

# Data Mining For Customer Relationship Management

Savitha S. Kadiyala, (Email: [savitha@gsu.edu](mailto:savitha@gsu.edu)), Georgia State University  
Alok Srivastava, (Email: [alok@gsu.edu](mailto:alok@gsu.edu)), Georgia State University

## Abstract

*Data mining has various applications for customer relationship management. In this article, we introduce a framework for identifying appropriate data mining techniques for various CRM activities. This article attempts to integrate the data mining and CRM models and to propose a new model of Data mining for CRM. The new model specifies which types of data mining processes are suitable for which stages/processes of CRM. In order to develop an integrated model it is important to understand the existing Data mining and CRM models. Hence the article discusses some of the existing data mining and CRM models and finally proposes an integrated model of data mining for CRM.*

## Introduction

Value Creation for the customer is the key determinant of a successful business. Customer satisfaction ensures profitability for businesses in the long run. Customer bases built over a period of time proved to be of immense help in increasing the reach of a particular business's product or service. However, the recent increase in the operating costs of business made it more compelling for businesses to increase loyalty among existing customers while trying to attract new ones. The processes by which an organization creates value for the customer, is often referred to as Customer Relationship Management (CRM).

According to Microsoft, CRM is “a customer-focused business strategy designed to optimize revenue, profitability, and customer loyalty. By implementing a CRM strategy, an organization can improve the business processes and technology solutions around selling, marketing, and servicing functions across all customer touch-points (for example: Web, e-mail, phone, fax, in-person)”. The overall objective of CRM applications is to attract, retain and manage a firm’s profitable (“right”) customers. Electronic support for such activities is provided by the firm’s business intelligence system (Anonymous 2001).

Business intelligence for CRM applications provides a firm with actionable information from the analysis and interpretation of vast quantities of customer/market related data. Databases for business intelligence include customer demographics, buying histories, cross-sales, service calls, website navigation experiences and online transactions. Through the appropriate use of analytical methods and software, a firm is able to turn data into information that leads to greater insight and development of fact-based strategies which in turn helps the firm gain competitive advantage by creating greater value for the customer.

Analogous to traditional mining, which involves searching for an ore in a mountain, data mining involves searching for valuable information in large databases. Both these processes involve either groping through a vast amount of material or intelligently probing the data to find the true value that lies hidden in data. Data mining involves not only the extraction of previously unknown information from a database but also the discovery of relationships that did not surface in the previous methods of data analysis. The “jewels” discovered from the data mining process include these non-intuitive hidden predictive relationships between variables that explain customer behavior and preferences. The predictive capabilities of data mining enable the businesses to make proactive,

---

*Readers with comments or questions are encouraged to contact the authors via email.*

knowledge-driven decisions. Data mining tools facilitate prospective analysis, which is an improvement over the analysis of past events provided by the retrospective tools. The emergence of large data warehouses and the availability of data mining software is creating opportunities for businesses to find innovative ways to implement effective customer relationship strategies.

The automation of data collection and the relative decrease in the costs of operating huge data warehouses has made customer data more accessible than ever. The analysis of data, which until a few years ago was associated with high-end computing power and algorithms decipherable by only professional statisticians, is increasing to become more popular with user-friendly tools available on desktops [Berger, 1999 #2]. Data mining plays an important role in the analytical phases of the CRM life cycle as well as the CRM process.

The purpose of this article is to introduce a framework for identifying appropriate data mining techniques for various CRM activities. This article attempts to integrate the data mining and CRM models and to propose a new model of Data mining for CRM. The new model specifies which types of data mining processes are suitable for which stages/processes of CRM. In order to develop an integrated model, it is important to understand the existing Data mining and CRM models. The next few sections of the article describe some of the existing data mining and CRM models and propose an integrated model of data mining for CRM.

