

Chapter 65

A Predictive Modeling of Retail Satisfaction: A Data Mining Approach to Retail Service Industry

M. Hemalatha

M.A.M. College of Engineering, India

ABSTRACT

Data mining allows managers to make more knowledgeable decisions by predicting future trends and behaviors. One of the most widely used areas of data mining for the retail industry is in marketing. There are also at least seven methods of analysis or statistical techniques that are commonly used in data mining. It is obvious that the approach to the data mining is the key determinant of the statistical technique to be used. Predictive modeling uses variety of techniques such as linear regression, logistic regression, and their extensions can be used to identify patterns, which can be used to predict the future. This research specifically focuses on the application of multiple regression technique a data mining tool in Indian retail industries to predict the retail satisfaction using store attributes as independent variables.

1. INTRODUCTION

Data analysis, especially statistics, has been used in the social life of human beings for more than two hundred years. Data analysis could not widely affect our daily life until the creation of computer in 1950s. Using computational and storage power

of computer, people gradually developed database, database management system, and data warehouse in the last 40 years. These computing environments provided the bases for us to efficiently use a number of mathematical tools, such as statistics, artificial intelligence, machine learning, and so on to solve larger-scale data analysis problems. Now, here comes data mining, the interface of some of the approaches. Strictly speaking, data mining discov-

DOI: 10.4018/978-1-4666-2455-9.ch065

ers the hidden patterns of data from a large-scale data warehouse by precise mathematical means (Cheng *et al.*, 2006).

Loosely speaking, data mining may refer to the analysis of the large quantities of data that were stored in computers. It can be viewed from three different angles: process, methodology, and mathematical methods. From the aspect of the process, data mining consists of four stages: (1) selecting, (2) transforming, (3) mining, and (4) interpreting. While stages (1) and (4) are mainly business-related problems, stages (2) and (3) involve mathematical modeling and computer algorithms. From the aspect of methodology, the functions of data mining can be classified as association, classification, clustering, predictions, sequential patterns, and similar time sequences. Each function can be applied to different nature of the problems. From the aspect of mathematical methods, the functions of data mining can be implemented through statistics decision tree, neural networks, rough set, fuzzy logic and mathematical programming (Cheng *et al.*, 2006).

Data mining has become a widely accepted process for organizations to enhance their organizational performance and gain a competitive advantage. Because the data mining process is a relatively new concept, it has been defined in various ways by various authors in the recent past. The definitions of data mining given by these authors are somewhat different but all have the same idea: to extract important information from existing data and enable better decision making throughout an organization. Not only data mining can improve decision making by searching for relationships and patterns from the extensive data collected by organizations, but it can also reduce information overload (Zhu *et al.*, 2001). Data mining enables an organization to focus on the most important information in the database, which allows managers to make more knowledgeable decisions by predicting future trends and behaviors. Before data mining was utilized, managers were not as capable of making such informed decisions

because searching through large amounts of data was too expensive and time-consuming.

One of the most widely used areas of data mining for the retail industry is in marketing. Market basket analysis is a marketing method used by many retailers to determine the optimal locations to promote products. Another marketing tactic employed by many retail stores is the use of loyalty cards. There are an increasing number of retail companies utilizing data mining for marketing purposes and benefiting from its use. Risk management is another area where data mining is used in the retail industry; however, there has not been as much research done in this area. Retail industries must also be aware that fraud detection is necessary. Fraud occurring at POS terminals is a concern for retailers, but can be reduced using data mining (Hormozi and Giles, 2004).

In marketing, data mining enables an organization to sort through vast amounts of customer data to target the right customers. In risk management, data mining helps an organization predict the number of customers who are likely to be lost to competitors, so that the business can be better prepared. Data mining is useful for fraud detection because, by detecting these fraudulent actions, an organization can keep from losing substantial amounts of money. Data mining is also important for customer acquisition and retention, enabling an organization to target new customers and retain the customers it already has. Those industries that effectively utilize data mining have the ability to develop or enhance a competitive advantage (Hormozi and Giles, 2004).

Throughout the literature, data mining operations are also classified in different ways. The following are a few categorizations of data mining operations, including (1) clustering/segmentation, (2) visualization, (3) predictive modeling, (4) link analysis, (5) deviation detection, (6) dependency modeling, and (7) summarization (Hormozi and Giles, 2004). This study made an attempt to identify the dimensions of retail satisfaction based on the retail strategy mix which influences the

13 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/predictive-modeling-retail-satisfaction/73495

Related Content

User-Centric Similarity and Proximity Measures for Spatial Personalization

Yanwu Yang, Christophe Claramunt, Marie-Aude Aufaure and Wensheng Zhang (2010). *International Journal of Data Warehousing and Mining* (pp. 59-78).

www.irma-international.org/article/user-centric-similarity-proximity-measures/42152

A Survey of Parallel and Distributed Data Warehouses

Pedro Furtado (2009). *International Journal of Data Warehousing and Mining* (pp. 57-77).

www.irma-international.org/article/survey-parallel-distributed-data-warehouses/1826

Remarks on a Fuzzy Approach to Flexible Database Querying, Its Extension and Relation to Data Mining and Summarization

Janusz Kacprzyk, Guy de Tré and Slawomir Zadrozny (2013). *Data Mining: Concepts, Methodologies, Tools, and Applications* (pp. 279-298).

www.irma-international.org/chapter/remarks-fuzzy-approach-flexible-database/73444

Gaussian-Stacking Multiclassifiers for Human Embryo Selection

Dinora A. Morales, Endika Bengoetxea and Pedro Larrañaga (2009). *Data Mining and Medical Knowledge Management: Cases and Applications* (pp. 307-331).

www.irma-international.org/chapter/gaussian-stacking-multiclassifiers-human-embryo/7539

Contextualized Text OLAP Based on Information Retrieval

Lamia Oukid, Nadjia Benblidia, Fadila Bentayeb, Ounas Asfari and Omar Boussaid (2015). *International Journal of Data Warehousing and Mining* (pp. 1-21).

www.irma-international.org/article/contextualized-text-olap-based-on-information-retrieval/125648