



Pergamon

Technology in Society 24 (2002) 483–502

Technology
In Society

www.elsevier.com/locate/techsoc

Data mining techniques for customer relationship management

Chris Rygielski^a, Jyun-Cheng Wang^b, David C. Yen^{a,*}

^a *Department of DSC & MIS, Miami University, Oxford, OH, USA*

^b *Department of Information Management, National Chung-Cheng University, Taiwan, ROC*

Abstract

Advancements in technology have made relationship marketing a reality in recent years. Technologies such as data warehousing, data mining, and campaign management software have made customer relationship management a new area where firms can gain a competitive advantage. Particularly through data mining—the extraction of hidden predictive information from large databases—organizations can identify valuable customers, predict future behaviors, and enable firms to make proactive, knowledge-driven decisions. The automated, future-oriented analyses made possible by data mining move beyond the analyses of past events typically provided by history-oriented tools such as decision support systems. Data mining tools answer business questions that in the past were too time-consuming to pursue. Yet, it is the answers to these questions make customer relationship management possible. Various techniques exist among data mining software, each with their own advantages and challenges for different types of applications. A particular dichotomy exists between neural networks and chi-square automated interaction detection (CHAID). While differing approaches abound in the realm of data mining, the use of some type of data mining is necessary to accomplish the goals of today's customer relationship management philosophy.

© 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Customer relationship management (CRM); Relationship marketing; Data mining; Neural networks; Chi-square automated interaction detection (CHAID); Privacy rights

* Corresponding author. Tel.: +1-513-529-4826; fax: +1-513-529-9689.

E-mail address: yendc@muohio.edu (D.C. Yen).

1. Introduction

A new business culture is developing today. Within it, the economics of customer relationships are changing in fundamental ways, and companies are facing the need to implement new solutions and strategies that address these changes. The concepts of mass production and mass marketing, first created during the Industrial Revolution, are being supplanted by new ideas in which customer relationships are the central business issue. Firms today are concerned with increasing customer value through analysis of the customer lifecycle. The tools and technologies of data warehousing, data mining, and other customer relationship management (CRM) techniques afford new opportunities for businesses to act on the concepts of relationship marketing. The old model of “design-build-sell” (a product-oriented view) is being replaced by “sell-build-redesign” (a customer-oriented view). The traditional process of mass-marketing is being challenged by the new approach of one-to-one marketing. In the traditional process, the marketing goal is to reach more customers and expand the customer base. But given the high cost of acquiring new customers, it makes better sense to conduct business with current customers. In so doing, the marketing focus shifts away from the breadth of customer base to the depth of each customer’s needs. The performance metric changes from market share to so-called “wallet share”. Businesses do not just deal with customers in order to make transactions; they turn the opportunity to sell products into a service experience and endeavor to establish a long-term relationship with each customer.

The advent of the Internet has undoubtedly contributed to the shift of marketing focus. As on-line information becomes more accessible and abundant, consumers become more informed and sophisticated. They are aware of all that is being offered, and they demand the best. To cope with this condition, businesses have to distinguish their products or services in a way that avoids the undesired result of becoming mere commodities. One effective way to distinguish themselves is with systems that can interact precisely and consistently with customers. Collecting customer demographics and behavior data makes precision targeting possible. This kind of targeting also helps when devising an effective promotion plan to meet tough competition or identifying prospective customers when new products appear. Interacting with customers consistently means businesses must store transaction records and responses in an on-line system that is available to knowledgeable staff members who know how to interact with it. The importance of establishing close customer relationships is recognized, and CRM is called for.

It may seem that CRM is applicable only for managing relationships between businesses and consumers. A closer examination reveals that it is even more crucial for business customers. In business-to-business (B2B) environments, a tremendous amount of information is exchanged on a regular basis. For example, transactions are more numerous, custom contracts are more diverse, and pricing schemes are more complicated. CRM helps smooth the process when various representatives of seller and buyer companies communicate and collaborate. Customized catalogues, personalized business portals, and targeted product offers can simplify the procurement process and improve efficiencies for both companies. E-mail alerts and new

product information tailored to different roles in the buyer company can help increase the effectiveness of the sales pitch. Trust and authority are enhanced if targeted academic reports or industry news are delivered to the relevant individuals. All of these can be considered among the benefits of CRM.

Cap Gemini conducted a study to gauge company awareness and preparation of a CRM strategy [1]. Of the firms surveyed, 65% were aware of CRM technology and methods; 28% had CRM projects under study or in the implementation phase; 12% were in the operational phase. In 45% of the companies surveyed, implementation and monitoring of the CRM project had been initiated and controlled by top management. Thus, it is apparent that this is a new and emerging concept that is seen as a key strategic initiative.

This article examines the concepts of customer relationship management and one of its components, data mining. It begins with an overview of the concepts of data mining and CRM, followed by a discussion of evolution, characteristics, techniques, and applications of both concepts. Next, it integrates the two concepts and illustrates the relationship, benefits, and approaches to implementation, and the limitations of the technologies. Through two studies, we offer a closer look at two data mining techniques: Chi-square Automatic Interaction Detection (CHAID) and Neural Networks. Based on those case studies, CHAID and neural networks are compared and contrasted on the basis of their strengths and weaknesses. Finally, we draw conclusions based on the discussion.