### **Data mining**

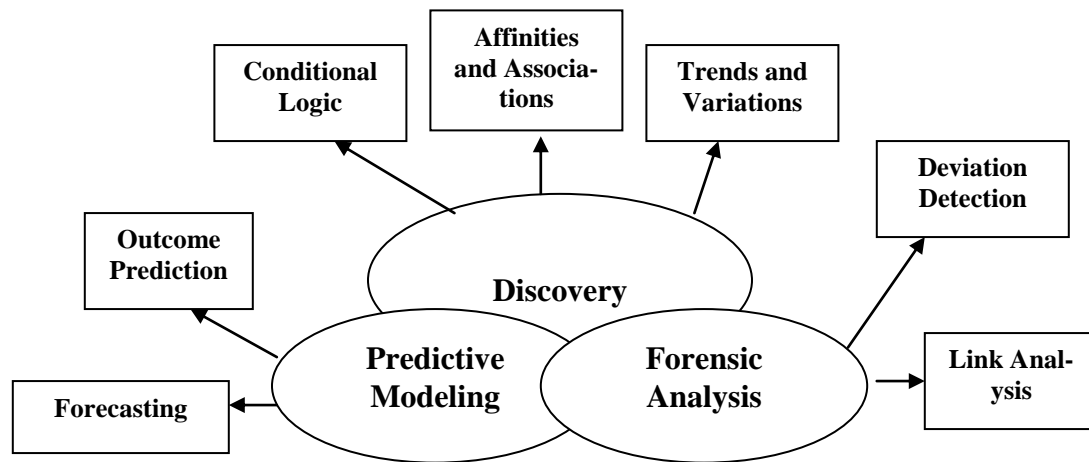
Data mining is "a decision support process in which we search for patterns in data so as to glean previously unknown information" (Parsaye 1997), (Thearling 1999). Such actionable information becomes instrumental in facilitating intelligent decisions to maximize an organization's ability to create value for its customers. In this light, customer data becomes a valuable corporate asset. In the recent years, we have made incredible strides in data warehousing technologies - thus making data assets available to managers for critical decision-making activities. Realizing such value from a firm's data resources will require an effective utilization of data mining techniques.

Business Intelligence, powered with sophisticated data analysis and mining techniques can greatly enhance all phases of customer relationship management. Timely and actionable information is critical in the acquisition and retention of profitable customers, and to maintain high levels of customer satisfaction. Investment in understanding appropriate data mining techniques for various CRM activities and their proper utilization can go a long way in maintaining a successful CRM strategy.

Several typologies have been presented to organize the domain of data mining techniques. Parsaye (Anonymous) proposed a classification of data mining processes on various grounds like the end-users, the functions they perform, the types of algorithms used, and whether they are supervised or unsupervised type of analytical systems. However, the classification of data mining processes based on the functions they perform is more relevant to the context of its application for implementing CRM strategies. From this perspective, data mining processes can be characterized into three types - discovery, predictive modeling, and forensic analysis. Based on Parsaye's classification of data mining processes, we present a more unified typology of Data mining techniques based on their functionality and utility. This typology is described in Figure 1.

Discovery encompasses techniques that explore the data to extract valuable information from the data. It involves the use of techniques like Conditional Logic with IF-THEN rules, affinities, and associations, and identification of trends and variations hidden in the data. Conditional logic allows the algorithm to return specific instances of the database, which satisfy a series of "if" conditions. Affinities and associations within the data can be discovered using clustering and automatic interaction detection algorithms. For example, cluster analysis could be used to identify distinct groups of customers that have similar tastes or preferences.

Figure 1: A Typology of Data Mining Processes.



Similarly, decision tree methods like CART and CHAID classify the data into based on a set of rules. These might reveal some unusual patterns and provide insights into the relationships among the different variables in the data. The classic market basket analysis is a case of association. For example, when customers buy a book on Amazon.com, they get suggestions such as “customers who bought “x” book also bought “y” and “z” books (Greening 2001). Apart from cross-selling opportunities, association between seemingly unrelated items like “paint and cat food” calls for further analysis to differentiate true association from coincidence (Crawford 1996). Data mining also involves the discovery of trends and variations in the data. A typical example of this would be to discover the change in consumer tastes reflected in buying behavior.

The next type of data mining processes can be categorized as the predictive modeling processes, which involve outcome prediction and forecasting. This type of data mining activity follows discovery and generally involves predicting outcomes and forecasting based on previous data patterns. For example, outcome prediction is widely used in predicting response rates of various campaigns. Outcome prediction is very popular in the catalog business and e marketing endeavors where it is important to know which of the customers would respond to which type of campaigns. This helps in better targeting those campaigns to customers who are likely to respond. Forecasting techniques are used to forecast bankruptcy, and other forms of default (Thearling). Data mining can also be used in sales forecasting and inventory management systems. Based on previous data about sales, retail stores can forecast their monthly sales and order inventory accordingly.

