S. Thiripura Sundari, A. Padmapriya / International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com
Vol. 2, Issue 4, July-August 2012, pp.317-320

Data Mining Application-Usage of Visualizing Association Rules in CRM System

S. Thiripura Sundari*, Dr.A. Padmapriya**
*(Department of Computer Science and Engg, AlagappaUniversity, Karaikudi-630 003)
** (Department of Computer Science and Engg, AlagappaUniversity, Karaikudi-630 003)

ABSTRACT

Data Mining, simply called DM, is a kind of method for data analyses at deep level. Data Mining can be described as this: It is, in light of the set target of an enterprise, an effective and advanced method in probing and analyzing huge numbers of enterprise data and revealing the unknown regularities or verifying the known disciplines and model them better. Association rule mining is one of the most popular data mining methods. However, mining association rules often results in a very large number of found rules, leaving the analyst with the task to go through all the rules and discover interesting ones. Sifting manually through large sets of rules is time consuming and strenuous. Visualization has a long history of making large amounts of data better accessible using techniques like selecting and zooming. In this paper we introduces the Visualizing Association Rules using CRM based data mining and the concept of Clustering Analysis, and then started from Customer relationship management's core values-Customer Value, deeply expound the meaning of customer value and it is in a key position in the customer relationship management.

Keywords - **Customer Relationship Management, Data mining**

I. INTRODUCTION

Many organizations are generating a large amount of transaction data on a daily basis. Enterprises must rely on customers to achieve profitable. In the fierce market competition, in order to maintain superiority and long-term and stable development, we must attach importance to customer relationship management. And only continued to gain and maintain valuable customers and achieve customer demand for personalized, in-depth excavation in order to achieve the corporation-customer win-win. This paper describes in depth the importance of customer value for the company, noting that customer value mining is the key to the successful implementation of CRM. And to use data mining techniques to customer analysis to help enterprise to better implement CRM and develop customized marketing strategies. This kind of customer relationship management, customer-centric business strategy, is based on information technology. It restructures the work processes in order to give businesses better customer communication skills, to maximize customer profitability.

The main objective of its management is to increase efficiency, expand markets and retain customers. Data mining is data processing in a high-level process, which from the data sets to identify in order to model to represent knowledge. The so-called high-level process is a multi-step processing, multi-step interaction between the repeated adjustments, to form a spiral process.

The knowledge of mining expresses as Concepts, Rules, Regularities and other forms. Clustering Analysis is a widely used data mining analytical tools. After clustering, the same types of data together as much as possible, while the separation of different types of data as possible. The enterprise customer classification, through the customer's buying behavior, consumption habits, and the background to identify the major clients, customers, and tap potential customers. There are many clustering algorithms, this means using K-means algorithm. K-means algorithm is a specific means: for a given number of classes K, the n objects assigned to a class K to go, making within-class similarity between objects the most, while the similarity between the smallest class.

II. BACKGROUND OF THE STUDY

Association rule mining is a well-known data mining task for discovering association rules between items in a dataset. It has been successfully applied to different domains especially for business applications. However, the mined rules rely heavily on human interpretation in order to infer their semantic meanings. In this paper, we mine a new kind of

association rules, called conceptual association rules, which imply the relationships between concepts. Conceptual association rules can convey more semantic meanings than those classical association rules.

In Data Mining, the usefulness of association rules is strongly limited by the huge amount of delivered rules. In this paper we propose a new approach to prune and filter discovered rules. Using Domain Ontologies, we strengthen the integration of user knowledge in the post-processing task. Furthermore, an interactive and iterative framework is designed to assist the user along the analyzing task. On the one hand, we represent user domain knowledge using a Domain Ontology over database.

K-means clustering is a popular clustering algorithm based on the partition of data. However, there are some shortcomings of it, such as its requiring a user to give out the number of clusters at first, and its sensitiveness to initial conditions, and its easily getting to the trap of a local solution et cetera. The global K-means algorithm proposed by Likas et al is an incremental approach to clustering that dynamically adds one cluster center at a time through a deterministic global search procedure consisting of N (with N being the size of the data set) runs of the K-means algorithm from suitable initial positions.

