Data mining techniques for customer relationship management

Feng Guo

Information School of Beijing Wuzi University, Beijing
E-mail:guofeng80050@sina.com

Data mining has made customer relationship management (CRM) a new area where firms can gain a competitive advantage, and it plays a key role in the firms' management decision. In this paper, we first analyze the value and application fields of data mining techniques for CRM, and further explore how data mining being applied to Customer churn analysis.

Keywords: Data mining; Customer Relationship Management (CRM); Customer Loyalty.

1. Introduction

A new business culture is developing today. The conventional production centered and sales purposed market strategy is gradually shifting to customer centered and service purposed. Customers' value orientation is increasingly affecting the firms'. And customer resource has become one of the most important strategic resources. Therefore, understanding customers' needs and discriminating the most contributed customers has become the driving force of most modern business.

2. Customer Relationship Management

Customer relationship management (CRM) is a customer centered business strategy. Firms abidingly accumulate, analyze and exploit customer comprehensive information to allocate resources rationally according to customer oriented working mode, and meet customer demand through meaningful communication, understanding and influence customer behavior. Thereby increasing customer satisfaction and loyalty, and to obtain the maximum profit. The general CRM covers customer resource management, marketing management, sales management, customer care and service management, and many other aspects. Its ultimate goal is to provide the basis for business management and decision-making through proactive market research and analysis and processing of customer feedback. Simple from the content of

customer relationship management, CRM is customer propaganda, tracking, revisit and classification. Of course, with the development of modern information technology, the pursuit of personalized service has become the new target in the field of after-sales service.

3. Overview of Data Mining Techniques

Firms want to acquaint customer behavior as much as possible, but as the number of customers accumulated in large quantities and customer information has become increasingly complex, more advanced technologies and tools are needed to manage and analyze customer information so as to find out the valuable knowledge of management decision-making. Data mining technology provides a good technical support for CRM to analyze large amounts of complex customer data and explore customers' value.

3.1. *Definition*

Data mining can extract potentially valuable knowledge, model and rules from mass of data. It is an effective tool to exploit potential associations and patterns. With the help of statistics and artificial intelligence technology, data mining can make further analysis of data and find out knowledge, and build different models in accordance with different business issues, thus provide a technical reference for the firms' decision-making. Data mining involves a considerable number of subject areas, most of them are mature algorithms and techniques in artificial intelligence such as machine learning, statistics, neural network and database techniques. According to the purpose of data mining, it can be divided into association analysis, time series analysis, classification, cluster analysis and data summary.

3.2. *Functionality*

Through the prediction for future trends and behavior, data mining can help firms make a forward-looking and knowledge-based decision-making. The goal of data mining is to find implied and meaningful knowledge from database. Its functionality is to designate type of pattern that data mining tasks are looking for, and can be divided into two categories: description and prediction. Descriptive mining tasks characterize the general characteristics of the data in the database, while predictive ones extrapolate the existing data to predict. Specifically, firms can discover most valuable customers, make the combined sales more efficient and find fraud with a smaller cost with the help of data mining techniques.

3.3. *Methods*

As an advanced information processing technology, the essential difference between data mining and traditional data analysis is that it is a process of exploration of data relationships, and in most cases with no assumptions and premises. There are numbers of different ways in data mining for users to embark on a comprehensive analysis of data from different perspectives.

(1) Correlation and regression analysis

Correlation analysis mainly analyzes the closeness of variables, and regression analysis mainly establishes dependencies between variables based on observational data. Both of them can reflect valuable associations and contacts among variables. So they can be collectively referred to as association analysis.

(2) Time series analysis

Similar to association analysis, time series analysis is aimed at tap the intrinsic link between data. The difference is that time series analysis focuses on chronological causal relationship of data, while association analysis focuses on parallel relationship.

(3) Classification and prediction analysis

Classification and prediction analysis are used to extract models describing important data. The model is then applied to determine and classify new observations, or predict future data trends.

(4) Cluster analysis

Cluster analysis categorizes data into more than one classes or clusters according to certain characteristics. Objects must have a high degree of similarity if they are from a same cluster, and the difference will be much larger if not. We could see that the cluster analysis is the reverse process of classification and prediction to a certain extent.

4. Value of Data Mining for CRM

CRM collect and accumulate a large number of market and customer information for the firms to build a huge data warehouse. The key to data mining is to distinguish truly valuable information and find the association between them from massive data.