Forensic analysis involves deviation detection and performing link analysis to identify the various connections between records (Gerber). Deviation detection builds on previously discovered patterns. Anomalies that surface in the data aid in identifying serious problems like fraud. Forensic analysis is popular in healthcare, credit, telecom, and insurance industries. Deviation detection plays a major role in identifying fraudulent activities in these industries. A set of variables is associated with fraud and when a record in the database matches this profile, the record is flagged as a fraudulent activity. Link analysis enables one to find relationships and associations among different objects, which, in turn helps investigating deviations. Link analysis has applications in law enforcement investigations, pharmaceutical industry, and telecommunications industry (Jensen 1998).

### Customer Relationship Management: A business strategy

Customer relationship management has been defined in various ways. The most comprehensive definitions of CRM are as follows:

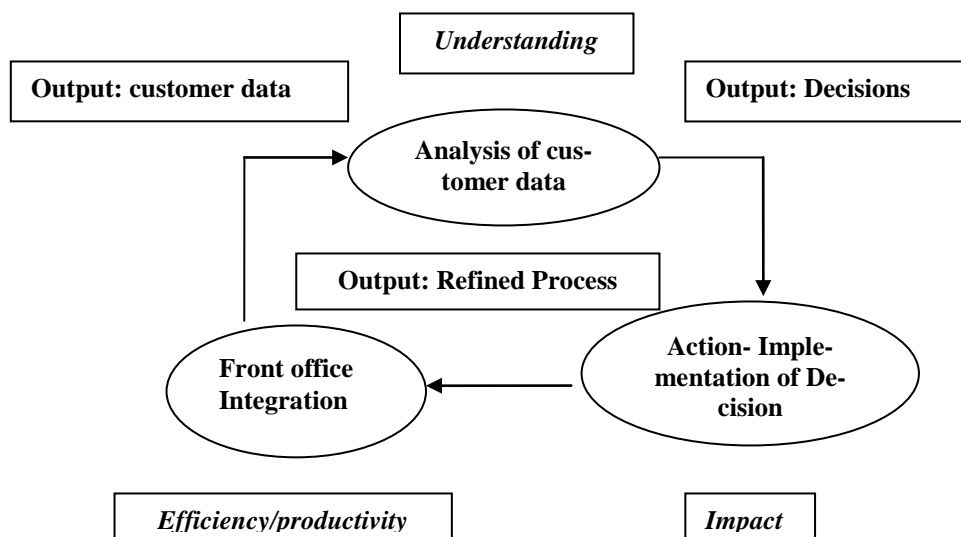
- "...A business strategy, which creates superior value and profitability through a comprehensive and systematic approach to efficiently acquire, collaboratively manage and efficiently retain the right customer." (Sweeney)
- "...A comprehensive approach which provides seamless integration of every area of business that touches the customer - namely, marketing, sales, customer service and field support - through the integration of people, process and technology, taking advantage of the revolutionary impact of the internet" (Chablo 2000).

As indicated by the above definitions, CRM is a business strategy that enables the firm to attract new customers, to retain profitable customers and build loyalty among the existing customers (Kiesnoski 1999). Thus it becomes important for organizations to find systematic ways to identify potential customers, target their marketing campaigns according to customer preferences, and prevent the loss of valuable customers.

According to Reichheld and Sasser (Reichheld 1991), a 5% improvement in a company's customer retention rate will increase its profits by 15% - 50% depending on the industry. Reichheld and Shea (Reichheld 1991) believe that an average company spends five to six times more money to attract a new customer than to keep a present one. Hence, it is more crucial for organizations to retain existing customers. These are some of the factors that explain the popularity of Customer Relationship Management (CRM), which has redefined the way businesses integrate the various functions involving the customer. On the one hand it benefits the customer from meaningful, relevant messages and promotional offers while on the other, it benefits the organizations by achieving a better understanding of the customer (Hill 1999). CRM encompasses all the functions from maintaining call centers to the more analytical functions of finding relationships in the customer data.