2. Data mining: an overview

2.1. Definition

“Data mining” is defined as a sophisticated data search capability that uses statistical algorithms to discover patterns and correlations in data [2]. The term is an analogy to gold or coal mining; data mining finds and extracts knowledge (“data nuggets”) buried in corporate data warehouses, or information that visitors have dropped on a website, most of which can lead to improvements in the understanding and use of the data. The data mining approach is complementary to other data analysis techniques such as statistics, on-line analytical processing (OLAP), spreadsheets, and basic data access. In simple terms, data mining is another way to find meaning in data.

Data mining discovers patterns and relationships hidden in data [3], and is actually part of a larger process called “knowledge discovery” which describes the steps that must be taken to ensure meaningful results. Data mining software does not, however, eliminate the need to know the business, understand the data, or be aware of general statistical methods. Data mining does not find patterns and knowledge that can be trusted automatically without verification. Data mining helps business analysts to generate hypotheses, but it does not validate the hypotheses.

2.2. The evolution of data mining

Data mining techniques are the result of a long research and product development process. The origin of data mining lies with the first storage of data on computers, continues with improvements in data access, until today technology allows users to navigate through data in real time. In the evolution from business data to useful information, each step is built on the previous ones. Table 1 shows the evolutionary stages from the perspective of the user.

In the first stage, Data Collection, individual sites collected data used to make simple calculations such as summations or averages. Information generated at this step answered business questions related to figures derived from data collection sites, such as total revenue or average total revenue over a period of time. Specific application programs were created for collecting data and calculations.

The second step, Data Access, used databases to store data in a structured format. At this stage, company-wide policies for data collection and reporting of management information were established. Because every business unit conformed to specific requirements or formats, businesses could query the information system regarding branch sales during any specified time period.

Once individual figures were known, questions that probed the performance of aggregated sites could be asked. For example, regional sales for a specified period could be calculated. Thanks to multi-dimensional databases, a business could obtain either a global view or drill down to a particular site for comparisons with its peers (Data Navigation). Finally, on-line analytic tools provided real-time feedback and information exchange with collaborating business units (Data Mining). This capa-

Table 1
Evolutionary stages of data mining

Stage	Business question	Enabling technologies	Product providers	Characteristics
Data Collection (1960s)	“What was my average total revenue over the last five years?”	Computers, tapes, disks	IBM, CDC	Retrospective, static data delivery
Data Access (1980s)	“What were unit sales in New England last March?”	Relational databases (RDBMS), Structured Query Language (SQL), ODBC	Oracle, Sybase, Informix, IBM, Microsoft	Retrospective, dynamic data delivery at record level
Data Navigation (1990s)	“What were unit sales in New England last March? Drill down to Boston”	On-line analytic processing (OLAP), multidimensional databases, data warehouses	Pilot, IRI, Arbor, Redbrick, Evolutionary Technologies	Retrospective, dynamic data delivery at multiple levels
Data Mining (2000)	“What’s likely to happen in Boston unit sales next month? Why?”	Advanced algorithms, multiprocessor computers, massive databases	Lockheed, IBM, SGI, numerous startups (nascent industry)	Prospective, proactive information delivery

Source: Pilot Software [17].

bility is useful when sales representatives or customer service persons need to retrieve customer information on-line and respond to questions on a real-time basis.

Information systems can query past data up to and including the current level of business. Often businesses need to make strategic decisions or implement new policies that better serve their customers. For example, grocery stores redesign their layout to promote more impulse purchasing. Telephone companies establish new price structures to entice customers into placing more calls. Both tasks require an understanding of past customer consumption behavior data in order to identify patterns for making those strategic decisions—and data mining is particularly suited to this purpose. With the application of advanced algorithms, data mining uncovers knowledge in a vast amount of data and points out possible relationships among the data. Data mining help businesses address questions such as, “What is likely to happen to Boston unit sales next month, and why?” Each of the four stages were revolutionary because they allowed new business questions to be answered accurately and quickly [4].

The core components of data mining technology have been developing for decades in research areas such as statistics, artificial intelligence, and machine learning. Today, these technologies are mature, and when coupled with relational database systems and a culture of data integration, they create a business environment that can capitalize on knowledge formerly buried within the systems.

2.3. Applications of data mining

Data mining tools take data and construct a representation of reality in the form of a model. The resulting model describes patterns and relationships present in the data. From a process orientation, data mining activities fall into three general categories (see Fig. 1):

- **Discovery**—the process of looking in a database to find hidden patterns without a predetermined idea or hypothesis about what the patterns may be.

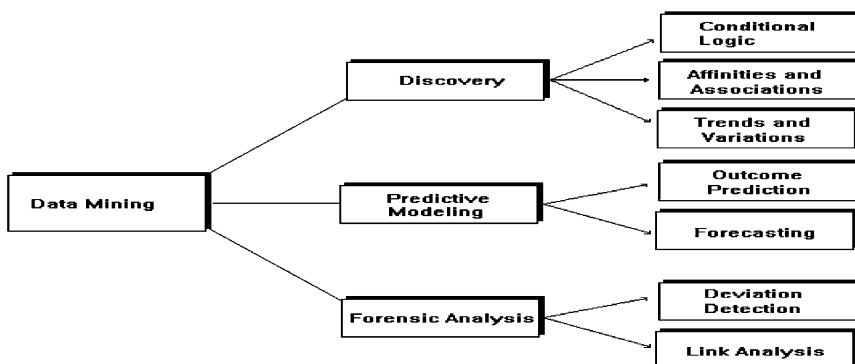


Fig. 1. Breakdown of data mining from a process orientation. Source: Information Discovery, Inc. [18].

