2014

ISSN : 0974 - 7435 Volume 10 Issue 22

BioTechnology

An Indian Journal

FULL PAPER

BTAIJ, 10(22), 2014 [13821-13826]

Evaluation on the meaning and value of applying data mining algorithm to bank financial products marketing

Shan Zhao

Jincheng College, Sichuan University, Chengdu, 611731, (CHINA)

ABSTRACT

Nowadays, more and more financial institutions are delving into the application of data warehouse and data mining techniques to bank financial products marketing. It has been a research hotspot in the financial sector of China. Currently, relevant research mainly focuses on data mining techniques and algorithm as well as on establishment of customer relationship management system. As a new research hotspot in database field, the data mining technique combines traditional data analysis technique with complex algorithm which processes big data. Meanwhile, it explores and analyzes new data type, as well as processing old date type with new methods. After elaboration on the background of data mining techniques, this paper discusses major analysis techniques concerning data mining and explains the most common algorithms. Finally, the application area and future trend of data mining techniques is also pointed out. Meanwhile, great attention should be paid to mechanism security and improvement of marketing resources allocation mechanism. Besides, marketing resources allocation for the individual should also be well arranged. Marketing system is to be positioned in a mechanism for exploitation and analysis of as well as response to customer demands. Moreover, the system should deal with the authority-responsibility relationship at all levels on scientific and reasonable basis. This paper focuses on evaluation on the meaning and value of applying data mining algorithm to bank financial products marketing. Meanwhile, a design plan for the marketing system of bank financial products is put forward with an effective mining algorithm of negative association rules employed innovatively.

KEYWORDS

Practical application of data mining algorithm; Bank financial products marketing; Vagueness issues; Accuracy issues.

INTRODUCTION

With the rapid development of internet and database technology along with its increasing capacity, enormous information stored in the database renders it difficult for us to exploit the information we need[1]. Data mining comes just in time to help solve this problem so as to exploit effective data and cut unnecessary costs for enterprises. In fact, data mining techniques have been employed in many companies to help identify the most valuable customer groups and reformulate promotion strategies so as to gain higher returns with less investment.

In this study, the typical process of merchandise sales in shopping malls is taken as an example for discussion convenience. Retailers keep detailed records of all the transactions, including transaction numbers, customer numbers and merchandise numbers. Now, relational analysis is conducted on customer purchasing behaviors with different types of merchandise taken into a group called "Item"[2]. The transaction records kept by retailers constitute a record set of relational analysis with each of the records composed of transaction numbers, customer numbers, merchandise numbers, the quantity, and date, as shown in TABLE 1.

Decision trees used in data mining can be employed for data analysis and predictions as shown in Figure 1. Common algorithms for data predictions include CHAID, Quest and C5.0. It' more likely to deal with variables through the establishment of decision trees merely after a scan of data in the database. But a tremendous decision tree is not suggested in consideration of the prediction results. The scale of the decision tree can be limited by setting its maximum height. Another way is to preset the minimum record number in a node, which will stop splitting when present records are less than the minimum number. Sometimes, the decision tree can be free to grow to its fullest, and then it will be trimmed to the required size. Accuracy of the decision tree should be maintained in the process. Greedy algorithm can be employed in node splitting of a decision tree. Its thread is as follows[3]: all the splitting goes on sequentially; the rationality of a splitting will be ignored after the nod has split, which means each splitting relies on the rationality of the last splitting. Due to this limitation, greedy algorithm is rarely used.

GENETIC ALGORITHM

Genetic algorithm is the generation of new descendant string population from its parent population iteratively[4], comprising the following steps: 1) determination of study object, which forms the basis of the whole data mining process; 2) establishment of database for collection, description, selection and other pretreatment of data; 3) data analysis-the common analysis techniques (such as relational analysis and sequential pattern analysis) introduced in this study can be employed for data analysis and feature induction; 4) data preparation, including selection, record and transformation of variables; 5) establishment of model, which needs proper mining algorithms to mine transformed data; 6) evaluation and explanation, which are due for results concluded from the model; 7) application (assimilation of knowledge)-analysis results are to be applied to the same type of system. Iterative algorithm based on survival of the fittest is used to make the search process approach the optimum solution.

