

ROLE OF DATA MINING IN RETAIL SECTOR

Bharati M. Ramageri

Assistant Professor, PES's Institute of Management and Career Development (MCA),
Pune, Maharashtra, India-411044
bharati_aramani@rediffmail.com

Dr. B.L. Desai

Project Manager, Capgemini Pvt. Ltd, MIDC, Talawade,
Pune, Maharashtra, India-412114
drblidesai@yahoo.com

Abstract— Data mining is proved to be one of the important tools for identifying useful information from very large amount of data bases in almost all the industries. Industries are using data mining to increase revenues and reduce costs. This article begins the concept of data mining that has emerged as a technique of discovering patterns to make better strategies and decisions. It also discusses standard tasks involved in data mining, reviews various data mining applications in different sectors. This paper attempts, how data mining can be applied in retail industry to improve market campaign.

Keywords- Data Mining Applications Review, Retail Industry, Market Campaign

I. INTRODUCTION

Today retailer is facing dynamic and competitive environment, with increase in globalization and competitiveness retailers are seeking better market campaign [24]. Retailer are collecting large amount of customer dialy transaction details. 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 is the extraction of hidden predictive information from very large databases. It is a powerful technology with great potential to help organizations focus on the most important information in their data warehouses [2,9,13]. Data mining tools predict future trends and behaviours, helps organizations to make proactive knowledge-driven decisions [7,15]. The automated, prospective analyses offered by data mining move beyond the analyses of past events provided by retrospective tools typical of decision support systems. Data mining tools can answer the questions that traditionally were too time consuming to resolve. They prepare databases for finding hidden patterns, finding predictive information that experts may miss because it lies outside their expectations.

From the last decade data mining have got a rich focus due to its significance in decision making and it has become an essential component in various industries [1,7,15]. The field of data mining have been prospered and posed into new areas such as manufacturing, insurance, medicine, retail etc [1,2,9,13]. Hence, this paper reviews the various trends of data mining and its relative applications from past to present and discusses how effectively can be used for targeting profitable customers in campaigns.

II. DATA MINING DEFINITION

Berry and Linoff (2000) [2] says data mining is “The process of exploration and analysis, by automatic or semiautomatic means, of large quantities of data in order to discover meaningful patterns and rules.”

III. DATA MINING PROCESS

The life cycle of a data mining project consists of six phases [23]. Moving back and forth between different phases is always required. It depends on the outcome of each phase. The main phases are as follows,

a) Business Understanding: This phase focuses on understanding the project objectives and requirements from a business perspective, then converting this knowledge into a data mining problem definition and a preliminary plan designed to achieve the objectives.

b) Data Understanding: It starts with an initial data collection, to get familiar with the data, to identify data quality problems, to discover first insights into the data or to detect interesting subsets to form hypotheses for hidden information.

- c) Data Preparation: It covers all activities to construct the final dataset from the initial raw data.
- d) Modeling: In this phase, various modeling techniques are selected and applied and their parameters are calibrated to optimal values.
- e) Evaluation: In this stage the model is thoroughly evaluated and reviewed. The steps executed to construct the model to be certain it properly achieves the business objectives. At the end of this phase, a decision on the use of the data mining results should be reached.
- f) Deployment: The purpose of the model is to increase knowledge of the data, the knowledge gained will need to be organized and presented in a way that the customer can use it. The deployment phase can be as simple as generating a report or as complex as implementing a repeatable data mining process across the enterprise.

IV. APPLICATION OF DATA MINING

Data mining can be implemented in many fields depending on the aim of the business. In today's business world, many industries can apply data mining technology such as [7, 9, 13] manufacturing, banking, healthcare, insurance, telecom, medicine, retail etc.

4.1 Manufacturing

Data Mining can be applied in manufacturing industries for manufacturing system modeling, manufacturing process control, safety evaluation, Process planning and scheduling and so on [15].

4.2 Banks and financial sector

Data mining in bank or any financial sector for loan payment analysis, customer credit policy analysis, target marketing, detection of money laundering and other financial crimes can be applied [13].

4.3 HealthCare

Data mining can help healthcare insurers to detect fraud and abuse, healthcare organizations make customer relationship management decisions, physicians identify effective treatments and best practices, and patients receive better and more affordable healthcare services [7].

4.4 Computer Science and Engineering

Data mining in computer science and engineering can be used to help monitor system status, improve system performance, isolate software bugs, detect software plagiarism, analyze computer system faults, uncover network intrusions and recognize system malfunctions [13].

4.5 Software and system Engineering

Data mining for software and system engineering can operate on static and dynamic data depending on whether the system dumps trace beforehand for post analysis or if it must react in real time to handle online data.

4.6 Telecommunication

Telecommunication can apply data mining for customer retention, fraud analysis, and churn management. It helps to identify telecommunication patterns, catch fraudulent activities, intrusion detection, make better use of resources and improve service quality.

