DATA MINING: A KEY TECHNOLOGY FOR MARKETING EDUCATORS

Edward J. Ryan, Jr., Millsaps College, 341 Mockingbird Lane, Madison, MS 39110 (601-853-2491)
Ajay K. Aggarwal, Millsaps College, 1701 North State Street, Jackson, MS 39210 (602-974-1270)
Philip D. Ryan, ConsultRyan, 22051 West Ripple Road, Buckeye, AZ 85326 (602-386-2763)

ABSTRACT

Data mining is an emerging technology with important implications for marketing educators. The authors define the notion, describe data mining components, present applications that should interest marketing educators, and mention some obstacles to its future development. They conclude that data mining has the potential to enhance a firm's marketing strategy and improve its performance. The authors also suggest marketing educators consider mentioning data mining when carrying out their teaching, publishing, and consulting tasks.

EDUCATORS' OVERVIEW OF DATA MINING

Marketing educators are constantly searching for new technological developments that effect their professional responsibilities. An emerging technology that should be of interest to members of the Western Marketing Educators' Association is data mining. Regarded by some an "amorphous" notion, data mining helps marketers discover certain patterns and correlations from data that were previously unknown. This in turn enables organizations to generate better marketing plans and new business opportunities (Foley, 1996).

Data mining can be the derivation of useful information from a data base through creative queries. It also can be the identification of relationships that would have gone undetected without the application of specialized approaches (Grupe, 1995). Mena (1996) describes data mining as a methodology for using software to analyze database records to discover patterns. Given a set of thousands of database records, the software searches for a pattern and rule to describe them. The knowledge extracted is then presented in usable business statements that managers who do not have extensive expertise in quantitative decision making can interpret.

Lebowitz (1996) says that data mining is one of five technologies that can enhance profitability in the financial services industry. Dickey (1996) believes that the growing importance of data mining shows a shift from product orientation to customer orientation. Market potential also reflects the importance of data mining. For example, the markets for data warehouses and related tools are growing fast. META Group predicts that the market for data mining will reach $300 million in 1997 and rise to $800 million by the year 2000 (Shoesmith, 1996). The overall market was $2 billion in 1995 (Kay, 1996).

Data mining software is an offshoot of university based artificial intelligence research and the most complex and technologically sophisticated of data query tools (Richman, 1996). It is small wonder then that marketing educators are only recently being exposed to data mining. The same is true in industrial and government sectors where data mining tools have been traditionally restricted to scientific and engineering settings.

PURPOSE AND METHOD

The purpose of this paper is to familiarize or update attendees at the 1997 Western Marketing Educators' Conference with the subject of data mining. The authors hope it will motivate members to become more aware of this important technological development and incorporate it into their teaching, publishing, and consulting tasks. This is consistent with the objectives of data mining hardware and software providers. They plan to bring data mining out of the realm of academia, artificial intelligence, and highly skilled power users and into the corporate middle class.

This conceptual paper is based on 193 articles that cover data mining. They were published during 1995 and 1996 and recorded in the ABI/Inform data base. No attempt was made to find information from other sources such as books or scientific journals. The intent was to compile a general overview of interest to marketing educators who do not specialize in the decision sciences.
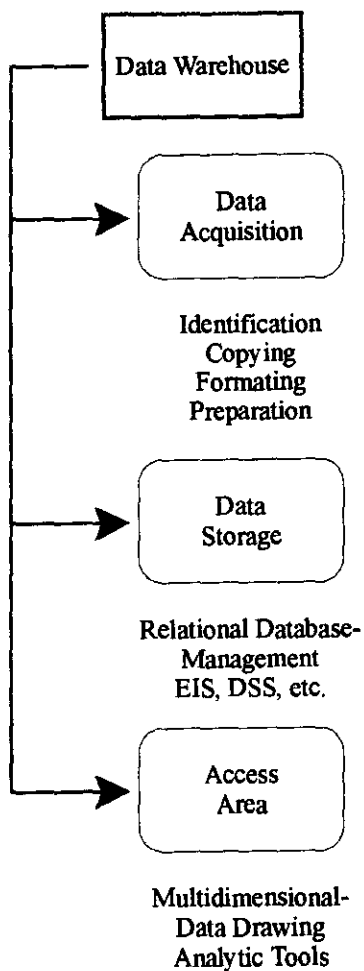
COMPONENTS OF DATA MINING

Allen (1996) noted three components of data mining when discussing mortgage banking: (1) production

systems that track customer interactions; (2) data warehouses to collect relevant data, and (3) extract engines that identify data elements. These components are similar to those described by authors who have investigated data mining in other industries. Production systems routinely collect data during an organization's day-to-day operations. An example is a firm's marketing information system that generates and stores internal corporate data, external competitive data, environmental data and marketing research data. This data is ultimately stored in a data warehouse along with that gathered from other sources.