Like any other business process, CRM follows a life cycle, and hence it is necessary to recognize the importance of each of the various stages in the life cycle for CRM to succeed. Researchers, mainly from the industry, have developed CRM models to enable us to understand the process of customer relationship management. Hyperion Solutions proposed a model to determine the life cycle of CRM processes (Anonymous).

**Fig. 2 CRM Life Cycle**

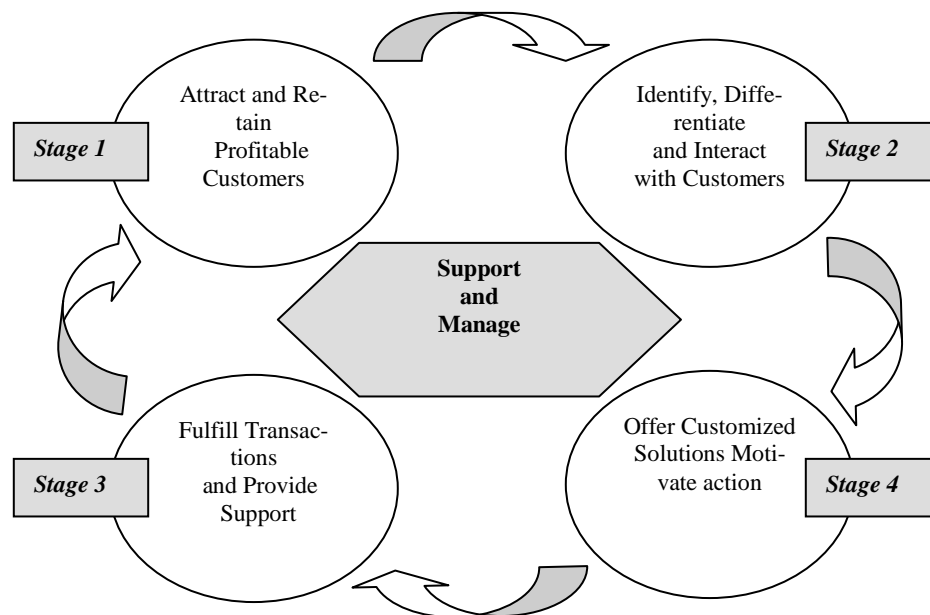


This model of CRM life cycle comprises of three stages. Stage I, the integration stage, consists of integrating the front office systems, and centralizing the customer data. While the output of this stage is a centralized source of all customer data, its benefits include Efficiency and productivity. Stage II in the life cycle is the analysis stage

where customer data are analyzed to understand the patterns in customer behavior. This helps in making decisions about the various strategies to be implemented in order to increase customer profitability. It is in this stage that data mining plays a major role in understanding and predicting the customer behavior. In the final stage, Action, the decisions are implemented which will have an impact on the business and organizational processes. The outcome of this stage is the refinement of these processes based on the improved customer understanding gained through analysis.

Customer Relationship Management Group LLC has modeled the CRM process, which is a detailed description of the manner in which CRM is conducted in organizations. This CRM model comprises of systems that support and manage four types of activities. In the first stage, new customers are attracted and profitable customers are retained. In the second stage, the customer preferences are identified. Customers are differentiated on the basis their preferences and strategies are devised to interact with the customers, recognizing the differences between the various segments of customers. In the next stage, the organization offers customized solutions to suit the needs/preferences of the customers to motivate action on their part. In the final stage the organization fulfills transactions and provides customer support. Although these are the different stages in the CRM process, they are not exclusive. Often there is an overlap among the various stages of the process and the process itself is iterative. For example, an identification of a new segment might trigger attraction strategies and the data from customer support might indicate new insights on whether the solution provided is apt or not.

**Figure 3. A process model for Customer Relationship Management**



Key elements of CRM process include:

- Market segmentation: The ability to classify customers into different groups to be managed differently, either tactically or strategically.
- Predictive marketing: Using data mining capabilities to discover new customers for specific marketing programs and evaluate the responses to offers.
- Campaign planning: Cataloging and organizing high-level campaign data so that multiple users can access what is happening across all campaigns.
- Campaign management: Target customer selections are made and contact lists are created -- all based on the analytical work done and the intelligence generated in the preceding steps of the marketing process.