- **Predictive Modeling**—the process of taking patterns discovered from the database and using them to predict the future.
- **Forensic Analysis**—the process of applying the extracted patterns to find anomalous or unusual data elements.

Data mining is used to construct six types of models aimed at solving business problems: classification, regression, time series, clustering, association analysis, and sequence discovery [3]. The first two, classification and regression, are used to make predictions, while association and sequence discovery are used to describe behavior. Clustering can be used for either forecasting or description.

Companies in various industries can gain a competitive edge by mining their expanding databases for valuable, detailed transaction information. Examples of such uses are provided below.

Each of the four applications below makes use of the first two activities of data mining: discovery and predictive modeling. The discovery process, while not mentioned explicitly in the examples (except in the retail description), is used to identify customer segments. This is done through conditional logic, analysis of affinities and associations, and trends and variations. Each of the application categories described below describes some sort of predictive modeling. Each business is interested in predicting the behavior of its customers through the knowledge gained in data mining [5].

2.3.1. Retail

Through the use of store-branded credit cards and point-of-sale systems, retailers can keep detailed records of every shopping transaction. This enables them to better understand their various customer segments. Some retail applications include [5]:

- *Performing basket analysis*—Also known as affinity analysis, basket analysis reveals which items customers tend to purchase together. This knowledge can improve stocking, store layout strategies, and promotions.
- *Sales forecasting*—Examining time-based patterns helps retailers make stocking decisions. If a customer purchases an item today, when are they likely to purchase a complementary item?
- *Database marketing*—Retailers can develop profiles of customers with certain behaviors, for example, those who purchase designer labels clothing or those who attend sales. This information can be used to focus cost-effective promotions.
- *Merchandise planning and allocation*—When retailers add new stores, they can improve merchandise planning and allocation by examining patterns in stores with similar demographic characteristics. Retailers can also use data mining to determine the ideal layout for a specific store.

2.3.2. Banking

Banks can utilize knowledge discovery for various applications, including [5]:

- *Card marketing*—By identifying customer segments, card issuers and acquirers

can improve profitability with more effective acquisition and retention programs, targeted product development, and customized pricing.

- *Cardholder pricing and profitability*—Card issuers can take advantage of data mining technology to price their products so as to maximize profit and minimize loss of customers. Includes risk-based pricing.
- *Fraud detection*—Fraud is enormously costly. By analyzing past transactions that were later determined to be fraudulent, banks can identify patterns.
- *Predictive life-cycle management*—Data mining helps banks predict each customer's lifetime value and to service each segment appropriately (for example, offering special deals and discounts).

2.3.3. Telecommunications

Telecommunication companies around the world face escalating competition which is forcing them to aggressively market special pricing programs aimed at retaining existing customers and attracting new ones. Knowledge discovery in telecommunications include the following [5]:

- *Call detail record analysis*—Telecommunication companies accumulate detailed call records. By identifying customer segments with similar use patterns, the companies can develop attractive pricing and feature promotions.
- *Customer loyalty*—Some customers repeatedly switch providers, or “churn”, to take advantage of attractive incentives by competing companies. The companies can use data mining to identify the characteristics of customers who are likely to remain loyal once they switch, thus enabling the companies to target their spending on customers who will produce the most profit.

2.3.4. Other applications

Knowledge discovery applications are emerging in a variety of industries [5]:

- *Customer segmentation*—All industries can take advantage of data mining to discover discrete segments in their customer bases by considering additional variables beyond traditional analysis.
- *Manufacturing*—Through choice boards, manufacturers are beginning to customize products for customers; therefore they must be able to predict which features should be bundled to meet customer demand.
- *Warranties*—Manufacturers need to predict the number of customers who will submit warranty claims and the average cost of those claims.
- *Frequent flier incentives*—Airlines can identify groups of customers that can be given incentives to fly more.

In the application examples discussed above, the use of forensic analysis was not as common. The banking example is the only one that was looking for deviations in the data. Banks and other financial institutions use data mining for fraud detection, which was not alluded to in the other examples even though there are similar uses of deviation detection in the other industries.

2.4. Internal considerations

For firms to integrate data mining into their decision-making process, the proper skillsets and technology must be available. Skillsets will vary with the variety of data-mining stakeholders in the organization (see Table 2). While data mining is frequently done centrally or regionally, people on the front lines need to have the knowledge gained through data mining. These workers sell to and service customers, manage inventory, supervise employees, and work to correct and prevent loss. Information derived from data mining can be communicated to operational employees in several forms:

- an algorithm for scoring
- a score for a particular customer, employee, or transaction
- a recommended action associated with a particular customer, employee, or transaction. [6].

2.5. Data mining techniques

A top-level breakdown of data mining technologies is based on data retention. In other words, is the data retained or discarded after it has been mined? (see Fig. 2). In early approaches to data mining, the data set was maintained for future pattern matching. The retention-based techniques only apply to tasks of predictive modeling and forensic analysis, and not knowledge discovery since they do not distill any patterns, as shown earlier in Fig. 1.

