term, the technology is Decision Support Systems (DSS) were praised for their great potential to supply executives with mountains of data needed to carry out their jobs. After1995s, corporate intranets were developed to support information exchange and knowledge management. The primary decision support tools in use included ad hoc query and reporting tools, optimization and simulation models, online analytical processing and data visualization. On the other hand, a data warehouse is the newest form of decision support system. Data mining was defined as one of the hottest technologies in decision support applications to date. Advances in data collection, the widespread use of bar codes for most commercial products, and the computerization of many business transactions have flooded us with information, and generated an urgent need for new techniques and tools that can intelligently and automatically assist us in transforming this data into useful knowledge. Today, there is a huge amount of information locked up in the mountains of data in companies' databases, information that is potentially important but has not yet discovered .Since almost all mid to large size retailers today possess electronic sales transaction systems, retailers realize that competitive advantage will no longer be achieved by the mere use of these systems for purposes of inventory management or facilitating customer checkout .In contrast, competitive advantage will be gained by those retailers who are able to extract the knowledge hidden in the data, generated by those systems, and use it to optimize their marketing decision making. In this context, knowledge about how customers are using the retail store is of critical importance and distinctive competencies will be built by those retailers who best succeed in extracting actionable knowledge from these data. Data Mining provides many different techniques to extract knowledge from data. It is an exciting multidisciplinary field of research which has many extremely useful applications. At present the techniques are becoming more commonly used but have not been applied adequately in the store layout. Store layout problem is motivated by applications known as market basket analysis to find relationships between items purchased by customers.

1.2 Store Layout Based on Knowledge

Store layout is an important retailing decision that can help to sales and store profitability. Store layouts are extremely important because they strongly influence in-store traffic patterns, shopping atmosphere, shopping behavior, and operational efficiency. Most retailers know that their retail store layout has great impact on their business. And because of this they contemplate changes to improve traffic flow, increase their shelf space, and increase their sales. The effects of retail store layouts are too big to overlook .The store layout has been well studied in both academic and practitioner literature. The store layout design is one of the more important determinants of store loyalty. Store layout design can play a key role not only in satisfying buyers 'requirements but also in influencing their wants and preferences. Store layout affects consumers' price acceptability, which is positively related to purchase intentions. They also report that superstores are currently revolutionizing the nature of retail service, mainly by creating more effective self-service arrangements as a result of improvements in store layout design .The routine store layout in the supermarkets based on the industrial logic implementation, which means putting products that share some functional characteristics or origins in the same area. So we will find product categories. Despite improvements, the store layout remains organized in product categories as defined by the manufacturers or category buyers. This approach is company oriented and it fails to respond to the needs of the time pressured consumer. Some retailers are trying to move from this organization to consumer oriented .This paper proposes a new store layout

approach based on the association rule mining. We assume that attractive store layout, navigational aids, sales people contact and in-store events induce transitions from recreational shopping to purchase-oriented shopping ,whereas retail crowding and time pressure engender shift from purchase-oriented shopping to recreational shopping. In addition, it is predicted that environmental design characteristics have greater impact on shopping path than on purchase decision marketing interventions exert more influence on purchase decision than on movement while contextual factors have comparable effect on shopping path and purchase decision. Retailers can utilize the proposed model to dynamically improve their in-store conversion rate.

1.3 Association Rules Mining

Association rules are if/then statements that help uncover relationships between seemingly unrelated data in a relational database or other information repository. An example of an association rule would be "If a customer buys a dozen eggs, he is 80% likely to also purchase milk. "An association rule has two parts, an antecedent (if) and a consequent (then). An antecedent is an item found in the data. A consequent is an item that is found in combination with the antecedent. Association rules are created by analyzing data for frequent if/then patterns and using the criteria support and confidence to identify the most important relationships. Support is an indication of how frequently the items appear in the database. Confidence indicates the number of times the if/then statements have been found to be true. In data mining, association rules are useful for analyzing and predicting customer behavior. They play an important part in shopping basket data analysis, product clustering, catalog design and store layout. Figure 1.1 to 1.7 shows steps 1 to 7 of Association Rule.