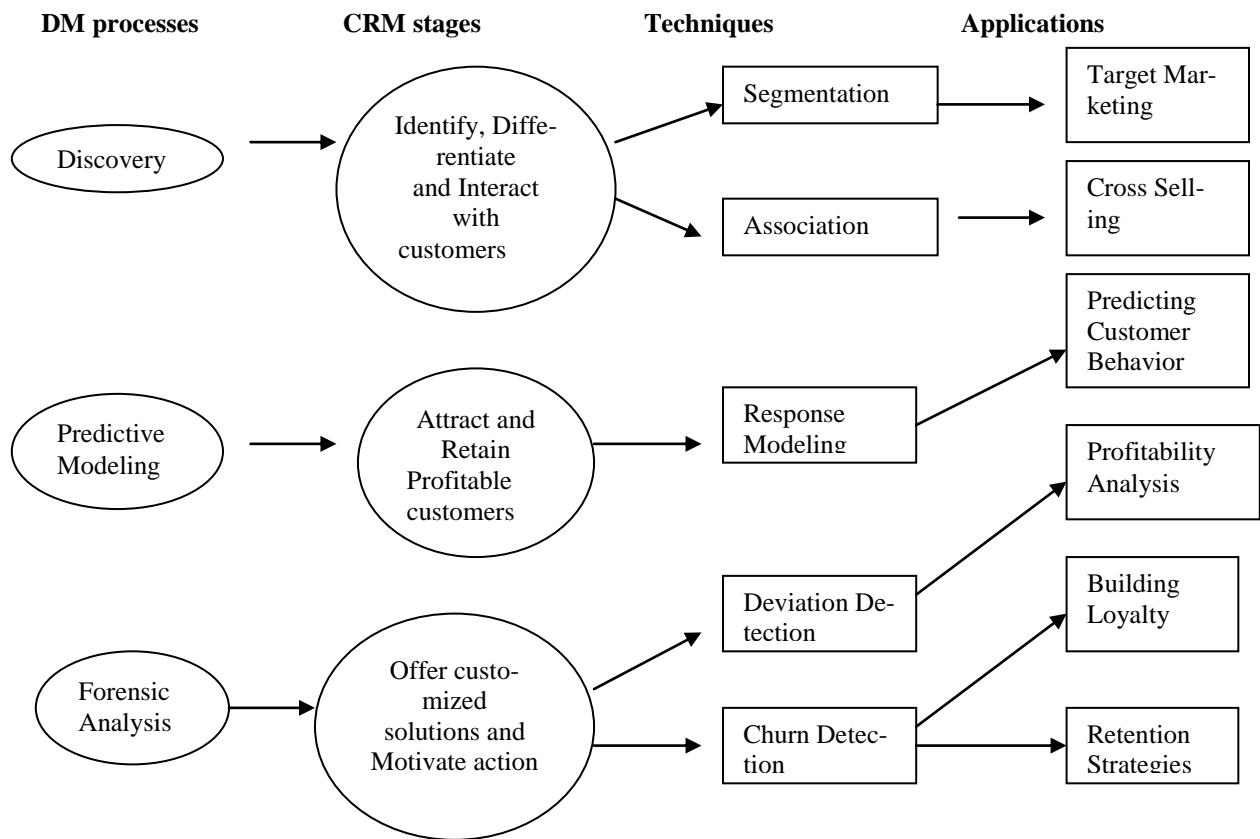
- Campaign review: The final step includes collecting, posting, summarizing, reporting, tracking and trending the results of each campaign and campaign segment (Thearling 2001).

**An Integrated model of Data mining for CRM:**

In this section of the article, an integrated model of data mining for CRM applications is presented, based on inputs from Parsaye's classification of Data mining processes and CRM Group's model of the CRM Process.

The integrated model consists of the three types of data mining processes, three stages of the CRM process, certain specific techniques of data mining that could be used in the various stages of the CRM process and the possible applications of those techniques. The applications of the data mining techniques become the important triggers for making strategic decisions related to the Customer Relationship Management process. The implementation of these actions is carried out in the action stage of the CRM life cycle.

