ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 1, January 2019

Application of Data Mining for understanding Storefront Operations in Retailing

Ashish B. Patel[1]

Research Scholar, Faculty of Computer Science, Pacific Academy of Higher Education and Research University, Udaipur, Rajasthan, India[1]

ABSTRACT: Data mining refers to using a variety of techniques to identify nuggets of information or knowledge in data, and extracting these for decision support, prediction, forecasting and estimation. The data is often voluminous, but as it stands of low value as no direct use can be made of it; it is the hidden information in the data that is useful. The essence of data mining usage in business applications is that it is useful in identifying relevant information and knowledge which can be used for better decision making in business area. This paper studies the use of data mining, as a tool, in understanding storefront operations in retail marketing which also happens to generate huge amount of data in its operations. Retailers can make use of this data in order to understand store segmentation, market basket analysis, category management and out of stock analysis.

KEYWORDS: Data Mining, Retail Marketing, Storefront Operations

I. INTRODUCTION

Data Mining is the process of discovering interesting patterns and knowledge from large amount of data. The data sources can include databases, data warehouses, the web, other information repositories, or data that are streamed into the system dynamically [J. Han and M. Kamber, J Pei (2014].

Data mining is commonly seen as a single step of a whole process called Knowledge Discovery in Databases (Figure 1). (KDD). According to Fayyad et.al, “KDD is the nontrivial process of identifying valid, novel, potentially useful and ultimately understandable patterns in data”. [Fayad et. al].

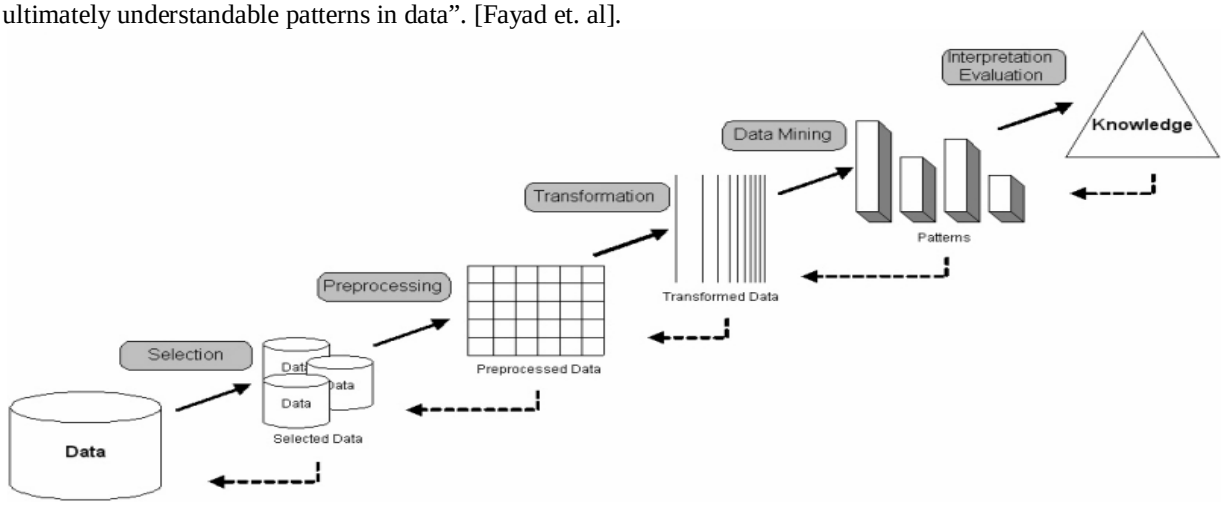

Figure 1: *Data Mining and the KDD Process.*

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 1, January 2019

As shown in the figure 1, KDD process consists of iterative sequence of following steps;

1. Data: It requires cleaning and integration of the collected data in order to remove noise and inconsistent data along with combining multiple sources of data.
2. Data Selection: selecting data relevant to the analysis task from the database.
3. Data Transformation: transforming data into appropriate forms to perform data mining.
4. Data mining: an essential process where intelligent methods are applied to extract data patterns.
5. Pattern evaluation: interpreting the patterns into knowledge by removing redundant or irrelevant patterns; translating the useful patterns into terms that human understandable.
6. Knowledge presentation: where visualization and knowledge representation techniques are used to present mined knowledge to users.