Figure 1 depicts the three functional areas of a data warehouse that are customized to meet business needs. The first handles acquisition of data from legacy systems and outside sources.

FIGURE 1
Data Warehouse Functional Areas

Data Warehouse

Data Acquisition

Identification
Copying
Formating
Preparation

Data Storage

Relational Database-
Management
EIS, DSS, etc.

Access Area

Multidimensional-
Data Drawing
Analytic Tools

A second area of the warehouse is the storage area and the third is the access area. The widest range of unique products can be found in the user access area. There different end-user PCs and workstations mine data from the warehouse (Mattison, 1996).

Several approaches are used to extract knowledge from existing data bases. Grupe (1995) noted five: (1) software such as DataLogic/R, IDIS:2; Recon, Nicel, and KnowledgeSeeker; (2) neural networks; (3) expert systems; (4) case-based reasoning systems; and (5) data base software extensions. The approaches are designed to identify relationships that would have gone undetected without their application.

EXAMPLES FOR MARKETING EDUCATORS

The authors contacted several executives whose organizations were involved with data mining. It was no surprise, given the idea of competitive advantage, that they refused to discuss it in detail. This probably also accounts for the sketchy descriptions of specific applications presented in the literature. We hope that the examples that follow will be sufficient to meet the paper's objective, which is to familiarize marketing educators with the topic.

Data mining has proven valuable in a wide range of applications. These include direct-marketing segmentation, credit-card fraud detection, financial and investment analysis, detection and prediction of chronic faults in large networks, analysis of healthcare information, intelligent instruments, and database cleaning (Greenfeld, 1996).

Within the financial services industry, data mining is labeled one of five technologies that holds promise of enhancing the industry's profitability. For example, Lebowitz (1996) said mortgage companies should regard data mining as strategic to their business. Nevertheless, most companies lack the information needed to make critical target market decisions. The companies have depended on databases whose design was based on products and not customers. In reality, customer buying patterns and attributes are different. Fanella (1996) agrees and notes that data mining helps develop customized products and services that are more in line with what customers want.

41

Another example concerns Mellon Bank that initiated a data-mining project with IBM Corporation. Its purpose is to develop computing tools that enabl the institution's retail marketing group to pull information from the bank's customer databases. The objective is to identify persons likely to respond favorably to direct-mail cross-selling efforts. The first mailing based on the project's findings promotes home equity loans.

Westergren (1996), president of the Bank Marketing Association, summarized the importance of data mining to the financial services industry when he commented: "Whatever you call it, [data mining] remains the key to segmenting your customer base into workable units that you design products for and deliver service to. You must know your customers' profiles and what their propensity to buy really is. This mining of your customer database can be accomplished either internally or externally by using an outside source. What is important is that you begin to do it now or the industry will leave you behind."

A variety of other industries and organizations use data mining. One instance is Pilot Software that recently introduced Pilot Discovery Server, a data mining component of its Decision Support Suite. It is aimed specifically at sales and marketing users rather than to statisticians and data mining experts. Its objective is to analyze a relational database, identify groups of customers who have similar characteristics and make predictions for those groups. For example, Lightbridge Corporation tested the Discovery Server in the cellular telephone industry. It helped the company determine which customers are most likely to defect to another company and why. This information has helped the firm determine how to keep more of them.

Mena (1996) pointed out two marketing applications of data mining to customer purchasing. First a set of records from a huge database of individuals who bought a product, and those who did not, was exposed to data mining software. The output was a set of "what-if" statements such as the following: IF credit line = (22,78) THEN = 65.9% potential client. This means that if the company asks for a down payment of 22%, with the remaining 78% financed, 65.9% of the potential clients will likely buy the product.

The second possibility suggested by Mena (1996) was to use data mining technology to discover associations in purchasing patterns. A supermarket retailer, for instance, might discover from its bar code database that 88% of the customers who buy more than $100 worth of groceries, including deli items, purchase expensive wine.

Within the distribution area, data mining is used to meet demands for smaller and more frequent shipments. An example reported by Hasty (1996) is Duckhead Apparel Warehouse. It is an $8-million state-of-the-art distribution technology built around the quantity-neutral premise. That is, distribution managers attempt to get down to the store level and determine a different product mix for each store. Duckhead managers expect the facility to pay for itself in two years.