**Fig. 4 Integrated model of Data Mining for CRM**



In the next few subsections of the article, the various data mining techniques that support CRM applications are discussed.

**Segmentation**

The discovery process of data mining aids in identifying new customers, differentiating the customers and finding best ways to interact with the customers. Data mining techniques like segmentation and classification help in achieving these goals. Organizations can gain insights on customer profiles through segmentation and in turn serve them better by offering them the right type of products and services (Anonymous). Through segmentation, customers could be divided into various groups based on their preferences. These segments of customers could be the bases for target marketing strategies, where one segment is targeted for marketing one type of product (Ananyan 2000),(Anonymous 2001), (Lubel 1998).

Target marketing might increase the probability that the customer actually responds to the marketing campaign, since the marketing is based on the knowledge of their preferences. Customer profiles are created and marketing campaigns are personalized for each segment of customers, thereby increasing the probability that they would respond to the campaign (Anonymous).

Or example, the data mining process at Fingerhut Corporation revealed that for customers from a particular zip code, the percentage of expenditure on gold per order was higher than any other product ordered from a particular catalog. This triggered an interest to further analyze the demographics associated with the Zip code. This analysis revealed that the majority of the population residing within the range of that zip code was Hispanic. Following this information Fingerhut revised their catalogs by including more gold jewelry in the catalogs sent out to Hispanic customers (Greening 2001).

AXA Financial uses data mining to gain a better perspective on customer behavior in devising CRM strategies. Analysis at AXA financial already includes measuring customer profitability using associations and calculating earnings in tune with the generally accepted accounting principles. Using data mining techniques enables the organization to segment their customers into profitability clusters. This in turn would help AXA develop various models like cost effective service levels, target marketing campaigns, customer scoring, retention, and also devise cross-selling and up-selling strategies (Davis 2001).

### **Association**

Another data mining technique that can be used to identify new customers is association. The knowledge about the related interests of the customers could provide good grounds for cross selling products to the different segments of customers. The data about a customer consistently revealing preferences to a certain mix of goods or services, provides information about which associated products the customer is more likely to buy or which types of campaigns he or she is more likely to respond to (Lubel 1998). The classic case of market basket analysis is an example for identifying associations. Cross selling is one main application that takes cues from association.

In the context of web data mining, the concept of cross selling has gained more importance in the form of personalized ad campaigns based on available data. For example, an online store, which analyzes the shopping baskets of their customers can better personalize their ad campaigns and increase sales (Bugher 2000). On the other hand, an online store can anticipate what its customers would need and suggest other products.

While association allows organization to cross-sell, it also helps in deciding warehouse layouts, which is a very popular application of data mining with the grocery and retail store chains (Lubel 1998). Apart from aiding in store layout, association also helps in identifying valuable relationships among the data as exemplified by a grocery chain in UK. The analysis of their customer database revealed that a particular product was being consumed by the top-spending 25% of its customers. This led to a decision to retain the product on the shelves instead of discontinuing it (Sethi).

### **Predictive Modeling**

The predictive modeling capability of data mining helps to attract and retain profitable customers. Using techniques like response models, the customer behavior could be predicted with a fair degree of certainty. Based on

the available customer data, response models could be built to predict the response of future customers or the responses of current customers to any new product or service offered (Anonymous 2000). This technique has many applications in the credit, insurance and catalog industries. Even in the case where prior data is not available, it is possible to create profiles of responding customers by studying the population (Cozine 2000). Based on predictions from the response model, a person with a certain mix of characteristics could be classified as either a worthwhile customer or not. If the customer's profile suggests that he or she is a profitable customer, strategies could be devised to attract or to retain him/her. This process of identifying profitable customers would save the organization an enormous amount of expenditures and minimize risks of having bad customers.

Response modeling also aids in more efficient target marketing and saves the finances of an organization (Sethi). Data mining techniques enable the marketing staff to make informed decisions about marketing campaigns without having to wait for the expert statisticians' advice. For example, Liverpool Victoria is a large insurance provider in UK and partnered with Quadstone to provide CRM solutions. Quadstone developed CRM models based on their customer data and made them accessible to the marketers. Now Liverpool Victoria's Marketers have a tool that gave them control of customer profiling, segmentation, and modeling process (Davis 2001). As a result of this effort, Liverpool Victoria is now able to improve retention rates and reduce attrition rates. The database marketers are now empowered to make most of the business decisions on their own without the help of professional statisticians. They are able to verify their business instincts about customer behavior- now more scientifically with the tool available from Quadstone.