Approaches based on pattern distillation fall into three categories: logical, cross-tabulation, and equational. These technologies extract patterns from a data set and then use the patterns for various purposes. They ask, “What types of patterns can be extracted and how are they represented?” The logical approach deals with both numeric and non-numeric data. Equations require all data to be numeric, while cross-tabulations work only on non-numeric data. Table 3 summarizes the pros and cons of these categories.

Table 2
User interactions with data mining technology and the user’s typical skillset

Stakeholder	Skill set
Miner	Analytics, model building, statistics, neural net development, research
Domain expert	Intensive business and data knowledge, experience, decision maker
Business user	Understands business and data, decision maker, user of mining results
IT	Supports analytic environment, data model for new DM components, integrates DM (tools, processes, results, models) into DW

Source: Cranford [19].

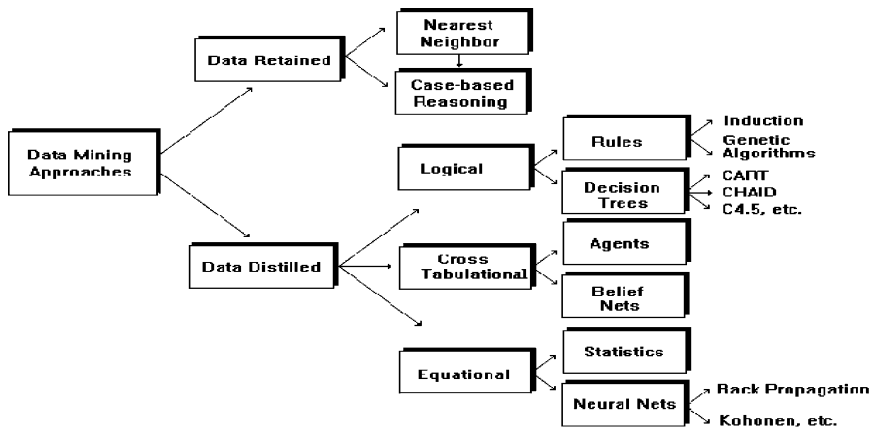


Fig. 2. Breakdown of data mining technologies based on retention of data. Source: Information Discovery, Inc. [18].

Table 3
Pros and cons to data mining approaches

Approach	Pros	Cons
Logical	Work well with multidimensional and OLAP data Able to deal with numeric and nonnumeric data in a uniform manner Able to deal with numeric and nonnumeric data in a uniform manner	Unable to work with smooth surfaces that typically occur in nature
Cross-tabulation	Simple to use with small number of nonnumeric values	Not scalable
Equational	Works well on large sets of data Works well with complex multi-dimensional models Ability to approximate smooth surfaces	Ability to handle numeric values Ability to handle conjunctions Require all data to be numeric (nonnumeric must be coded) System can quickly become a “black box”

Source: Information Discovery, Inc. [18].

3. Customer relationship management: an overview

3.1. Definition

Customer Relationship Management is defined by four elements of a simple framework: Know, Target, Sell, Service [7]. CRM requires the firm to know and understand its markets and customers. This involves detailed customer intelligence in order

to select the most profitable customers and identify those no longer worth targeting. CRM also entails development of the offer: which products to sell to which customers and through which channel. In selling, firms use campaign management to increase the marketing department's effectiveness. Finally, CRM seeks to retain its customers through services such as call centers and help desks.

CRM is essentially a two-stage concept. The task of the first stage is to master the basics of building customer focus. This means moving from a product orientation to a customer orientation and defining market strategy from outside-in and not from inside-out. The focus should be on customer needs rather than product features. Companies in the second stage are moving beyond the basics; they do not rest on their laurels but push their development of customer orientation by integrating CRM across the entire customer experience chain, by leveraging technology to achieve real-time customer management, and by constantly innovating their value proposition to customers [7].

3.2. Components of customer relationship management

Customer relationship management is a combination of several components. Before the process can begin, the firm must first possess customer information. Companies can learn about their customers through internal customer data or they can purchase data from outside sources. There are several sources of internal data:

- summary tables that describe customers (e.g., billing records)
- customer surveys of a subset of customers who answer detailed questions
- behavioral data contained in transactions systems (web logs, credit card records, etc). [8].

An enterprise data warehouse is a critical component of a successful CRM strategy. Most firms have massive databases that contain marketing, HR, and financial information. However, the data required for CRM can be limited to a marketing data mart with limited feeds from other corporate systems. Experience with CRM will dictate when to aggregate data for reasons of simplicity and when to keep the data granular. External sources of data or purchased databases can be a key source for gaining customer knowledge advantage [9]. Some examples of external data sources include lookups for current address and telephone number, household hierarchies, Fair-Isaacs credit scores, and Webpage viewing profiles [8].

Next, the CRM system must analyze the data using statistical tools, OLAP, and data mining. Whether the firm uses traditional statistical techniques or one of the data mining software tools, marketing professionals need to understand the customer data and business imperatives. The firm should employ data mining analysts who will be involved but will also make sure the firm does not lose sight of their original reason for doing data mining. Thus, having the right people who are trained to extract information with these tools is also important. The end result is segmentation of the market, and individual decisions are made regarding which segments are attractive [9].