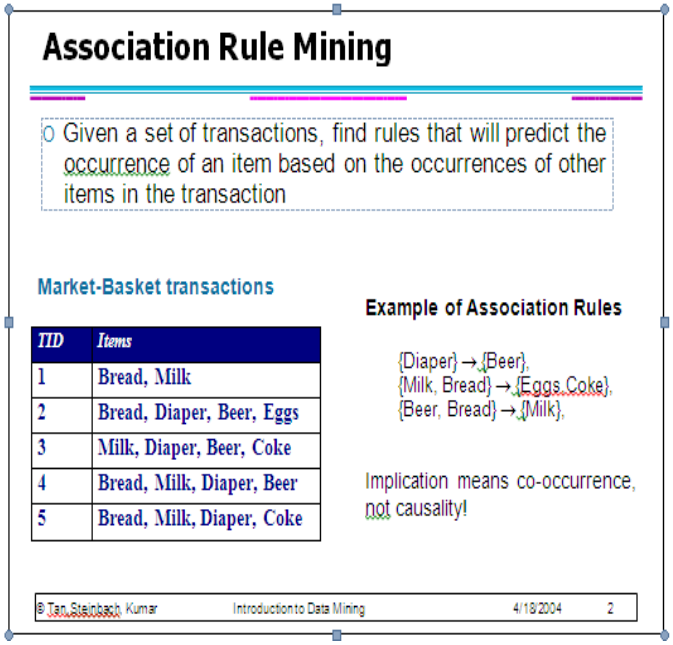

Fig.1.1 Association Rule Step1

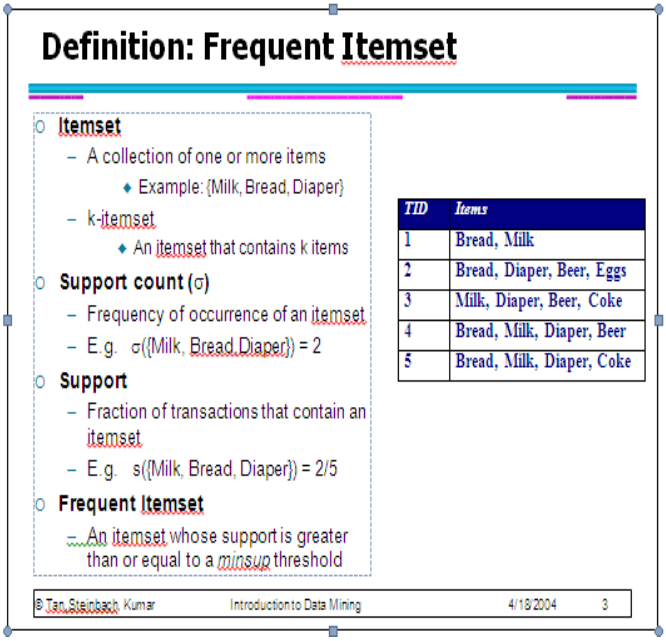

Fig.1.2 Association Rule step 2

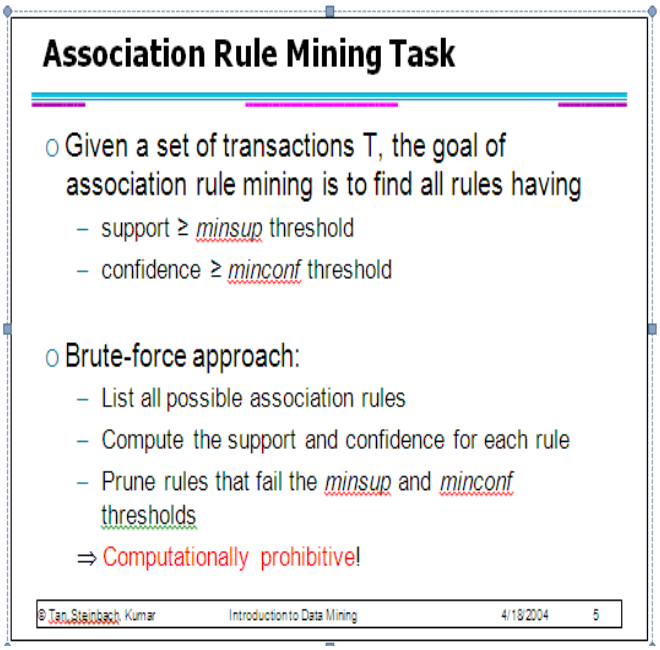

Fig.1.3 Association Rule step 3

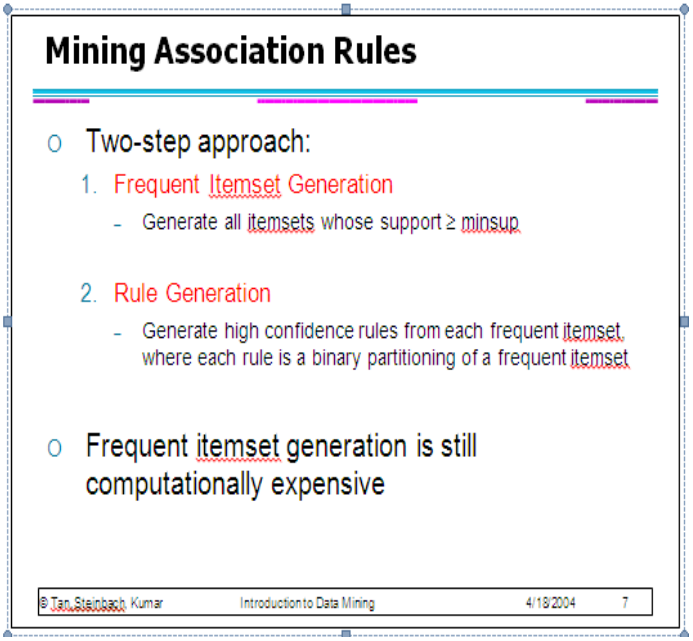

Fig.1.5 Association Rule step 5

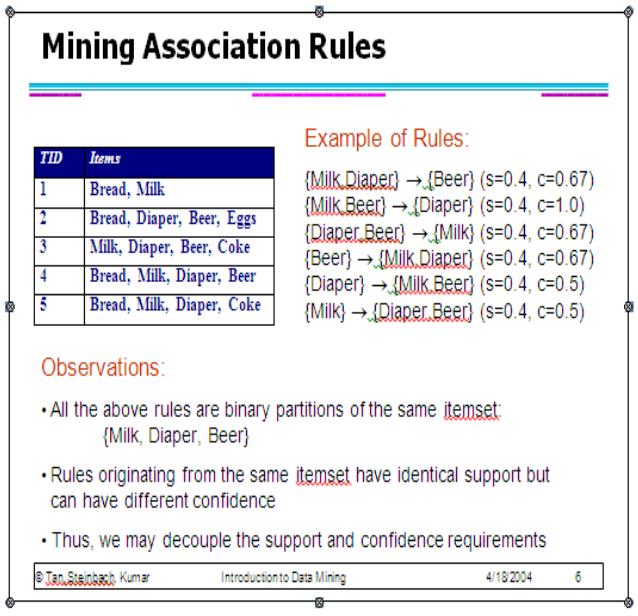

Fig. 1.4 Association Rule step 4

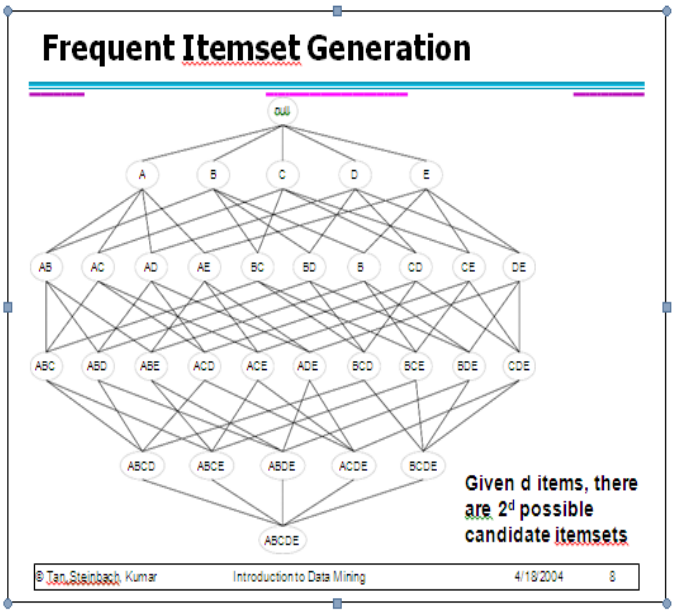

Fig .1.6 Association Rule step 6

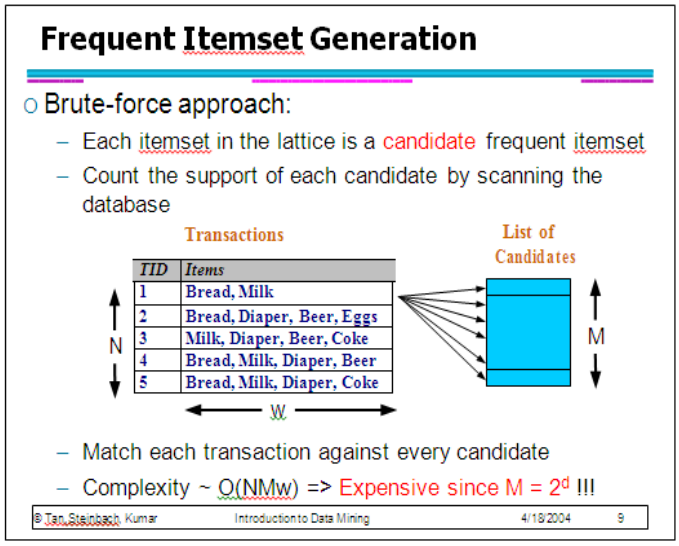

Fig.1.7 Association Rule step 7

2. PROPOSED SYSTEM

In our project contains following models

1) Data Input.

2) Market Basket Analysis.

3) Applying Association Rule.

4) Graphical representation.

2.1 Data input

In data input model we are using database of the supermarket .In that we are storing information about categories of the products, price of the products, quantity of products and names of the products .

2.2 Market basket understanding

Market Basket Analysis is used to determine which products sell together, the input data to a Market Basket Analysis is normally a list of sales transactions, where each has two dimensions, one represents a product and the other represents a customer, depending on whether the goal of the analysis is to find which items sell together to the same person. Apriori algorithm is one of the most