C&A is a European fashion retailer, which traditionally outsourced its database marketing operations. When a decision was made to implement CRM strategies in the organization, there was a need for a tool that would integrate the available customer data into the CRM strategy. Once again, with the help of Quadstone's CRM suit, the marketing analysts are now empowered to conduct analyses in-house and C&A has already shown an improvement in the mailing response rates by about 6% (Anonymous 1999). C&A also gained by being able to better target their high responding customers by making marketing offers tailored to their needs. Hence, C&A used the predictive modeling techniques to improve their direct mailing response rates.

### **Deviation Detection**

Deviation detection is probably one of the most important uses of data mining to CRM. Forensic analysis helps one to identify deviations from the norm. For example, it plays a valuable role in identifying true delinquencies and then devising strategies to classify the customer behavior as a delinquent one or as an exception to the behavior pattern. It also has implications for detecting fraud.

With respect to CRM, forensic analysis might reveal an unusual pattern about a particular customer and after conducting a profitability analysis (Anonymous) the organization could determine if retaining the customer is worthwhile. If the customer turns out to be worthwhile, customized solutions could be offered to that customer and motivate him/her to stay with the organization. On the other hand, if the customer does not turn out to be profitable, then the organization could make an informed decision of not employing any retention strategies to retain the unprofitable customer. Fraud detection plays a major role in the health care and credit industries (Sethi), (Anonymous 2000). Deviations detected through data mining could be followed up resulting in huge savings for companies.

### **Churn detection**

In the telecommunication industry churn is a common problem [Berger, 1999 #2], (Lubel 1998), (Anonymous 2000). Churn refers to the process where customers switch to the services of a competitor. Usually in these types of industries, customer acquisition is more expensive than customer retention. If an organization is losing many customers to its competitors, it can analyze the profiles of the lost customers and predict which of the present customers are likely to follow suit. This will enable the organization to devise customer retention strategies and loyalty programs to retain profitable customers. Also, once the churn is identified, it is necessary to perform a profitability analysis to determine if the customer needs to be retained. This would save the organization from investing in retaining a not so profitable customer.



Data mining tools help decision-makers in making informed decisions about their CRM strategies. IBM's Intelligent Miner is one example of the vast array of tools available for understanding customer behavior and formulating CRM strategies. Similarly, SAS's enterprise miner allows the analysis of vast data to detect hidden patterns. Other solution providers include Seibel, MarketSwitch and MarketMiner, which provide target marketing and retention modeling tools for CRM applications (Anonymous).

### **eCRM**


With the increasing power of the Internet and many organizations making their virtual presence felt, CRM has a new direction to focus on. E-commerce solutions are being deployed at a pace like never before and with them is increasing the data about e-customers. Web Data Mining, has gained popularity with many Internet companies to better target their marketing efforts and build loyal customer bases (Mena 1999). This is within the purview of eCRM. Web data mining follows the traditional data mining processes, performing functions like segmentation, association, and response prediction. The data for such types of analyses could be collected from the log file, cookies, and forms. The log files and cookies collect information about the customers without their knowledge, while the forms collect more specific information with the knowledge and cooperation of customers. Click stream analysis aids in keeping track of the sequence of the web pages visited and drawing conclusions about customer preferences.

Web data mining technique is very popular today with e-businesses. eCRM not only involves the analysis of the customer data but also targets improving customer relationships by improving service levels and improving customer support. Organizations like Kana Communications offer eCRM solutions for increased productivity and improved customer support. The famous retailer/auction site eBay is one of Kana's customers. Its success story includes a huge improvement in productivity, decrease in additional hiring, and improved decision making with the help of better reporting and analysis tools. Kana Response, a program that provides templates for e-mail responses- has dramatically reduced the time taken by the eBay's customer service representatives to respond to online queries. This in turn allowed eBay to provide 24-hour turnarounds on most of the online queries (Anonymous 2000). Kana's Response also tracks all contact with the customer, which provides for detailed information on all the transactions with the customer, thereby improving decision-making.

