Integrating e-commerce and CRM with Data Mining: A new era -- Dr.S.Arumuga perumal,Head, Department of Computer Science,S.T.Hindu College,Nagercoil-2,Tamilnadu, India, Email: visvenk@yahoo.co.in Dr.S.Arumuga perumal working as a Reader and Head of the department of computer science,S.T.Hindu college,Nagercoil for the last 19 years. He completed his M.S(Software Systems) in BITS,Pilani Rajesthan.His area of interest is data mining, image compression and Biometrics.

--- Abstract Due to the fast development of information technology there is a remarkable raise in the growth of the new-fangled area data mining, which leads to revolutionize the business and Customer relationship management. The aim of this paper is to make an association with e-commerce, CRM and Data mining. E-commerce is successful only if CRM is efficient. With out CRM there is no chance of improvement in the e-commerce. In this paper some suggestions are made to get effective CRM with data mining, which will help, for the improvement of e-commerce. Data Mining involves the use of data analysis tools, which include statistical model to discover previously unknown valid patterns, and relationships in large sets. Also discuss the new method of e-business strategy for the successful business. ---

Introduction

Data mining involves the use of sophisticated data analysis tools to discover previously unknown, valid patterns and relationships in large data sets. These tools can include statistical models, mathematical algorithms, and machine learning methods (algorithms that improve their performance automatically through experience, such as neural networks or decision trees). Consequently, data mining consists of more than collecting and managing data; it also includes analysis and prediction. Data mining can be performed on data represented in quantitative, textual, or multimedia forms. Data mining applications can use a variety of parameters to examine the data. They include association (patterns where one event is connected to another event, such as purchasing a pen and purchasing paper), sequence or path analysis (patterns where one event leads to another event, such as the birth of a child and purchasing diapers), classification (identification of new patterns, such as coincidences between duct tape purchases and plastic sheeting purchases), clustering (finding and visually documenting groups of previously unknown facts, such as geographic location and brand preferences), and forecasting (discovering patterns from which one can make reasonable predictions regarding future activities, such as the prediction that people who join an athletic club may take exercise classes).

Application

As an application, compared to other data analysis applications, such as structured queries (used in many commercial databases) or statistical analysis software, data mining represents a difference of kind rather than degree. Many simpler analytical tools utilize a verification-based approach, where the user develops a hypothesis and then tests the data to prove or disprove the hypothesis. For example, a user might hypothesize that a customer, who buys a hammer, will also buy a box of nails. The effectiveness of this approach can be limited by the creativity of the user to develop various hypotheses, as well as the structure of the software being used. In contrast, data mining utilizes a discovery approach, in which algorithms can be used to examine several multidimensional data relationships simultaneously, identifying those that are unique or frequently represented. For example, a hardware store may compare their customers' tool purchases with home ownership, type of automobile driven, age, occupation, income, and/or distance between residence and the store. As a result of its complex capabilities, two precursors are important for a successful data mining exercise; a clear formulation of the problem to be solved, and access to the relevant data.

Predicting customer behavior with technology

One of the most important functions performed by a Customer Relationship Management system is to help predict customer behavior. Customer behavior prediction in most CRM consists of the following activities. * Capturing all relevant customer information * Customer behavior modeling * Customer value assessment * Capturing relevant customer information

In the past, information related to a single customer was distributed across the entire company within its different departments. This information has to be harmonized from both the business and technical point of view. Due to fast development of communication development, customer interacts with a company in a number of media such as telephone, fax, e-mail, EDI, Web etc. Hence, the analytical solution must include the ability to flexibility and consistently integrate all data from a variety of customer interactions across all touch points into consolidated view of the customer. This consolidated view should be available to every employee who needs to interact with the customer, should include information from external sources, interaction information captured at different touch points and information from the back office system of the company. CRM solution integrates information from multiple sources to create a consolidated customer view and then make this customer knowledge base available as source data for the numerous CRM analytical applications. The analytical process can then generate new insights on customer behavior with every new customer interaction.

Customer behavior modeling

Customer behavior modeling involves the following activities * Observe customer behavior as depicted by the consolidated customer view contained in the customer knowledge base * Identify relevant behavioral patterns from the observed customer behavior using profiling and scoring techniques * Create predictive models that can be used to acquire, grow and retain attractive and profitable customers.

Figure 1 Customer behavior model

Analytical techniques for customer behavior modeling

Association: Association techniques identify affinities among the collections as reflected in the examined records. These affinities are often expressed as rules. Clustering/Segmentation:Clustering is the method by which like records are grouped together. Usually this is done to give the end-user a high-level view of what is going on in the database. This technique segments records in a database into subsets (or clusters) based on a set of attributes. Clustering can be created either statistically or by using artificial intelligence methods, and can be analyzed automatically by a program or by using visualization techniques. Scoring:Scoring is the process of using a predictive model created from data mining software based on the historical data to make predications about the behavior and the predication is called score. Scoring Engine:Scoring involves a software application often called the scoring Engine that can evaluate mathematical functions on a set of data inputs. The scoring engine takes a predictive model and a dataset and produces a set of scores for the records in the dataset. Scoring process: * A marketing user identifies a segment of customers of interest in the customer database. The records representing the customer segments might be copied into a separate database table. * The selected group of customers is then scored by using a predictive model * The scores are then placed in a database table and then sorted by their score value. * The top x percent are then chosen for promotion. The information necessary for promotion is pulled out of the data warehouse for sending the brochures.

e-business Strategy

The company's e-business strategy comprised of building Internet applications to run e-business, usage of the Internet to build e-business and services and educate employees. The first step which was to formulate an IT strategy which was to be implemented in four steps namely-communication, coordination, business intelligence, and e-commerce, as they believed that the seamless flow of information held the key to a successful business

Communication

As communication was first and the core step with the main agenda being provision of connection at all levels in both office and other locations. As the company firmly believed information to be a strategic business enabler, it is decided to build a robust communication network via LAN, WAN and VPN.