The last component of a CRM system is campaign execution and tracking. These are the processes and systems that allow the user to develop and deliver targeted messages in a test-and-learn environment. Implementation of decisions made as a result of data mining and OLAP is done through campaign execution and tracking. Today there are software programs that help marketing departments handle this complex feedback procedure. Campaign management software manages and monitors customer communications across multiple touchpoints, such as direct mail, telemarketing, customer service, point-of-sale, e-mail, and the Web [10]. While campaign management software may be part of the overall solution, it is primarily the people and processes that contribute to smooth interactions between marketing, information technology, and the sales channels [9].

4. Data mining and customer relationship management

It should be clear from the discussion so far that customer relationship management is a broad topic with many layers, one of which is data mining, and that data mining is a method or tool that can aid companies in their quest to become more customer-oriented. Now we need to step back and see how all the pieces fit together.

4.1. *The relationship*

The term “customer lifecycle” refers to the stages in the relationship between a customer and a business. It is important to understand customer lifecycle because it relates directly to customer revenue and customer profitability. Marketers say there are three ways to increase a customer’s value: (1) increase their use (or purchases) of products they already have; (2) sell them more or higher-margin products; and (3) keep the customers for a longer period of time [8].

However, the customer relationship changes over time, evolving as the business and the customer learn more about each other. So why is the customer lifecycle important? Simply put, it is a framework for understanding customer behavior. In general, there are four key stages in the customer lifecycle:

1. **Prospects**—people who are not yet customers but are in the target market
2. **Responders**—prospects who show an interest in a product or service
3. **Active Customers**—people who are currently using the product or service
4. **Former Customers**—may be “bad” customers who did not pay their bills or who incurred high costs; those who are not appropriate customers because they are no longer part of the target market; or those who may have shifted their purchases to competing products.

The customer lifecycle provides a good framework for applying data mining to CRM. On the “input” side of data mining, the customer lifecycle tells what information is available. On the “output” side, the customer lifecycle tells what is likely to be interesting [8].

Looking first at the input side, there is relatively little information about prospects except what is learned through data purchased from outside sources. There are two exceptions: one, there are more prospecting data warehouses in various industries that track acquisition campaigns directed at prospects; two, click-stream information is available about prospects' behavior on some websites. Data mining can predict the profitability of prospects as they become active customers, how long they will be active customers, and how likely they are to leave [8].

In addition, data mining can be used over a period of time to predict changes in details. It will not be an accurate predictor of when most lifecycle events occur. Rather, it will help the organization identify patterns in their customer data that are predictive. For example, a firm could use data mining to predict the behavior surrounding a particular lifecycle event (e.g., retirement) and find other people in similar life stages and determine which customers are following similar behavior patterns [8].

The outcome of this process is marketing data intelligence, which is defined as "Combining data driven marketing and technology to increase the knowledge and understanding of customers, products and transactional data to improve strategic decision making and tactical marketing activity, delivering the CRM challenge" [11]. There are two critical components of marketing data intelligence: customer data transformation and customer knowledge discovery. Raw data extracted and transformed from a wide array of internal and external databases, marts or warehouses and the collecting of that total data into a centralized place where it can be accessed and explored is data transformation. The process is continued through customer knowledge discovery, where the information is mined, and usable patterns and inferences can be drawn from the data. The process must be measured and tracked to ensure that the results fed to campaign management software produce information that the models created by data mining software find useful and accurate [11].

Data mining plays a critical role in the overall CRM process, which includes interaction with the data mart or warehouse in one direction, and interaction with campaign management software in the other direction. In the past, the link between data mining software and campaign management software was mostly manual. It required that physical copies of the scoring from data models be created and transferred to the database. This separation of data mining and campaign management software introduced considerable inefficiency and was prone to human error. Today the trend is to integrate the two components in order to gain a competitive advantage [12].

Firms can gain a competitive advantage by ensuring that their data mining software and campaign management software share the same definition of the customer segment in order to avoid modeling the entire database. For instance, if the ideal segment is high-income males between the ages of 25 and 35 living in the northeast, the analysis should be restricted to just those characteristics. In addition, the selected scores from the predictive model should flow directly into the campaign segment in order to form targets with the highest profit potential [13].

4.2. *Data mining and customer privacy*

While data mining techniques help businesses address more questions than ever before, this capability may add to the risk of invading customer privacy. On one hand, mining customer data can help build an intimate relationship between businesses and their customers. On the other, databases can be used against customers' wishes or to their detriment. However, personalization of CRM is far from invasion of an individual's privacy. Personal information collected by businesses can be classified in two categories: data that are provided and accessible to the users, and data that are generated and analyzed by businesses. Before data mining became popular among businesses, customers' data was generally collected on a self-provided or transactional basis. Customers themselves provide general descriptive data which contains demographic data about themselves. Transactional data refers to data obtained when a transaction takes place, such as product name, quantity, location, and time of purchase. These data are collected from registration forms, order forms, computer cookies, log files, surveys, and contests.

The power of data mining helps turn customer data into customer profiling information. This kind of information belongs to the second category and is accessible to businesses, although this fact may not be known to consumers. It may include customer value, customer targeting information, customer rating, and behavior tracking. Once this information is obtained by marketers or businesses, consumers may periodically receive timely and personalized information. However, when abused, people may also suffer from certain forms of discrimination (such as insurance) or loss of career. Without proper scrutiny when applying and releasing profiling information, consumers may turn away from any effort to maintain a closer customer relationship. The central issue of privacy is to find a balance between privacy rights for consumer protection and while still providing benefits to businesses. Several advocacy groups and private efforts have been formed to promote the responsible use of technology for personalizing consumer and business relationships.