TABLE 1 : L_1 Rule Generated by GA_PNAR Algorithm

X_1	=>	Y_1	Supports	Confidence	Fitness
Average Fund Balance	=>	Third Party Depository	0.519608	0.697368	0.596138
Bank Securities Link	=>	Mortgage Loan	0.401961	0.488095	0.366865
Demand Deposit Interest Rate	=>	Third Party Depository	0.401961	0.97619	0.645545
Bank Securities Link	=>	Third Party Depository	0.607843	0.738095	0.738095
Mortgage Loan	=>	Third Party Depository	0.205882	0.35	0.118548
Bank Securities Link	=>	Demand Deposit Interest Rate	0.411765	0.5	0.5
Bank Securities Link	=>	Average Fund Balance	0.735296	0.892857	0.881109
Demand Deposit Interest Rate	=>	Average Fund Balance	0.35294	0.857143	0.406015
Mortgage Loan	=>	Average Fund Balance	0.352941	0.857143	0.350877

Genetic algorithm mainly applies genetic operators such as selection, crossover and mutation to an initially random population. Currently, a lot of biomedical research is focused on data analysis of DNA sequence, based on which have the causes of enormous disease and defective genes been discovered. Relational analysis is used to identify gene sequence occurring simultaneously; path analysis is employed to discover the pathogenic gene sequence at different stages of diseases; visual tools and genetic data analysis techniques can also be used to exhibit complex structures and sequences of genes. The crossover probability can automatically adjust itself according to following equations as (1) and (2).

$$P_c = \begin{cases} P_{c1} - \dfrac{(P_{c1} - P_{c2})(f' - f_{avg})}{f_{\max} - f_{avg}}, & f' \geq f_{avg} \\ P_{c2}, & f' < f_{avg} \end{cases} \quad (1)$$

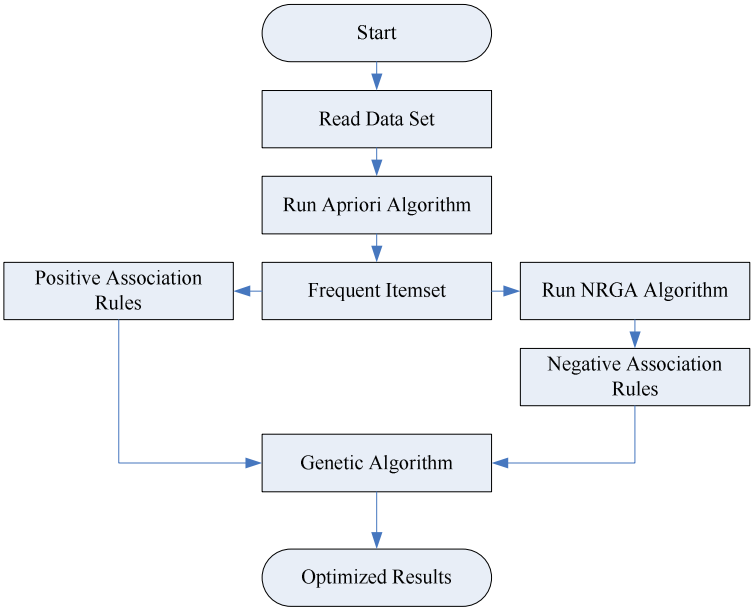

Figure 1 : GA_PNAR algorithm flowchart

$$P_m = \begin{cases} P_{m1} - \dfrac{(P_{m1} - P_{m2})(f_{\max} - f')}{f_{\max} - f_{avg}}, & f' \geq f_{avg} \\ P_{m2}, & f' < f_{avg} \end{cases} \quad (2)$$

In this study, the genetic algorithm for association rules mining is used based on Apriori algorithm[5], as shown in Figure 1.

Many rules can be found out from TABLE 1. Bank and financial institutions offer various savings service, credit service, investment and financing service, insurance service etc[6]. Based on the business and data features, typical data mining techniques are shown in the following: (1) design and building of database according to multidimensionality of data; (2) data analysis by loan repayment prediction technology and customer credit policy analysis; (3) segmentation and cluster analysis on customers of specific target market.

APPLICATION OF DATA MINING ALGORITHM TO BANK FINANCIAL PRODUCTS MARKETING

Existing data mining techniques, along with some new and improved algorithms, can greatly boost the cross-selling accuracy of bank financial products so as to improve the profit margin[7]. However, how to put these techniques into practice to design an effective and powerful marketing analysis system for bank financial products, providing decision basis for banks to implement profit-oriented strategies, it's one of the practical issues worth studying.