Every firm wants to quickly and accurately extract the most important and valuable customer information through the use of certain tools and instruments. After a deep analysis for a huge data warehouse, data mining techniques can find hidden information and knowledge conducive to business operations, thus improve the competitiveness of the firm. Data mining can help firms manage customer life cycle stages, including acquiring new customers, making existing

customers create more profits, maintaining valuable customer, etc. It can do in-depth analysis of customer needs, to meet firms' needs of customer relationship in individual market segments. So the firms are able to determine customer characteristics, provide them with targeted services and increase business opportunities.

5. Application Fields Of Data Mining For CRM

Data mining can filter customer information, and mining implicit, unknown and potentially valuable knowledge on business decisions and rules from a huge number of customer data. To help firms identify new opportunities, predict success in business strategies and make quick decisions. The following aspects of its application are particularly prominent.

5.1. *Customer character analysis*

In addition to customers' address, gender, age, occupation, income, education and other basic information, access, such as hobbies, marriage, spouse, health, home environment and other features can help firms get a more detailed understanding of customers, observe their behavioral regularities and thus better develop customer strategies to improve their campaign response speed. The main idea is to use classification and clustering techniques, divide customers into different group with different demand and transaction practices according to their age, gender, income, trading behavior, etc. And finally arrive at customers' concern, targeted to the development of personalized marketing strategy.

5.2. *Customer loyalty analysis*

Customer loyalty analysis is to analyze and categorize upscale customers, stable customers, valuable customers, more consumption demand customers, etc. To help firms make customers persistent and stability analysis and respond quickly to customer needs, so that each customer will get a highly personalized service. Statistically, firms must pay much higher cost to get new customers than to keep old customers. The gap is recognized to be more than six to eight times, no matter what business they focus on. Meanwhile, according to 80/20 principles of marketing, that 20% of customers contributed 80% of sales, more strategies should be adopted to develop high consumption customers' loyalty.

5.3. *Cross marketing analysis*

The previously customer information may contain the key or even critical factors to determine customer behavior. Using data mining techniques, firms can get the key factors affecting customer purchasing behavior from the customer information especially former purchase behavior, and build prediction model to predict customer future purchase behavior. Association analysis is often used to help companies find some implicit, subtle and great commercial value of relationship. This mining process is generally divided into two steps: First, find all the frequent item sets whose frequency should be at least equal to the minimum support frequency, thereby identify all related merchandises that may be purchased together; Then, generate strong association rules according to the frequent item sets. These rules must meet the minimum threshold of confidence; thereby identify all related merchandises that are very likely to be purchased together. To sum up, firms can discover hidden relationship between seemingly independent events; find factors that affect customer behavior through data mining technology. And then target to expand marketing strategy and promote other products in a timely manner.

5.4. *Customer acquisition and churn analysis*

The growth and expansion of the firm need continuously maintain old customers and draw new customers. Through classification, clustering, decision tree and other techniques, it can extract mass customer information, identify potential customer, and determine what kind of customer are most likely to loss and what features they have. Thus build customer churn prediction model to more accurately identify easily lost customers. This would help firms making plans to take appropriate marketing measures in advance to hold old customers.

5.5. *Tracking evaluation*

Through clean and centralize, customers' feedbacks are automatically store in data warehouse. So the firm can track customer behavior, analyze customer satisfaction, credit rating and so on, in order to evaluate and optimize the existing strategy, such as take different credit terms on different credit rating customers. This could maintain customer loyalty and avoid unnecessary risks at the same time. By tracking the evaluation to ensure that the firm customer relationship management to achieve the goals, also establish a good customer relationship.

6. Customer Churn Analysis

Statistically, firms must pay much higher cost to get new customers than to keep old customers. The gap is recognized to be more than six to eight times, no matter what business they focus on. So firms should take strategy to develop high consumption and high loyalty customers' loyalty, and use data mining techniques to identify customers that are likely to loss and take measures to prevent the loss. Commercial enterprises usually record sales information and customer information through transaction records, sweepstakes, membership and other activities. And then dig out customer preferences merchandise mixes, customer habits and characteristics of lost customers from this information. The following will take customer churn analysis as an example to illustrate the process of data mining in CRM.

6.1. *Data selection*

The target variables are customers' loss states. In another word, by analyzing a specific problem, select a combination of one or a plurality of known amount variables as target variable. What specific consumption characteristics lead to customer loss? Is that income change? Or preference change? Or the demand of shopping environment change?

The input arguments include static and dynamic data. The static data is the data that changes infrequently, such as the customer basic information (gender, marital status, occupation, residential areas, etc.); the dynamic data is the data that changes frequently or regularly, such as the consumption amount in a month, consumption characteristics, etc. There are intrinsic links between the input arguments and the target variables, but cannot be quantified. This provides data mining a play space. If we are unable to determine whether some sort of data is associated with customer churn probability, we could temporarily select it in the model, and examine each variable distribution and correlation in the subsequent step to determine whether to keep it.