Data mining gained popularity especially in the last two decades when advances in computing power provided us with the possibility to mine voluminous data. Extracting knowledge and hidden information from data using a whole set of techniques found its applications in various contexts. Knowledge discovery is widely used in marketing to identify and analyse customer groups and predict future behaviour. Data mining is an effective way to provide better service to customers and adjust offers according to their needs and motivations.

Data mining Applications:

Here is the suggestive list of areas where data mining is widely used: Financial data analysis, retail marketing, telecommunication industry, biological data analysis, scientific applications, intrusion detection, data mining applications in transportation, data mining applications in medicine, data mining applications in health care and insurance, visual and audio data mining.

II. RELATED WORK

DATA MINING AND RETAIL MARKETING

Retailing includes all activities involved in selling goods or services directly to the final consumer for personal or non-business use [Kotler, Kevin Lane Keller, Abraham Kosly, Mithileshwar Jha (2013)]. Any organisation that directs its marketing efforts towards satisfying the final consumer in selling goods and services as a means of distribution is carrying out retailing function [David Gilbert (2003)]. In other words, retailing is the sale of goods and services to the ultimate consumer for personal, family or household use. Retailing involves selling of tangible as well as intangible goods. Retailers are facing dynamic and competitive environment. With increase in globalization and competitiveness, retailers are seeking better market campaign. Retailers are collecting large amount of data. This data collected requires proper mechanisms to convert it into knowledge, using this knowledge retailer can make better business decision. Retail industry is looking strategy where in they can target right customers who may be profitable to them.

Data mining tools perform analyses that are very valuable for business strategies, scientific research and getting to know your customers better. Managerial insights are no longer the only factor trusted when it comes to decision-making. Data driven decisions can lead to better firm performance. Data-based implications are gaining popularity while the gut instinct of managers is remaining in the background. Analysing data not only improves firm performance but gives us accurate insights on different aspects of the business.

Data mining is widely used in marketing for spotting sales trends, developing better marketing campaigns and finding the root cause of specific problems like customer defection or fraudulent transactions, for example. It is also used for prediction of behaviour: which customers are most likely to leave us (customer churns) or what are the things that an individual will be most interested to see in a website.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 1, January 2019

Some of the Retail Applications of Data Mining are in following areas:

Customer behaviour and customer relationship management, supply chain management and procurement, storefront operations, alternative sales analysis.

In this paper, we focus on storefront operations

Store Segmentation

Retail Marketers and analysts use segmentation to organize and manage market and customer data. A segment is built by analyzing data and grouping (clustering) elements that share sought characteristics (in the case of retail, demographic, expenditure, lifestyle and media preference information). Segmentation is essential for cost effective, accurately targeted direct marketing.

By organizing customer and market data in this way, organizations are able to understand the differences and similarities between groups of customers and prospects, and develop effective messaging, products and distribution channels appropriate to the specific needs and wants of a given segment.

Market Basket Analysis

Market Basket Analysis is a modelling technique based upon the theory that if you buy a certain group of items, you are more (or less) likely to buy another group of items. Market basket analysis can be used in deciding the location and promotion of goods inside a store.

Category Management

Category management is a retailing and purchasing concept in which the range of products purchased by a business organization or sold by a retailer is broken down into discrete groups of similar or related products; these groups are known as product categories (examples of grocery categories might be: tinned fish, washing detergent, toothpastes). It is a systematic, disciplined approach to managing a product category as a strategic business unit [27].

Out of Stock Analysis

Out of stock analysis can help supply and demand planners and sales directors to analyze out-of-stock situations at retailer locations. Out-of-stock data is aggregated at product level. This analysis also helps to find the cumulative net sales values that retailers have lost due to out-of-stock situations.

III. LITERATURE REVIEW

Marketing literature provides many examples of market segmentation research and numerous bases for segmentation have been proposed (Mangaraj and Senuer, 2001; Lake, 2007; McKinsey et al. 2000; Green and Krieger, 1991; Grover and Srinivasan, 1987; Kamakura and Russell, 1989). One technique commonly used in domestic market segmentation is cluster analysis. Cluster analysis groups objects by minimizing the within group differences and maximizing between group differences. Cluster analysis is often based on consumer attitude towards the products, perceived benefits, purchase propensities, lifestyle, or demographics (Punj and Stewart, 1983).