Faced with fierce competition, commercial banks in China find it hard to continue its traditional growth model by relying too much on intermediary business income and interest margin. With the development of various financial services in commercial banks of China, increasing importance has been attached to customer relationship management. How to achieve the optimum match between customer groups and bank financial products has been an urgent problem to address. Association rule mining provides effective technical means for the discovery of potential information in practical business data. In this paper, information about financial customers and the financial products they have purchased are taken as research objects, constituting the customer-financial-product association model. Through this, the purchasing behaviors and habits of customers on bank financial products can be dug out to provide decision support information for marketing departments of bank financial products so as to better promote their financial services.

To cope with challenges mentioned in this paper, meet the need of banks for data sharing and profit-oriented service, and offset the deficiency in uncovering and analysis of product sales information, we propose the following system design objectives.

Effective Customer Segmentation: it's generated automatically through collecting the natural information, value information and behavior information of customers by means of data mining tools. Based on the different segmentation results, banks can effectively master the current status, actively improve customer relationship management style, discover and elevate customer value so as to form the basis for offering professional and personalized service to customers.

Analysis and Prediction of Customer Value: data mining techniques are employed to analyze customers' contribution to bank profit with a view to discovering customer groups of high value and potential value. Analysis on the

natural properties and behavior properties of these customer groups can help provide decision support for banks to allocate resources, optimize product structure and improve services.

Analysis of Customers' Purchasing Preferences: connections in customers' purchasing behaviors can be discovered by study on abundant customer information. Then banks can provide effective sales strategy and value-added service for customers based on their preference and properties, so as to improve the efficiency and value of personal business marketing for banks.

Professional Customer Service: Banks are supposed to find new service mode from multiple perspectives and levels through comprehensive application of the mining tools mentioned above. They aim at providing professional, personalized and fine service for each customer group with their service quality and results evaluated at the same time.

Logical architecture design for bank financial products marketing system

The main application and research area for data mining lies in retail industry, which has accumulated abundant sales data, such as customers' purchasing records, consumption and service records, input and output records of merchandise, etc. Data mining techniques for retail industry help merchants to identify customers' purchasing behaviors and mode, so that they can improve the service quality, increase the sales ratio of merchandise and cut their costs. Data mining techniques concerned include the following: designing and building database by data mining techniques; multidimensional analysis of sales volume, customers, merchandise, time and regions; analysis of customers' purchasing loyalty. The marketing analysis system works under traditional c/s (client/server) framework, as shown in Figure 2. The overall framework includes the following aspects: establishment of database for collection, description, selection and other pretreatment of data; data analysis-the common analysis techniques (such as relational analysis and sequential pattern analysis) introduced in this study can be employed for data analysis and feature induction; data preparation, including selection, record and transformation of variables.

The logical architecture of bank financial products marketing analysis system includes data acquisition layer, information integration layer, analysis application layer, and information interaction layer. Based on the business and data features, typical data mining techniques are shown in the following: 1. design and building of database according to multidimensionality of data; 2. data analysis by loan repayment prediction technology and customer credit policy analysis; 3. segmentation and cluster analysis on customers of specific target market. The main application and research area for data mining lies in retail industry, which has accumulated abundant sales data, such as customers' purchasing records, consumption and service records, input and output records of merchandise, etc. Data mining techniques for retail industry help merchants to identify customers' purchasing behaviors and mode, so that they can improve the service quality, increase the sales ratio of merchandise and cut their costs. Data mining techniques concerned include the following: designing and building database by data mining techniques; multidimensional analysis of sales volume, customers, merchandise, time and regions; analysis of customers' purchasing loyalty.