Ross (1996) reports use of data mining in retailing to reduce employee-retail shrink. Managers aim several new applications at identifying dishonest personnel who account for an estimated 38 percent of retail shrink. They also identify those who may be making errors at the point of sale. The software helps retailers find fraud patterns by collecting and "remembering" information about incidents occurring at the point of sale. For instance, they can program "self-ringing employee" alerts when employees ring up returns for themselves.

The favorable examples of data mining covered in this paper suggest that it can be easily employed in practice. This is not so since obstacles will hinder its application as discussed in the next section.

OBSTACLES TO DATA MINING GROWTH

The greatest obstacles to the growth of data mining are probably consumer mistrust, privacy, and government intervention. Lucas (1996), a proponent of data mining, notes that surveys and complaints to government regulators suggest that consumer tolerance for telemarketing may be eroding. Consumer fraud and billing cost telephone companies huge amounts of money (Urquhart, 1996). Another major concern is how marketers gather and use information about children from the Internet. Consequently, the government is examining data stockpiles and how they are being used. The issue is privacy. During June 1996 the Federal Trade Commission held a workshop to

decide if additional guidelines should be set by the government. The FTC was particularly concerned with how companies use the data they collect over the Internet. Advocates are pressing proposals to regulate the way marketers garner and use information. Legislation is also pending on Capitol Hill.

Many technical considerations also present obstacles to full implementation of data mining. However, they are beyond the scope of this paper and are not presented. Despite its limitations an effective strategy for data mining is to use it to deliver business-critical customer information faster, in greater volumes, and with greater precision (Jeffery, 1996). The technology could enhance an organization's marketing strategy as illustrated in the previous section.

SUMMARY

This paper is addressed to members of the Western Marketing Educators' Association. Its purpose is to familiarize or update them with the subject of data mining. It is also presented to help members determine if the topic should be incorporated into their teaching, publishing, and consulting tasks. Data mining is an offshoot of university-based artificial intelligence research and the most complex and technologically sophisticated of data query tools. Therefore, data mining has been confined to the scientific and technical community. It has only recently received attention from general marketing and business administration academicians and practitioners. The technology has multiple applications in marketing including direct-marketing segmentation, credit-card fraud detection, financial and investment analysis, detection and prediction of chronic faults in large networks, analysis of healthcare information, intelligent instruments, and database cleaning. Major obstacles to data mining include consumer mistrust, privacy, and government intervention. Despite these challenges, data mining has the potential to both enhance an organization's marketing strategy and improve its performance when employed in a socially responsible manner.

REFERENCES

Allen, L.E. (1996). Mining Gold from Data Bases. Mortgage Banking. 56:99-100.

Dickey, S. (1996). Mine Your Own Business.

Midrange Systems. 9:36-38+.

Fanella, J.A. (1996). Meet Customers' Needs Using Data Warehousing. Best's Review (Property/Casualty). 96:80.

Foley, J. (1996). Mining Data for a Marketing Plan. Informationweek. 587:87.

Greenfeld, N. (1996). Data Mining. UNIX Review. 14:9-14.

Grupe, F.H. (1995). Data base mining: Discovering New Knowledge and Competitive Advantage. Information Systems Management. 12:26-31.

Hasty, S.E. (1996). Mining Databases. Apparel Industry Magazine. 57:4.

Jeffery, B. (1996). Getting to Know the Customer Through a Customer Information System. Information Strategy: The Executive's Journal. 13: 26-36.

Kay, E. (1996). Decisions Decisions. Informationweek. 584:73-78.

Lebowitz, J. (1996). The Dawning of a New Era. Mortgage Banking. 56:54-66.

Lucas, P. (1996). Calling All Cardholders. Credit Card Management. 9:44-48.

Mattison, R. (1996). Warehousing Wherewithal. CIO. 9:58-63.

Mena, J. (1996). Mining Your Business. Network World. 13:39.

Richman, D. (1996). Data Mining Chisels Its Niche. Computerworld. 30:49+.

Ross, J.R. (1996). New Rule-based Systems Tackle Employee Shrink. Stores. 78:71-72.

Shoesmith, J. (1996). Vendors Hype Treasures in Data Mines. Computing Canada. 22:18.

Urquhart, J. R. (1996). Mining for Gold. Telephony. 23:148-154.

Westergren, J.M. (1996). The New Bank Marketing. Bank Marketing. 28:140.