widely used and famous techniques for finding association rules. This algorithm was chosen primarily due to the speed of application. Simulation models, online analytical processing and data visualization. On the other hand, a data warehouse is the newest form of decision support system.

2.3 Applying Association Rule

The applications of data mining in retail trade enterprises are mainly concentrated in association rules mining. Association rule mining is an initial data exploration approach that is often applied to extremely large data set. An example is grocery store market basket data. Association rules mining provides valuable information in assessing significant correlations. By mining association rules, marketing analysts try to find sets of products that are frequently bought together. They have been applied to a variety of fields. Market Basket Analysis is used to determine which products sell together, the input data to a Market Basket analysis .We use the buying association measure to create a category correlation matrix and we apply the multi-dimensional scale technique to display the set of products in the store space.

2.4 Graphical representation

Data modeling is where the data mining software is used to generate results for various situations. The data processing in Clementine is done through the use of nodes, which are then connected together to form a stream frame. In addition, data visualization can be presented to users after the mining process has been done in the graphical format.

3. METHODOLOGY

3.1 Proposed Methodology

This system develops a relational database and uses Apriori algorithm techniques as methodologies for the store layout. Knowledge extraction by association rule mining results is illustrated as knowledge patterns/rules and clusters in order to propose suggestions and solutions to the case firm for store layout.

First of all, we need to measure the relationship among products million transactions we required as two text file. The first file contains transactions during the period. Each transaction has

the date, cache no, receipt no, and barcode. The second database file contains product data. The first step toward a data mining is to transform text files into SQL Server. This study established relational data base tables and transferred them on MS SQL Server within OBDC environment in order to implement the data table. The text formatted data of transaction records was loaded into a relational database for querying. In the product data file there are barcode, product name, and group item code. We have chosen all the categories, which are 16 different categories, to construct correlation matrix.