4.7 Retail Sector

One can also apply data mining for the following purposes,

4.7.1 Acquiring and Retaining Customer

It is more costly to reach new customers than to get existing one [24]. So by knowing existing customers' purchase behaviour, direct marketer can predict customers need and interest in buying particular product. Using this type of prediction retailer can retain existing customers by providing discounts or offer, attract customers and acquire customers [12].

4.7.2 Market Basket Analysis

Market basket analysis is a technique in understanding what items are likely to be purchased together according to association rule [6]. It provides valuable indications about customers, shopping patterns by showing associations among various items. This type of item association is useful for shelf design, deciding the location and promotion of items by means of combination. So that customers can easily locate item and this analysis helps in product cross-selling.

4.7.3 Customer Segmentation and Target Marketing

Segmentation is to divide the market into several parts by certain characters. Data mining can be used in grouping or clustering customers based on the behavior [26]. This type of information is useful to define similar customers in a cluster, holding on good customers and identify likely responders for target marketing.

Many industries have already applied data mining technology and gaining advantages over its competitors [4].

Data mining in manufacturing industry [1] gives deeper insight into the processes allows for the prediction of machine failure and for preventive maintenance. So data mining supported to improve quality and reduce costs due to damages or loss of prediction.

Data mining to analyze complex and difficult medical data, REMIND (Reliable Extraction and Meaningful Inference from Non-structured Data) system [20] integrates the structured and unstructured clinical data in patient records to automatically create high quality structured clinical data. The high quality of structuring allows existing patient records to be mined to support guidelines compliance and to improve patient care [20].

[5] Successful application in healthcare arena where data mining implemented includes health infection control, ranking hospital, identifying high risk patients.

A surveillance system was developed using data mining technology to identify new and interesting patterns in infection control data was implemented at University of Alabama.

Data mining technology have been implemented to examine reporting practice. With the use of international classification and diseases, 9th revision, codes (risk factors) and by reconstructing patients profiles, cluster and association analysis, shows how risk factors are reported.

A robust data mining and model building solution identifies patients who are trending towards a high risk condition. This information gives nurse care coordinators a head start in identifying high risk patients so that steps can be taken to improve the patient's quality of healthcare to prevent health related problems in the future.

Data mining facilitates software maintenance engineers to comprehend the structure of a software system and assess its maintainability [14]. The clustering algorithm effectively used to produce overviews of systems by creating mutually exclusive groups of classes, member data or methods, according to their similarities and hence reduces the time required to understand the overall system.

A number of adaptive supervised learning networks have been developed for traders and portfolio management [16]. LBS Capital Management Inc.1 , a fund-management firm, uses neural networks, genetic algorithms, and expert systems to manage portfolios worth US\$600 million. The system has outperformed the overall stock market.

HNC Software [17] uses a BP neural network with three layers in its system for detecting credit card fraud. The input layer receives information from an external source, the hidden layer uncovers common features of the information that the input layer provides, and the output layer learns combinations of the features that are associated with known behaviors or decisions.

Many industries have already applied data mining technology such as Wal-mart, proctor and Gamble, Coke Pepsi and American Greetings have been using data mining for category management for several years [18].

J Crew Group Inc wanted to determine what clothes, shoes and accessories customers most purchase together, so they used market basked analysis combining click stream analysis from its website along with point of sale (POS) data from retail locations to perform product affinity analysis. The data was then used to make complementary product suggestions for online shoppers [25] (Whiting 2001).

Customer relationship management (CRM) and customer profiles: Federated department stores are combining customer and transaction data to identify best customers and offer exclusive extras. For example Macys West by identifying customer spending habits can predict life changing events such as marriage and children within one year [11] (Chain store age 1999).

ZCMI a department store chain based in Utah is using data mining to integrate customer data with multiple other merchandising systems to identify popular products and specific customer categories [19] (Rabinovitch 1999).

Neiman Marcus is using data mining technology to build and enhance customer relationship.

We found research using data mining technology in retail industry to target customers for market campaign is very much limited [8]. Most retailers have difficulty in identifying the right customers to engage in successful

campaigns. So far, customer segmentation is a popular method that is used for selecting appropriate customers for a launch campaign [3,12,26].

This investigation identifies customer behaviour using a recency, frequency and monetary (RFM) model and then uses a customer life time value (LTV) model to evaluate proposed segmented customers[3].

Identification of profitable customers for market campaign or new product launch is a present issue. As the existing research in retail is limited to targeting customers based on RFM and LTV model [3], reward point based target for market campaign is being suggestion here. Suppose customers are given a card, and as and when they make certain purchases reward points will be added up. Based on customer RFM, and behaviour using a LTV and reward point details, retailer can target potentially profitable customers for campaigns. If this type methodology is applied retailer can easily find potentially profitable customer for campaign programs we can further correlate between customer values and campaigns.