Association rule mining is a data mining task that identifies relationships among items in a transactional database. Association rules have been widely investigated in the literature for their role in several application domains such as Market Basket Analysis, recommender systems, diagnosis decisions support, telecommunication, intrusion detection, etc. The competent discovery of such rules has been a key focus in the data mining research community. The standard

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 1, January 2019

apriori algorithm has been modified for the improvements of association rule mining algorithms [Pei-ji Wang et al., 2009]. Weiyang Lin et al., (2004) described an efficient Adaptive-Support Association Rule Mining for Recommender Systems. The author examined the usage of association rule mining as a fundamental technique for collaborative recommender systems. Association rules have been used with sensation in other domains. Nevertheless, most currently existing association rule mining algorithms were designed with market basket analysis in mind.

Shyue-Liang Wang et al., (2003) proposed an effective data-mining approach for discovering Adaptive-Support Association Rules from databases. Adaptive-support association rules are constrained association rules with application to collaborative recommendation systems. To find out association rules for recommendation systems, a particular value of target item in association rules is normally assumed and no minimum support is specified in advance. Min Wang et al., (2011) developed a technique for mining multi-dimension association rule depending on the Adaptive Genetic Algorithm with crossover matrix and mutation matrix. In this association rule mining system, selection, mutation, and crossover are all parameter-free in evolution process. Results show that: combined with the adaptive genetic algorithm, the precision and efficiency of mining association rules is improved.

Taboada et al., (2006) proposed a method of association rule mining using genetic network programming (GNP) with a self-adaptation mechanism in order to improve the performance of association rule extraction systems. GNP is a kind of evolutionary methods, whose directed graphs are evolved to find a solution as individuals. Self-adaptation behavior in GNP is related to adjust the setting of control parameters such as crossover and mutation rates. It is called self-adaptive because the algorithm controls the setting of these parameters itself - embedding them into an individual's genome and evolving them. The aim is not only to find suitable adjustments but to do this efficiently. Our method can measure the significance of the association via the chi-squared test and obtain a sufficient number of important association rules. Extracted association rules are stored in a pool all together through generations and reflected in three genetic operators as acquired information. Further, our method can contain negation of attributes in association rules and suit association rule mining from dense databases.

Kahn et al. recognize that our model does not explicitly account for other possible factors that influence the relationship between assortment variety and demand: the space devoted to a category and the presence or absence of a favourite item influence the perception of variety (Kahn and Lehmann 1991, Broniarczyk et al. 1998) as well as the arrangement, complexity, and presence of repeated items in an assortment (Hoch et al. 1999, Huffman and Kahn 1998, Simonson 1999). Assortment planning has attracted researchers from both operations and marketing fields. See Kok et al. (2006) for a recent review of this literature. van Ryzin and Mahajan (1999), Smith and Agrawal (2000), and Kok and Fisher (2004) study assortment selection and stocking decisions for a group of substitutable products in a single category assuming that store traffic is exogenous. Agrawal and Smith (2003) extend this work to the case with basket shopping consumers. Cachon et al. (2005) partially relaxes the exogenous store traffic assumption by considering consumer search behavior. The customers can choose to purchase an item at the store or to continue to search, which means that the fraction of “no-purchase” customers depends on the assortment. Chong et al. (2001) present an empirically-based modelling framework for managers to assess the revenue and lost sales implication of alternative assortments. Hopp and Xu (2005) study the impact of product modularity in the optimal product line length from a manufacturer’s perspective. Hopp and Xu (2006) study price, service, and assortment competition in a single category between two retailers and find that the retailers provide less variety and lower prices in competition.

IV. CONCLUSION

Data mining is the process of analyzing data from different angle or perspective and collecting it to get useful information that can be used to increase revenue costs or both, data mining allows backend processors to analyze data from many different dimensions, categories it & summarize the relationships identified.

In retail business, the goal of storefront operation is to help decision makers for better organizing customer and market data. In this way, organizations are able to understand the differences and similarities between groups of customers and

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 1, January 2019

prospects, and develop effective messaging, products and distribution channels appropriate to the specific needs and wants of a given segment. Market basket analysis can be used in deciding the location and promotion of goods inside a store. Category management is disciplined approach to managing a product category as a strategic business unit. Out of stock analysis can help supply and demand planners and sales directors to analyze out-of-stock situations at retailer locations.