3.1.1Apriori Algorithm

Apriori is designed to operate on databases containing transactions (for example, collections of items bought by customers, or details of a website frequentation).Each transaction is seen as a set of items (an *itemset*). Given a threshold, the Apriori algorithm identifies the item sets which are subsets of at least transactions in the database.

Apriori uses a "bottom up" approach, where frequent subsets are extended one item at a time (a step known as candidate generation), and groups of candidates are tested against the data. The algorithm terminates when no further successful extensions are found. Apriori uses breadth-first search and a Hash tree structure to count candidate item sets efficiently

3.2 Design Approaches

3.2.1 DFD diagram

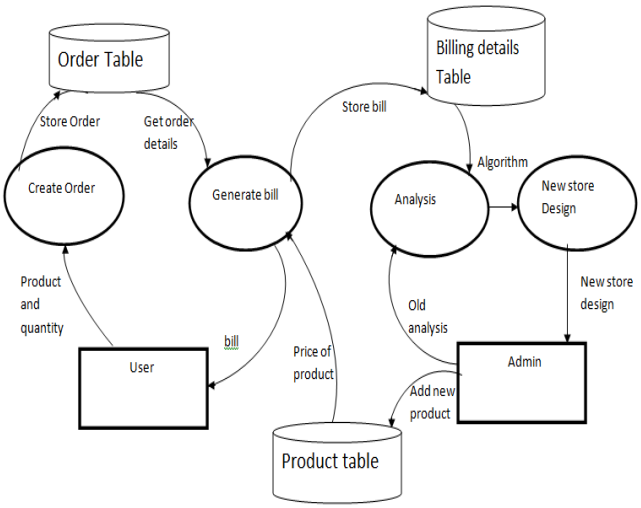

Figure 3.2.1 DFD Level 2

This DFD show the flow of data between the process. Billing detail table stores the bill of customer and in this project we are using Apriori algorithm on that table data .The result of the analysis is use to the design the store layout for the supermarket.

3.2.2 Use Case Diagram:

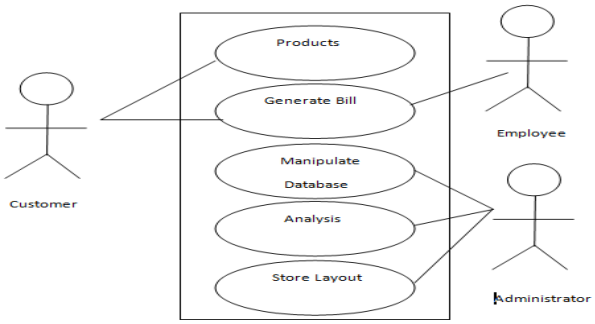

Figure 3.2.2 Use Case Diagram

4. CONCLUSION

This paper discuses association rules for data mining extract knowledge from a database and a new supermarket store layout based on the association among categories. This approach allows supermarkets to cluster products around meaningful purchase opportunities related to use association.

REFERENCES

[1] Atkin, Charles K., "Observation of Parent Child Interaction in Supermarket Decision- Making," Journal of Marketing, October, 1978. February
[2] M. Hahsler, K. Hornik, "Building on the arules infrastructure for analyzing transaction data with R", in R. Decker, H.-J. Lenz (eds.), Advances in Data Analysis, pp. 449-456, Berlin: Springer, 2007.
[3] J. Han, M. Kamber, Data Mining: Concepts and Techniques, p. cm. London: Academic Press, 2001.
[4] Chen, Y.-L., Tang, K., Shen, R.-J., Hu, Y.-H.: Market basket analysis in a multiple store environment, Decision Support Systems, 2004.
[5] Berry, M.J.A., Linoff, G.S.: Data Mining Techniques: for Marketing, Sales and Customer Relationship Management (second edition), Hungry Minds Inc., 2004.
[6] Agrawal R, Srikant R, Fast algorithms for mining association rules. In: Proceedings of the 20th VLDB conference,1994, pp 487–499.