Coordination

The second step was collaboration, which translated into information sharing, discussion database, interactive communication and proper synchronization of workflow applications, setting up a corporate intranet and providing seamless connectivity across the organization. The whole purpose of this e-business solution was driven by the goal of expansion in the market; improve efficiency and lastly retaining customers.

Business intelligence

The business intelligence is the process of integrating its business process re-engineering and the supply chain management process. The company expects not in coordination but has also lowered production costs, through the integration of various functions. The processes are expected to help the company to improve its efficiency-planning and scheduling, procurement, manufacturing, deployment, financial management and internal efficiency.

e-commerce

E-commerce is the last step to make the organization a totally IT savvy one. Once the company has fully set up an e-business model, it plans to use it to implement knowledge management processes and create a knowledge university, which will help individuals who plan to make a career in an organization. It also plans to make the use of B2B website for e-commerce related activities. Other categories of e-commerce activities are C2C - Consumer to Consumer : Transaction between customers C2B - Consumer to Business : Customer transaction to company B2C - Business to Customer : Transaction between company to customer B2B - Business to Business : Transaction between company to company

Why e-commerce?

Business organizations are paying attention in e-commerce because of two reasons * Revenue * Cost The revenue profit and cost are related by the Focus on revenue: Profit =Revenue-cost * Crate new markets for existing off-lien products * Create new application for existing markets * Create new product for the online environment * Improve the quality of existing products i.e. use customer service to attract, serve and keep customers. Focus on Cost: Profit=Revenue-cost * Reduce the cost of delivering existing product/services * To facilitate easier collaboration and information exchange * To improve decision making by standardizing information flow * To coordinate business processes or workflows

Business needs

Create New markets, Distribution, and Sales Channels Create New Products for the Online Environment Improve customer service and support

Integrated structure

Integrated structure consists of three main components * Business data definition * Customer interaction * Analysis The data flows between these components are stage data, build data warehouse and deploy results. The relation between these components are illustrated in Figure 2

Figure 2: System Architecture

In the business data definition component the e-commerce business user defines the data and metadata associated with their business. The data includes merchandising information, content information and business rules. From a data mining perspective the key to the business data definition component is the ability to define a rich set of attributes

(meta data) for any type of data. The customer interaction component provides the interface between customers and the e-commerce business The analysis component provides an integrated environment for decision support utilizing data transformations, reporting, data mining algorithms, and visualization and OLAP tools. (Online Analytical Processing)

Applications to Indian Economy

Business trends:1.Today's businesses must be more profitable, react quicker, and offer higher quality services than ever before, and do it all fewer people and at a lower cost. With these types of expectations and constraints, data mining becomes a fundamental technology in enabling businesses to predict opportunities and risks generated by their customers and their customer's transaction more accurately.2.Given a customer and the product he/she uses, predict whether the customer will switch brands.

Banking Sectors * Detect patterns of fraudulent credit card use * Identifying 'Loyal' Customers * Predicting customers likely to change their credit card affiliation * Determine credit card spending by customer groups * Finding hidden correlation between different financial indicators * Identifying stock trading rules from historical market data

Conclusion

According to the Internet and Online Association(IOA), revenues from online shopping in India was at Rs570 crore in 2004-05 and that is expected t more than double and reach Rs 1180 crore in 2005-06 and climb to Rs2300 crore by 2006-07.Also the average number of transactions per month in India has gone up from 2 lakh in 2003-04 to 4.4. lakh in 2004-05. It is expected to double to 7.95 lakh transactions per month in fiscal 2005-06.Indian customers are increasingly comfortable with online shopping, and there is a higher acceptability of the concept in the forth coming years, The integration effectively solves several major problems associated with horizontal data mining tools including the enormous effort required in preprocessing the data before it can be used for mining and making the results of mining actionable. The tight integration allows for automated construction of data warehouse with the analysis component. The shared metadata across the three components further simplify this construction and coupled with the rich set of data mining algorithm and analysis tools also increases the efficiency of the knowledge discovery process.

References

1. Ralph Kimball, "he data ware housing Tool Kit: Practical Techniques for building dimensional Data Ware houses", John wiely & sons, 1996. 2. Pitkow. J ,"In search of reliable usage data on the WWW", Sixth international World wide Web conference,1997 3. Michael J.A. Berry and Gordan Linoff,"Data mining Techniques: For marketing,Sales and Customer support,Joh wiely & Sons,2000 4. http://www.lon-capa.org 5. Express computer, "IT Business Weekly, September to Novermber Issues.