However, privacy is more of a policy issue than a technology one. One basic principle for businesses using personalized technology is to disclose to their consumers the kinds of information they are seeking and how that information will be used. Some groups list objectives for ethical information and privacy management. Others have developed a Privacy Bill of Rights that includes fair access by individuals to their personal information, responsible linkage of online and off-line information, suitable criteria for opt-in and opt-out privacy options, standardizing the disclosure to consumers of any existing privacy policy, independent verification of implementation and execution of privacy and security policies, and fair mechanisms for resolving disputes by a trusted third party.

Customer privacy can be better protected when customers do not have to reveal their identities and can remain anonymous even after data mining probing. One way to achieve this goal is to create an anonymous architecture for handling customer information. In this architecture, identity information is processed with an additional encryption procedure whenever data are fed into a data mining module for analysis. The encrypted identity information remains unique for each individual but does not

diminish the power of data mining while keeping the customer's identity information protected under a firewall. Some third-party organizations also take responsibility for handling identity information, becoming a surrogate for executing targeted marketing efforts, such as mail promotion messages to the targeted individual.

5. Case studies

We have chosen two particular data mining techniques, CHAID and neural nets, and will illustrate their use through two case studies. There are various data mining tool providers in the marketplace today, and each provider has a different combination of data mining tools that can be used to help their clients. There are no instances where one provider chose to use only one data mining technique; to the contrary, providers often choose a group of similar methods for accomplishing their goals.

In this section we examine neural networks through NeoVista Solutions Inc., and CHAID through Applied Metrix, Ltd. Both case studies come from the respective websites of NeoVista and Applied Metrix, and the identities of their clients are withheld at the client's request.

5.1. Case study: neural networks

NeoVista Solutions, Inc. provides comprehensive, enterprise-level data mining solutions and professional services. NeoVistas Solutions' Decision Series suite of knowledge discovery tools solves data mining challenges in a variety of markets, including retail, insurance, telecommunications, and healthcare. The Decision Series suite includes pattern discovery tools based on neural networks, clustering, genetic algorithms, and association rules (see Fig. 3) [5].

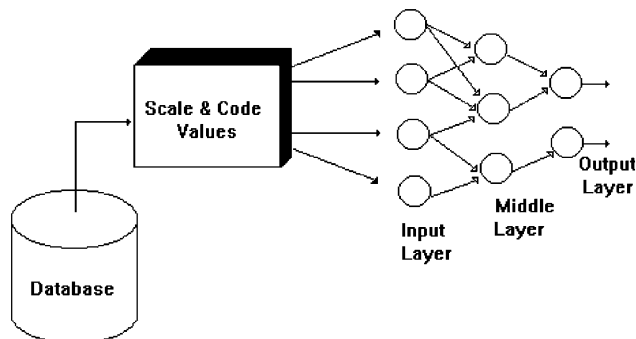


Fig. 3. Diagram of a typical neural network. Source: Information Discovery, Inc. [18].

5.1.1. *The problem*

A large retailer, with over \$1 billion in sales, found its profits were suffering due to less-than-optimal seasonal product demand forecasting. The retailer was over-stocked on its slow-moving products and under-stocked on its most popular items at critical selling periods.

5.1.2. *The solution*

NeoVista designed and implemented a solution that combined elements of the clustering and neural network technology which enabled the retailer to automatically review its point-of-sale history and equate store groupings to sales patterns. Management is now able to explore the lowest level of detail and forecast stocking requirements for individual stock keeping units (SKUs) on a store-by-store basis. In addition, by combining neighborhood demographics with historic sales patterns, management receives exact data which enables them to continuously fine-tune their replenishment system.

5.1.3. *The results*

Management is able to forecast seasonal trends at the store-item level. Additionally, the Decision Series tools showed that clustering similar items into actionable groups streamlined the ordering process. Therefore, the company can now predict demand for SKUs and operate a just-in-time inventory program far more effectively. In the year since implementation, the company has increased revenues by 11.6% while reducing inventory by 2%, counter to the industry trend [14].

5.2. *Case study: CHAID*

Applied Metrix helps companies increase their competitive advantage and margins through by optimizing their marketing and sales productivity. Applied Metrix's focus is to lower the cost of customer acquisition and maximize the lifetime value of customers. Applied Metrix uses a combination of CHAID segmentation and logistic regression response probability modeling to establish predictive models that are deployed over a proprietary Internet system [15].

5.2.1. *The problem/goal*

The client was a home equity marketer that extended home equity lines of credit at the national level. The client's goal was to increase the efficiency of targeting current mortgage customers who might be interested in the client's service. The client set a goal of at least a 10% increase in targeting efficiency, which in turn would lead to a seven-fold payback in the first year's gross profits [16].

5.2.2. *The solution*

Applied Metrix used CHAID for the initial segmentation modeling. They used "tree" segmentation to identify important interactions among predictors of response to promotion. The CHAID process identified market segments that were formed by interactions among predictors of a chosen criterion variable. For example, customer

age may predict response-to-promotion differently within different categories of household income. In this case, CHAID identified 16 distinct market segments, where the segments represented combinations of individual predictors of response to promotion. In particular, one particular segment accounted for 65% of responses to the mailing, which enabled them to identify where potential profits actually declined.