IBM's SPSS, SAS's Analytics, Microsoft's XLSTAT, Megaputer's PolyAnalyst 6.0 etc., are the tools which can be used for such applications.

V. CONCLUSIONS

In this paper an attempt is made to define data mining as a tool used to extract important information from large databases so that various sectors can make better business decisions. A wide range of industries have deployed successful applications of data mining. Data mining in retail industry can be deployed for market campaigns, to target profitable customers using reward based points. The retail industry will gain, sustain and will be more successful in this competitive market if adopted data mining technology for market campaigns.

REFERENCES

- [1] Andre Bergmann, "Data Mining for Manufacturing: Preventive Maintenance, Failure Prediction, and Quality Control".
- [2] Berry, M. J. A. and Linoff, G. Data mining techniques for marketing, sales and customer support, USA: John Wiley and Sons,1997
- [3] Chu Chai Henry Chan, "Intelligent value based customer segmentation method for campaign management: A case study of automobile retailer", science Direct expert system with applications 34, (2008) 2754-2762.
- [4] Customer Successes in your industry from http://www.spss.com/success/?_source=homepage&hpzone=nav_bar, retrieved on Feb 2011.
- [5] David Biranbaum, "Application of Data mining Techniques to Healthcare data", Infection control and hospital epidemiology, statistics for hospital epidemiology, august 2004.
- [6] Dhanabhakayam and Dr. M Punithavalli, "A Survey on Data mining Algorithm for Market Basket Analysis": Global Journal of Computer Science and Technology" Vol-11 Issue -1 ver. 1.0 July 2011.ISSN: 0975-4350.
- [7] Dr. Gary Parker, vol 7, 2004, Data Mining: Modules in emerging fields, CD-ROM.
- [8] E.W.T Ngai,LI Xiu, D.C.K Chau, "Application of data mining techniques in customer relationship management: A literature review and classification", Expert system with applications (Elsevier) 36(2009), 2592-2602.
- [9] Fayyad, U. M; Piatetsky-Shapiro, G.; Smyth, P.; and Uthurusamy, R. 1996. Advances in Knowledge Discovery and Data Mining. Menlo Park, Calif.: AAAI Press.
- [10] Ghani, R., Probst, K., Liu, Y., Krema, M., Fano, A., "Text Mining for Product Attribute Extraction", SIGKDD Explorations Volume 8, Issue 1.
- [11] Growing Lifetime Value, Chain store Age(75), May 1999
- [12] Jayanthi Ranjan, IMT Ghaziabad, "A Review of Data Mining Tools in Customer Relationship Management", Journal of Knowledge Management Practice Vol 9, No 1, March 2008.
- [13] Jiawei Han and Micheline Kamber (2006), Data Mining Concepts and Techniques, published by Morgan Kaufman, 2nd ed.
- [14] Kanellopoulos, Y., Dimopoulos, T., Tjortjis, C., Makris, C. "Mining Source Code Elements for Comprehending Object-Oriented Systems and Evaluating Their Maintainability", SIGKDD Explorations Volume 8, Issue 1.
- [15] Literature Review: Data mining, <http://nccur.lib.nccu.edu.tw/bitstream/140.119/35231/8/35603108.pdf>, retrieved on June 2012.
- [16] L. Xu and Y. M. Cheung, "Adaptive supervised learning decision networks for traders and portfolios," presented at the IEEE/IAFE (Computational Intelligence for Financial Engineering), New York, 1997.
- [17] M. Weatherford, "Mining for fraud," IEEE Intell. Syst., vol. 17, pp. 4-6, Jan./Feb. 2002.
- [18] Parks L, Information resource inc, drug store news, June 2000, 219-222
- [19] Rabinovitch L "American' first department store mines customer data", Direct Marketing(62),8,42-44 Dec 1999
- [20] Rao, R. B., Krishnan, S. and Niculescu, R. S., "Data Mining for Improved Cardiac Care", SIGKDD Explorations Volume 8, Issue 1.
- [21] Rubbermaid, "Mines for profits"(1999, march), Discount stores new (38).
- [22] Shortland, R., Scarfe, R., "Data Mining Applications in BI", BT Technology 16. Journal, Vol.25, Nos 3&4, July/October (2007).
- [23] Step by step Data Mining guide, <http://www.crisp-dm.org/CRISPWP-0800.pdf>, retrieved on Jan 2010
- [24] "The 6 biggest challenges retailer Face today", www.onStepRetail.com, retrieved on June 2011
- [25] Whiting R, "Companies boost sales efforts with predictive analysis information", Feb 2002, 62-64.
- [26] Wedel, M. & Kamakura, W., 2000. Market segmentation Conceptual and Methodological Foundations 2nd ed., Kluwer Academic.