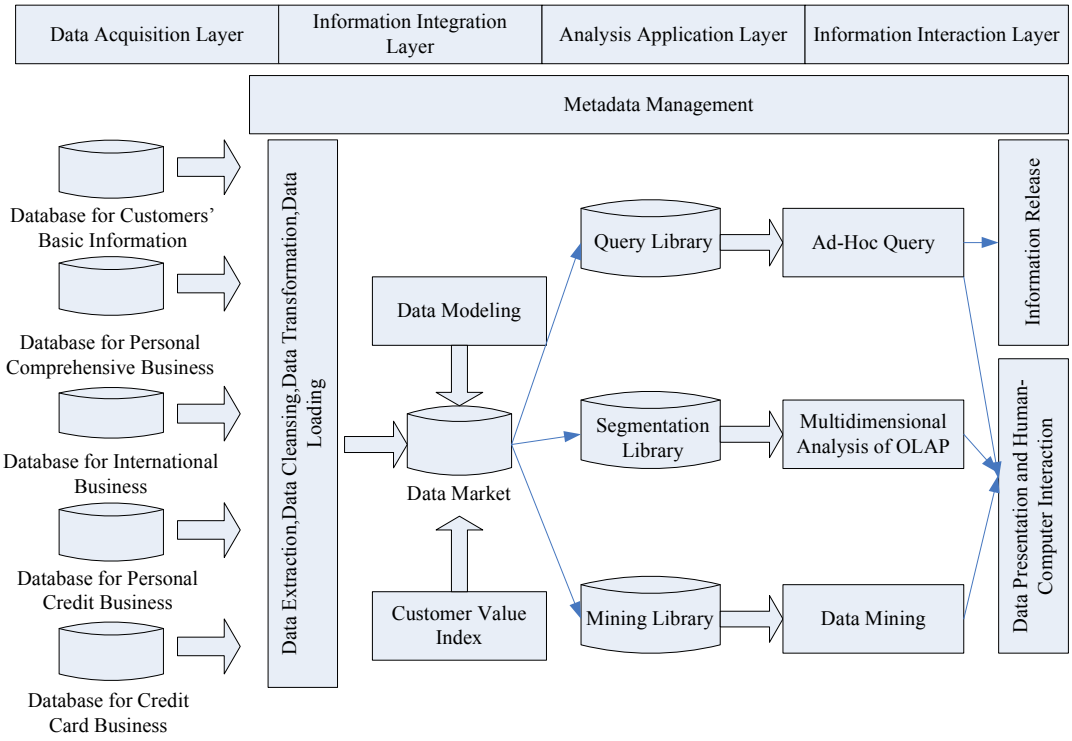

Figure 2 : Logical architecture of bank financial products marketing analysis system

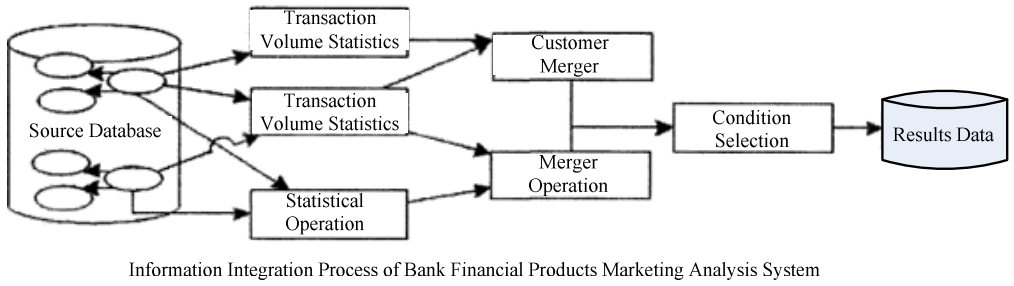

Figure 3 : Information integration process of bank financial products marketing analysis system

Data acquisition and integrated module design

To meet the needs for customer information analysis, decision trees are widely applied. Now they can be used to determine the rules for the way a certain value is acquired. For example, when dealing with mortgage loans, the bank needs to judge the risk brought by the loans.

Decision trees used in data mining can be employed for data analysis and predictions as shown in Figure 1. Common algorithms for data predictions include CHAID, Quest and C5.0. It' more likely to deal with variables through the establishment of decision trees merely after a scan of data in the database. But a tremendous decision tree is not suggested in consideration of the prediction results. The scale of the decision tree can be limited by setting its maximum height. Another way is to preset the minimum record number in a node, which will stop splitting when present records are less than the minimum number. Sometimes, the decision tree can be free to grow to its fullest, and then it will be trimmed to the required size. Accuracy of the decision tree should be maintained in the process. Greedy algorithm can be employed in node splitting of a decision tree. Its thread is as follows[3]: all the splitting goes on sequentially; the rationality of a splitting will be ignored after the nod has split, which means each splitting relies on the rationality of the last splitting. Due to this limitation, greedy algorithm is rarely used. Customer group analysis includes asset distribution features, age distribution features and regional distribution features of customers. The process for information integration can be seen in Figure 3.