After the segments were identified, the next step was to employ logistic regression response probability modeling. This step allowed measurement of the unique influence of each predictor on a specific criterion, and also allowed scoring of individual customers or prospects in the marketing database. In addition, logistic regression assigned a probability of response as a percentage to each record in the database. The CHAID segmentation became an independent variable in the regression model [16].

5.2.3. *The results*

The model was implemented with amazing success. The highest-rated group from the predictive model had by far the highest response rate to the equity line of credit campaign—85% above average for the direct mailing, while those who were predicted to be poor responders were 49% below average. Therefore, the analysis showed a strong correlation between predicted scores and the actual response rates.

The model's success transferred to the bottom line. The goal of the program was a 10% increase in response rate, but the actual response rate increased 30%. The firm was able to increase sales by \$36 million and profits by over one million dollars in the first year after implementation and review of the old system. Thus, it helped the firm make better decisions, cut down on waste, and made money immediately after implementation [16].

6. CHAID vs. neural nets

6.1. *Clarity and explicability*

The form of a CHAID model is understandable as a set of rules, whereas a neural network is obscure, with weights that have no intuitive meaning. It is possible to apply background domain knowledge to a CHAID model because it should be easy to explain to a domain expert or business user.

6.2. *Implementation/integration*

It is much easier for a CHAID model to be implemented than a neural network. Moreover, the risk of missing code by an IT department is slim for a CHAID model and higher for a neural network. The performance of an implemented CHAID model will be significantly faster than an implemented neural network.

6.3. *Data requirements*

More data must be provided for a CHAID model to ensure that there is critical mass in the leaf nodes following many branches. The data for both techniques

requires some pre-processing. Neural networks require the data be transformed into binary format. Before using CHAID, any continuous independent variables must be banded.

6.4. Accuracy of model

Neural networks provide more accurate (i.e., powerful and predictive) models, especially for complex problems. However, there is also a risk of finding sub-optimal solutions and over-fitting.

6.5. Construction of model

CHAID is easier and quicker to construct, whereas neural networks have many parameters that must be set and require more skilled manipulation to ensure over-fitting does not occur. It is harder to apply background domain knowledge using neural networks, whereas it is easier to see mistakes and over-fitting in a CHAID tree.

6.6. Costs

Building a neural network is more costly than building a CHAID model. A neural network requires more time and a higher level of building skills than a CHAID model.

6.7. Applications

Both CHAID and neural networks can create predictive models. Such models include attrition, churn, propensity to purchase, and customer lifetime value. Yet in general, the application of neural networks is wider than CHAID. The reason for this is that neural networks can be applied to both directed (supervised) and undirected (unsupervised) data mining. Neural networks can handle both categorical (e.g., marital status) and continuous (e.g., income) independent variables, but these have to be transformed to 0/1 input variables. When all or most of the independent variables are continuous, neural networks should perform better than CHAID. When all or most of the independent variables are categorical with high cardinality (i.e., implicit “containing” relationships), CHAID should perform better than neural networks.

In addition to the more common predictive models of marketing, both neural networks and CHAID can be used to solve sequence prediction problems, for example, predicting share prices in the stock market; however, major effort is required to pre-process the time series data.

Neural networks can be used to solve estimation problems (with continuous outcomes); whereas CHAID provides good solutions to classification problems, can be used for exploratory analysis (perhaps prior to another modeling technique), and can provide descriptive rules.

CHAID models are easier to build and implement than neural networks and also are less costly. Theoretically, neural networks should provide models that are better

than CHAID in terms of power and accuracy. That means they should be more powerful at discriminating between groups that fit the target (for instance, churn) and that they should predict correctly more often. Currently, this is not the case, perhaps because of the problems of over-fitting and sub-optimal solutions.

The exact workings of a CHAID model can easily be seen, as it is intuitive; a neural network has been described as a “black-box” because it is impossible to explain how each outcome is determined. For this reason, the US credit industry is prohibited from using neural networks to determine credit risk because the lack of clarity means that unfair prejudice cannot be ruled out of the credit decision.

In terms of predictive modeling, CHAID wins against neural networks, yet in the future if neural networks become easier to build and the methods for producing rules that explain a neural network improve, neural networks could become the winners. An interesting development is the combination of these two techniques to create “neural trees”. These could use the CHAID method to identify sub-populations on which neural networks could be built to predict a particular target.

7. Conclusions

In choosing a suitable technology for personalization or CRM, organizations must be aware of the tradeoffs when considering differing data mining software applications. The choice among different options is not as critical as the choice to use data mining technologies in a CRM initiative. Data mining represents the link from the data stored over many years through various interactions with customers in diverse situations, and the knowledge necessary to be successful in relationship marketing concepts. In order to unlock the potential of this information, data mining performs analysis that would be too complicated and time-consuming for statisticians, and arrives at previously unknown nuggets of information that are used to improve customer retention, response rates, attraction, and cross selling. Through the full implementation of a CRM program, which must include data mining, organizations foster improved loyalty, increase the value of their customers, and attract the right customers.