Select the modeling data. There two kinds of customer churn: subside and transfer to competitors. The formal is always caused by firms' relocation, or customer service upgrade such as ordinary customers upgrade to the VIP. But the latter is what firms really care about. We must choose the latter loss of customer data to carry out the modeling.

6.2. *Data reduction*

Take data reduction before building the model, on the one hand, can ensure the correctness and validity of the modeling data, and on the other hand, can make the data more in line with the modeling needs through the adjustment of data format and content. The main tasks include data conversion, integration, sampling, randomization, missing value handling, etc. Such as proportionally extract the information of existing customers and lost customers and merge these two types of data as data source for modeling. The model need large amount of data to test, so the data sample is divided into two part, one for modeling and the other for model checking and correction.

6.3. *Model building*

Before modeling, we should use correlation analysis to identify the correlation between each input argument and the probability of loss of customers, delete the less correlation variable, thereby reducing the degree of complexity of the model, thus shorten the modeling cycle. There exist many data mining methods, such as neural networks, regression, decision tree, association analysis, clustering, etc. We can use a variety of methods pre-build some models then compare the pros and cons of these models and select the most suitable for customer churn analysis. There are also a variety of approximation algorithms to simplify the model optimization process. The core part of data mining process is adjustment, which is usually complete by data analysis experts. Different business issues and data distribution properties will affect the modeling and adjust its strategy; therefore business experts are required in order to avoid inappropriate optimization resulting in loss of business information.

6.4. *Model evaluation and interpretation*

To obtain accurate result, we can use the data not involved in the modeling for model evaluation. The evaluate matric is to use the mode to predict the already known customer state, and compare the predicted result and the actual customer state. The model with the highest prediction accuracy is the best. Through making a reasonable business explanation to the model, firms can identify some potential law to guide businesses behavior. In turn, the explanation can also prove the correctness and effectiveness of the mathematical model.

6.5. *Model application*

To make the model practical, we can first select a pilot to implement the model, and pay attention to the earnings of the model application during the pilot. If an exception occurs, stop the application and correct the model. After the end of the pilot, if the model is proved to be good, then it can consider a wide range of promotion. The promotion should pay attention to regional differences, and cannot just copy the model. The model bias may increase with the growth of time and significant changes in economic, and then we should consider re-establish a more applicable model.

The churn model can monitor the loss of customer, if the possibility of loss is too high, firms can take measures to increase customer loyalty, and prevent the occurrence of loss of customers, which completely change the situation that firms cannot monitor the loss of customers and cannot effectively achieve customer care even they successfully get many customers.

7. Conclusion

The CRM system using data mining technology has become the key for firms to winning in the increasingly competitive knowledge economy and e-commerce economy mode. Data mining as the premise and foundation of CRM is no longer just confined to customer contact level customer relationship management. On the contrary, it has been deep into customer behavior and customer preferences, and can anticipate their needs and meet their expectations. Thus firms can insight into customer, understand customer value from a more in-depth and comprehensive perspective. Effective use of data mining in CRM, can continually promote a single customer value enhancement and expansion of the customer base, guide high-level decision makers to work out best marketing strategy, reduce operating costs, increase profits, strengthen customer relationships, improve customer satisfaction, improve, improve commodity promotions, enhance information exchange of goods, improve customer loyalty, accelerate the development of firms, and enable firms to gain long-term benefits.

References

1. Xianyan Lin. A Research on Telecom Customer Loss Based on Data Mining [J]. Journal of Mobile Communication, 34(8). P71-75.
2. Guoen Xia. Customer Churn Prediction Based on Simple Support Vector Machine [J]. Research on Application of Computer, 2010, 27(3).

3. Liuyin Guo, Zhi Fang. Application of Data Mining on Retaining Customers in Telecom Industry [J]. China New Telecommunications, 2005(11).
4. Hongxia Wang, Gencai Chen. Data Mining Techniques Applied in Retailing's CRM [J]. Computer and Digital Engineering, 2006(1).
5. Alex Berson, Stephen Smith, Kurt Theading. Building Data Mining Applications for CRM [M]. Beijing: People's Posts and Telecommunications Press, 2001.
6. MeiZhen Zhu. To Promote Customer Relationship Management Data Mining [J]. Technology and Innovation Management, 2006(4).
7. Sheng Tian. The Application of Data Warehouse and Data Mining in CRM [J]. Journal of Information, 2005, 24(9). P18-20.
8. Ning Dong. Data Mining Techniques for CRM [J]. Computer Engineering and Design, 2007, 28(6).