Data mining technology has emerged as a means for identifying patterns and trends from large quantities of data. Mining encompasses various algorithms such as clustering, classification, and association rule mining. In this paper we take advantage of the genetic algorithm (GA) designed specifically for discovering association rules. We propose a novel spatial mining algorithm, called ARMNGA(Association Rules Mining in Novel Genetic Algorithm), the ARMNGA algorithm avoids generating impossible candidates, and therefore is more efficient in terms of the execution time.

1.1. Existing Methodology

A straight-forward visualization of association rules is to use a scatter plot with two interest measures (typically support and confidence) on the axes. This is introduced a special version of a scatter plot called Two-key plot. Here support and confidence are used for the x and y-axes and the color of the points is used to indicate “order,” i.e., the number of items contained in the rule.

Graph-based techniques visualize association rules using vertices and directed edges where vertices typically represent items or itemsets and edges connect antecedents and consequents in rules. Graph-based visualization offers a very clear representation of rules but they tend to easily become cluttered and thus are only viable for very small sets of rules.

Parallel coordinates plots are designed to visualize multidimensional data where each dimension is displayed separately on the x-axis and the y-axis is shared. Each data point is represented by a line connecting the values for each dimension.

Double decker plots use only a single horizontal split. Introduced double decker plots to visualize a single association rule.

1.2. Drawbacks in Existing Method

Two key plot - There is not enough space in the plot the display the labels of the items in the rules. this still makes exploring a large set of rules time-consuming.

This visualization is powerful for analyzing a single rule; it cannot be used for large sets of rules.

III. PROPOSED CRM SYSTEM USING VISUALIZING ASSOCIATION RULES

Matrix-based visualization techniques organize the antecedent and consequent item sets on the x and y-axes, respectively. Matrix-based visualization is limited in the number of rules it can visualize effectively since large sets of rules typically also have large sets of unique antecedents/consequents

Grouped Matrix-based Visualization- Groups of rules are presented by aggregating rows/columns of the matrix. The groups are nested and organized hierarchically allowing the user to explore them interactively by zooming into groups. This paper attempts to use data mining - clustering method, combined with customer life cycle assessment system for the value of customer segmentation. Through data mining technology can effectively support the use of appropriate evaluation methods to make up a simple application of conventional evaluation methods lack of customer value.

1.1. The Advantages are listed below

Matrix-based visualization is very limited when facing large rule sets.

Since minimum support used in association rule mining leads in addition to relatively short rules resulting in extremely sparse data.

IV. PERFORMANCE ANALYSIS

In Data Mining, the subject to be studied is the core of the whole process. It drives the whole

process and is the basis of examining the final results and directing the analyst to complete mining.

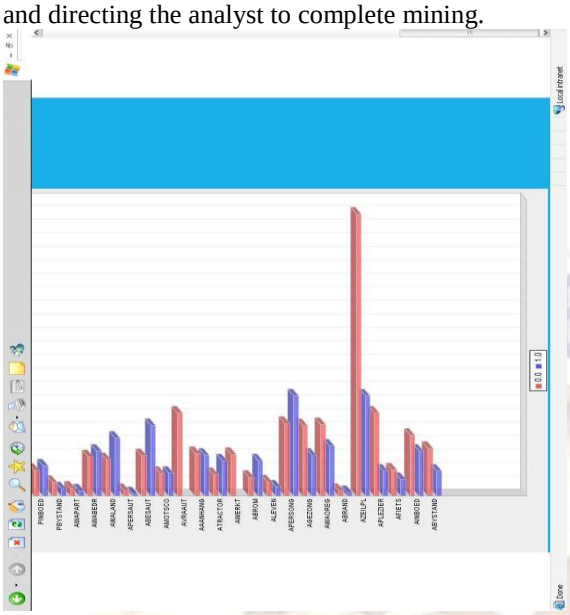

Fig. 1 Graph bar chart

1.1. Clustering or Grouping of Data

The data collected using any one of the prediction or description method and that data are stored as a record. The collected records called as a training set that data are called as training set data. Each record contains a set of attributes; one of the attributes is class. Find a model for class attribute as a function of the values of other attributes.

Attribute values are numbers or symbols assigned to an attribute. Same attribute can be mapped to different attribute values. Different attributes can be mapped to the same set of values.