Query module design

Business query software mainly offers the following functions: customer information input and query, financial advice by customers, financial products input and query, and system maintenance, as shown in Figure 4. Clustering is a process in which a set of physical or abstract objects are grouped into different clusters consisting of similar objects. The input set of cluster analysis techniques is a set of marked records, which means the input set hasn't been classified. It means to reasonably divide record sets based on specific rules and classify them according to explicit or implicit rules. Generally, various algorithms can be used in cluster analysis techniques, so different division results may be produced from the same record set. The classification analysis mentioned above can be also applied to cluster analysis.

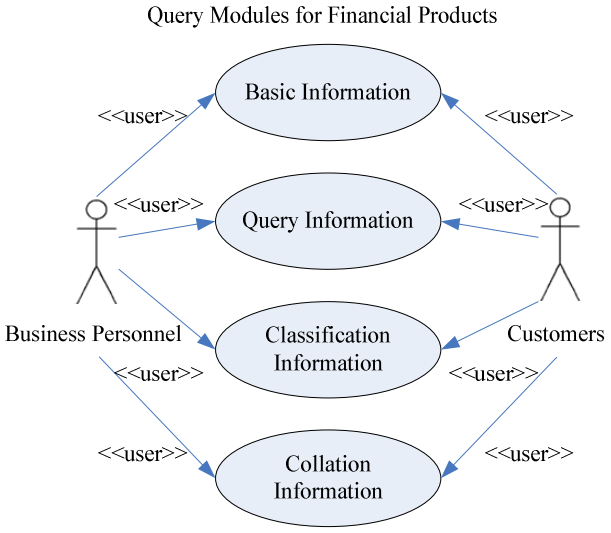

Figure 4 : The interaction diagram of use cases for financial products query

In fact, classification analysis and cluster analysis complement each other in many aspects. For example, in initial data analysis, analysts can mark and classify the data to be analyzed based on experience or general rules. Then classification analysis techniques are used for data analysis to get general descriptions of each category. These descriptions serve as new classification rules and classify the clusters again to obtain better results. Analysts can use these two analysis techniques in

turn to get satisfying results so as to help customize financial plans for bank customers. The main application and research area for data mining lies in retail industry, which has accumulated abundant sales data, such as customers' purchasing records, consumption and service records, input and output records of merchandise, etc. Data mining techniques for retail industry help merchants to identify customers' purchasing behaviors and mode, so that they can improve the service quality, increase the sales ratio of merchandise and cut their costs. Data mining techniques concerned include the following: designing and building database by data mining techniques; multidimensional analysis of sales volume, customers, merchandise, time and regions; analysis of customers' purchasing loyalty.

CONCLUSION

This paper focuses on evaluation on the meaning and value of applying data mining algorithm to bank financial products marketing. Meanwhile, a design plan for the marketing system of bank financial products is put forward with an effective mining algorithm of negative association rules employed innovatively. As a new research hotspot in database field, the data mining technique combines traditional data analysis technique with complex algorithm which processes big data. Meanwhile, it explores and analyzes new data type, as well as processing old date type by new methods. After elaboration on the background of data mining techniques, this paper discusses major analysis techniques concerning data mining and explains the most common algorithms. Finally, the application area and future trend of data mining techniques is also pointed out.

REFERENCES

[1] Yonghong Xie; Customers marketing and development of financial products--study on application of improved apriori algorithm to bank financial marketing [J], Heilongjiang Finance, **30(7)**, 46-50 **(2012)**.

[2] Bin Yang, Tian Tian; Customer classification and bank financing-study on application of associative classification algorithm to bank financial marketing [J], Finance Forum, **26(2)**, 67-72 **(2011)**.

[3] Tong Zhang; Empirical study of identified matrix of bank personal financial customers-based on the balance of perceived risk and perceived benefit [J], Journal of Shanxi Finance and Economics University, **28(11)**, 67-72 **(2011)**.

[4] Yajing Si, Hongping Du; Analysis on relationship marketing strategies in bank personal financial business [J], The Merchandise and Quality, **21(1)**, 140-146 **(2011)**.

[5] Wen Lu; Establishment of satisfaction model for commercial bank customers by means of data mining techniques [J], Financial Computer of China, **26(5)**, 32-33 **(2009)**.

[6] Ping Feng, Huiyu Xuan; Data mining techniques and its application to marketing [J], Journal of Beijing Institute of Light Industry, **12(6)**, 131-136 **(2011)**.

[7] Yue Hao, Yueping Du; Analysis of commercial bank financial product brands and customers' investment choice [J], Commercial Research, **25(8)**, 130-133 **(2010)**.