As customers and businesses interact more frequently, businesses will have to leverage on CRM and related technologies to capture and analyze massive amounts of customer information. Businesses that use customer data and personal information resources effectively will have an advantage in becoming successful. However, businesses must also bear in mind that they have to use technology responsibly in order to achieve a balance between privacy rights and economic benefits.

Different technologies vary in terms of effectiveness and ease of use. It is businesses and managers who determine how to exploit collected data, in other words, more of a policy issue than a technology issue. Several precautions have to be taken by business to assure consumers that their privacy will be respected and personal information will not be disclosed without permission. Businesses also have a duty to execute their privacy policy so as to establish and maintain good customer relationships. For such a sensitive issue as privacy, the burden is on businesses when it comes

to building and keeping trust. The nature of trust is so fragile that once violated, it vanishes. Current CRM solutions focus primarily on analyzing consumer information for economic benefits, and very little touches on ensuring privacy. As privacy issues become major concerns for consumers, surely an integrated solution that streamlines and enhances the entire process of managing customer relationships will become even more necessary.

References

- [1] Pettrissans A. Customer relationship management: the changing economics of customer relationships. White Paper prepared by Cap Gemini and International Data Corporation, 1999.
- [2] Newton's Telecom Dictionary, Harry Newton, CMP Books, <http://www.cmpbooks.com>
- [3] Edelstein H. Data mining: exploiting the hidden trends in your data. DB2 Online Magazine. <http://www.db2mag.com/9701edel.htm>
- [4] Data Intelligence Group Pilot Software. An overview of data mining at Dun & Bradstreet. DIG White Paper, 1995 <http://www3.shore.net/~kht/text/wp9501/wp9501.shtml>
- [5] Decision Support Solutions: Compaq. Object relational data mining technology for a competitive advantage. <http://www.tandem.com/brfs~wps?odmadvwp/odmadvwp.htm>
- [6] Garvin M. Data mining and the web: what they can do together. DMReview.com, 1998. <http://www.dmreview.com/editorial/dmreview/print~action.cfm?EdID=420>
- [7] IDC & Cap Gemini. Four elements of customer relationship management. Cap Gemini White Paper.
- [8] Freeman M. The 2 customer lifecycles. *Intelligent Enterprise* 1999;2(16):9.
- [9] Hill L. CRM: easier said than done. *Intelligent Enterprise* 1999;2(18):53.
- [10] Thearling K. Data mining and CRM: zeroing in on your best customers. *DM Direct*. December, 1999. <http://www.dmreview.com/editorial/dmreview/print~action.cfm?EdID=1744>
- [11] Chablo E, Marketing Director, smartFOCUS Limited. The importance of marketing data intelligence in delivering successful CRM, 1999. <http://www.crm-forum.com/crm~forum~white~papers/mdi/sld01.htm>
- [12] Thearling K, Exchange Applications, Inc. Increasing customer value by integrating data mining and campaign management software. Exchange Applications White Paper, 1998. <http://www.crm-forum.com/crm~forum~white~papers/icv/sld01.htm>
- [13] Thearling K, Exchange Applications, Inc. Increasing customer value by integrating data mining and campaign management software. Exchange Applications White Paper, 1998. <http://www.crm-forum.com/crm~forum~white~papers/icv/sld02.htm>
- [14] Book J. Customer relationship management: what is it that separates CRM contenders from CRM pretenders? *DMReview.Com* September, 1999. <http://www.dmreview.com/editorial/dmreview/print~action.cfm?EdID=1368>
- [15] Neoweb Website. <http://neoweb.accrue.com/retail/case~studies.htm>
- [16] Applied Metrix website. <http://www.appliedmetrix.com/CaseStudies/bank.html>
- [17] Pilot Software. An introduction to data mining: discovering hidden value in your data warehouse. <http://www3.shore.net/~kht/text/dmwhite/dmwhite.htm>
- [18] Information Discovery, Inc. A characterization of data mining technologies and processes: an Information Discovery, Inc. White Paper. <http://www.dmreview.com/portal~ros.cfm?NavID=92&WhitePaper=60&PortalID=9>
- [19] Cranford S. Knowledge through data warehousing: data mining—the intelligence component" *DMReview.com* May, 1998. <http://www.dmreview.com/editorial/dmreview/print~action.cfm?EdID=630>

Chris Rygielski received a MBA degree with a concentration in Management Information Systems on May 2001 from Miami University at Oxford, OH and is currently working in the information systems area. His research interests include customer relationship management, data communications, and electronic commerce.

Jyun-Cheng Wang is currently Associate Professor of Information Management at National Chung-Cheng University in Taiwan. He holds a PhD degree from the University of Wisconsin-Madison. His research interests are mainly in the area of electronic commerce, automatic negotiation, intelligent agent, and the information economy.

David C. Yen is Professor and Chair of the Department of Decision Sciences and Management Information Systems at Miami University. He received a PhD degree in MIS and a Master of Science degree in Computer Science from the University of Nebraska. Professor Yen is active in research, has published two books and over 150 articles which have appeared in *Communications of the ACM*, *Information & Management*, *International Journal of Information Management*, *Journal of Computer Information Systems*, *Interface*, *Telematics and Informatics* and *Internet Research* among others. He was also one of the co-recipients of a number of grants including grants from the Cleveland Foundation (1987–1988), GE Foundation (1989), and Microsoft Foundation (1996–1997).