### **Pattern Management**

The data generated about by an online store is very huge compared to a brick and mortar store or a catalog store. One of the solutions developed by the industry to deal with the data explosion problem is pattern management, which is analogous to data base management. Pattern management involves extracting patterns from the data and storing them in a pattern base. This allows the organizations to discard the multitude of historical data after extracting patterns from it. New data are matched against the existing patterns to look for consistency. If there is a deviation from the existing patterns, then a new pattern could be extracted to be stored in the pattern base. Pattern Query Language (PQL), similar to Structured query language (SQL) has been developed to query specific patterns from this pattern base and applied to the data available on hand.

### **Summary**

The article defined the concepts of data mining and CRM, explained the current models of CRM. Parsaye's classification of data mining processes was used as a base to identify the place of data mining technologies in the CRM process. The new model of data mining for CRM proposed in this article, is an integration of the existing models of data mining and CRM, with a reference to the various techniques that could be adopted as well as the various applications of the results of the analysis. 

### **References**

1. Ananyan, S. (2000). Data Mining for Direct Marketing. *DM Review*.

2. Anonymous Analysis for Banking Channel Strategy.
3. Anonymous AXA Financial Strikes Gold with SAS® Enterprise Miner™, SAS.
4. Anonymous Bank of New Zealand Global Plus Home Loan.
5. Anonymous The CRM Life Cycle, Hyperion Solutions.
6. Anonymous CRM: Customer Relationship Management Solutions, kdnuggets.com.
7. Anonymous (1999). Quadstone's Decisionhouse Increases Profitability Of C&A's Database Marketing Program. *DSSStar*. 3.
8. Anonymous (2000). Data mining: Using Data Mining Techniques for Fraud Detection.
9. Anonymous (2000). KANA success stories-eBAY.
10. Anonymous (2000). Mining New Markets in Telecommunications.
11. Anonymous (2000). Reducing Customer Churn: Customer Application Story.
12. Anonymous (2001). Customer Relationship Management Basics.
13. Anonymous (2001). Datamining Increases Campaign Results 50%.
14. Bugher, G. (2000). Market Basket Analysis of the Sales Data for a Client of Cambridge Technology Partners.
15. Chablo, E. (2000). The Importance of Marketing Data Intelligence In Delivering Successful CRM, to CRM-forum.
16. Cozine, C. A. (2000). Cactus Strategies uses PolyAnalyst to predict customer response for a communications client.
17. Crawford, J. a. C. F. (1996). *Data Mining in a Scientific Environment*. AUUG 96 & Asia Pacific World Wide Web, Australia.
18. Davis, C. (2001). QUADSTONE Insures Liverpool Victoria CRM Strategy With Scalable Technology.
19. Gerber, C. Dissecting data mining, Datamation.
20. Greening, D. R. (2001). Data Mining on the Web: There's Gold in that Mountain of Data. *Webtechniques*.
21. Hill, L. (1999). CRM: Easier Said Than Done. *Intelligent Enterprise Magazine*. 2.
22. Jensen, D. a. G., Henry (1998). *Artificial Intelligence and Link Analysis*. AAAI.
23. Kiesnoski, K. (1999). Customer Relationship Management. *Bank Systems & Technology*.
24. Lubel, K. S. (1998). Data Mining: A New Way to Find Answers, University of Maryland European Division.
25. Mena, J. (1999). Mining E-Customer Behavior. *DB2 Magazine*.
26. Parsaye, K. (1997). "A Characterization of Data Mining Technologies and Processes." *Journal of Data Warehousing*.
27. Reichheld, F., F., and Shea, Mary Beth (1991). *Customer Retention Strategies*. The Planning Forum's International Conference.
28. Reichheld, F. F., and Sasser, W., Earl (1991). Zero-Defections: Quality Comes to Service. *Harvard Business Review*.
29. Sethi, I. K. Data Mining: An Introduction: 82.
30. Sweeney, J. Customer Relationship Management.
31. Thearling, K. An Introduction to Data Mining: Discovering hidden value in your data warehouse.
32. Thearling, K. (1999). Data Mining and CRM: Zeroing in on Your Best Customer. *DM Direct*.
33. Thearling, K. (2001). Campaign Management.