REFERENCES

[1] Analysis." The Journal of Marketing. Vol. 55(4):20-31.
[2] Cachon, G. P., C. Terwiesch, Y. Xu. 2005. Retail assortment planning in the presence of consumer search. Manufacturing Service Oper. Management 7(4) 330–346.
[3] Chong, J. K., T.-H. Ho, C. S. Tang. 2001. A modeling framework for category assortment planning. Manufacturing Service Oper. Management 3(3) 191–210.
[4] David Gilbert (2003), "Retail Marketing Management", Pearson Education, Indian Reprint.
[5] Green, Paul E. and Abba M. Krieger. (1991). "Segmenting Markets with Conjoint
[6] Grover, Rajiv and V. Srinivasan. (1987). "A Simultaneous Approach to Market Segmentation and Market Structuring." Journal of Marketing Research. Vol. 24: 139-53. Hoch, S. J., E. T. Bradlow, B.Wansink. 1999. The variety of an assortment. Marketing Sci. 18(4) 527–546.
[7] Hopp, W. J., X. Xu. 2006. A static approximation for dynamic demand substitution with applications in a competitive market. Working paper, Rutgers University, New Brunswick, NJ.
[8] Huffman, C., B. E. Kahn. 1998. Variety for sale: Mass customization or mass confusion? J. Retailing 74 491–513.
[9] J. Han and M. Kamber, J Pei (2014), "Data Mining: Concepts and Techniques," Morgan Kaufmann.
[10] K. Taboada, K. Shimada, S. Mabu, K. Hirasawa & J. Hu (2006), "Self-Adaptive Mechanism in Genetic Network Programming for Mining Association Rules", International Joint Conference SICE-ICASE, Pp. 6007–6012.
[11] Kahn, B. E., D. R. Lehmann. 1991. Modeling choice among assortment. J. Retailing 67 274–275.
[12] Kamakura, Wagner A. and Gary J. Russell. (1989). "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure." Journal of Marketing Research. Vol. 26:379- 90.
[13] Kok, A. G., M. L. Fisher, R. Vaidyanathan. 2006. Assortment planning: Review of literature and industry practice. N. Agrawal, S. A. Smith, eds. Invited chapter to appear in Retail Supply Chain Management. Kluwer, Amsterdam, The Netherlands.
[14] Lake, Laura. 2007. Market Segmentation for the Small Business. Accessed Jan. 16, 2015 at: http://marketing.about.com/cs/sbmarketing/a/smbizmrktseg.htm
[15] Mangaraj, Sand and Senauer, Ben. 2001. A Segmentation Analysis of U.S. Grocery Store Shoppers. Department of Applied Economics, University of Minnesota, St. Paul, MN.
[16] Mckinsey, James H. Segmentation and Positioning for Strategic Marketing Decisions. Chicago, America Marketing Association, 2000.
[17] Min Wang, Qin Zou & Caihui Liu (2011), "Multi-dimension Association Rule Mining based on Adaptive Genetic Algorithm", International Conference on Uncertainty Reasoning and Knowledge Engineering, Vol. 2, Pp. 150–153.
[18] Pei-ji Wang, Lin Shi, Jin-niu Bai & Yu-lin Zhao (2009), "Mining Association Rules based on Apriori Algorithm and Application", International Forum on Computer Science-Technology and Applications, Vol. 1, Pp. 141–143.
[19] Philip Kotler, Kevin Lane Keller, Abraham Kosly, Mithileshwar Jha (2013), "Marketing Management": Pearson India, ISBN 978-81-317-6716-0.
[20] Punj, Girish and David W. Stewart. (1983). "Cluster Analysis in Marketing Research: Review and Suggestions for Application." Journal of Marketing Research. Vol. 20:134 -48.
[21] Shyue-Liang Wang, Mei-Hwa Wang, Wen-Yang Lin & Tzung-Pei Hong (2003), "Efficient Generation of Adaptive-Support Association Rules", IEEE International Conference on Systems, Man and Cybernetics, Vol. 1, Pp. 894–899.
[22] Simonson, I. 1999. The effect of product assortment on buyer preferences. J. Retailing 75 347–370.
[23] Smith, S. A., N. Agrawal. 2000. Management of multi-item retail inventory systems with demand substitution. Oper. Res. 48(1) 50–64.[24] Usama Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth. (1996). From Data Mining to Knowledge Discovery in Databases. American Association for Artificial Intelligence.
[25] Van Ryzin, G., S. Mahajan. 1999. On the relationship between inventory costs and variety benefits in retail assortments. Management Sci. 45(11) 1496–1509.
[26] Weiyang Lin, Sergio A. Alvarez & Carolina Ruiz (2004), "Efficient Adaptive-Support Association Rule Mining for Recommender Systems", Journal on Data Mining and Knowledge Discovery, Vol. 6, No. 1, Pp. 83–105.
[27] "Industry Resources: Glossary". Nacsonline.com. Retrieved on December 16, 2018.