1.2. Data classification

The method used here to obtain a predictive model, and to answer the questions above, is naive Bayesian learning. Boosting was tried also, and some derived attributes were added to the training set. After two important derived attributes were added, boosting did not give any significant increase in accuracy on a validation set taken from the training set, and neither did adding more derived attributes. Therefore, the results here do not use boosting, and only use two derived attributes.

1.3. Analysis with graph bar chart

The results are shown using Graph chart. Here we have used 2 numerical parameters that parameters are 0 and 1.The result shows 1 means the customer ready to get insurance policy. The result shows 0 means the customer not ready to get insurance policy. The pie chart analysis all the input values and calculates the position of the customer using Naive Bayesian Classification after shows the result. So we easily get which one will get insurance pother than deep analysis.

V. CONCLUSION

The method addresses the problem that sets of mined association rules are typically very large by grouping antecedents and allowing the user to interactively explore a hierarchy of nested groups. Coloring and the position of elements in the plot guide the user automatically to the most interesting groups/rules. Finally, the ability to interpret the grouped matrix-based visualization can be easily acquired since it is based on the easy to understand concepts of matrix-based visualization of association rules and grouping. Customer value analysis is the enterprise implementation of customer relationship management core.

For companies to develop customized marketing strategies, it is necessary to segment customer groups. This paper analyzes the shortcoming of traditional method of evaluation of customer value. Through the introduction of data mining-clustering algorithm to improve the traditional methods of evaluation, to enable enterprises to more accurate positioning for target groups with limited enterprise resources to the implementation of targeted personalized marketing policy for the enterprise-to help customers achieve a win-win.

Unlike methods based on a single clustering of the data, the approach computes the probability that two genes belong to the same cluster while taking into account the main sources of model uncertainty, including the number of clusters and the location of clusters.

.References

[1] Abd. Manan Ahmad, AG. Noorajis Ag. Nordin, Emrul Hamide Md. Saaim, Fairol Samaon and Mohd Danial Ibrahim, "*An architecture design of the intelligent agent for speech recognition and translation*," in IEEE, 2004.

[2] Bala, J., Weng, Y., Wiliams, A.I., Gogia, B. K. and H. K. Lesser, "*Applications of distributed mining techniques for knowledge discovery in dispersed sensory data*", in 7th Joint Conference on Information Sciences- JCIS,2003.

[3] Bose, R. and V. Sugumaran, "*Application of intelligent agent technology for managerial data analysis and mining*," in The DATABASE for Advance in

Information Systems, 3 (1): pp. 77-94, 1999.

[4] Freitas, A. *Data mining and knowledge discovery with evolutionary algorithms*, Natural Computing Series, Springer: Berlin, 2002.

[5] Han, J. and Kamber, M. *Data Mining : Concept and Techniques*, Morgan Kaufmann Publisher: San Francisco, USA, 2001.

[6] Kehagias, D., Chatzidimitriou, K.C., Symeonidis, A.L. and P.A. Mitkas, "*Information agents cooperating with heterogeneous data sources for customer-order management*," in *ACM Symposium on Applied Computing*: pp. 52-57, 2004.

[7] Lourenço, A., Gonçalves, J. and O. Belo, "*Agent-based knowledge extraction services inside enterprise data warehousing systems environments*," IEEE: pp. 887- 891, 2001.

[8] Ping Luo, Qing He, Rui Huang, Fen Lin and Zhongzhi Shi, "*Execution engine of meta-learning system for kdd in multi-agent environment*," in AIS-ADM 2005: pp. 149-160, 2005M.

[9] Witten, I. H. and E. Frank. *Data mining: practical machine learning tools and techniques with java implementation*. Morgan Kaufmann Publishers, San Francisco,USA,2000.http:/www.cs.waikato .ac.nz/ml/weka/.

[10] Fayyad, U., Piatetsky-Shapiro, G. and P. Smyth. *The KDD process for extracting useful knowledge from volumes of data*. Communications of the ACM, 39(11): 27-34, 1996.

S.Thiripura Sundari, a very enthusiastic and energetic scholar in M.Phil degree in Computer Science. Her research interest is in the field of data mining. It will further be enhanced in future. She is equipping herself to that by attending National Conferences and by publishing papers in International Journals.	**Dr.A.Padmapriya** is the Assistant Professor in the Department of Computer Science and Engineering, Alagappa University, Karaikudi. She has been a member of IACSIT. She has presented a number of papers in both National and International Conferences.