NC STATE UNIVERSITY

JTATM

JOURNAL OF TEXTILE AND APPAREL, TECHNOLOGY AND MANAGEMENT

Volume 2, Issue 3, Summer 2002

Enhanced Decision Making using Data Mining: Applications for Retailers

Joan Anderson, Ph.D.
Department of Apparel, Merchandising, and Interior Design
Washington State University

Antigone Kotsiopulos, Ph.D.
College of Applied Human Sciences
Colorado State University

ABSTRACT

As the economy has tightened, retailers have been challenged in recent years to be ~~be~~ more strategic in their planning. They struggle to find answers to:

- *Who can I consider a loyal customer?*
- *What kind of marketing strategy is most likely to increase sales?*
- *What can customer-purchasing patterns reveal about improving inventory control?*
- *What is the most effective way to manage customer relations to increase revenues? (Rabinovitch, 1999).*

With the exponential growth in the amount of data being collected, improvements in technology, and research in machine learning, retailers are now able to reduce the ever growing difficult and complex decision making process by recruiting the efforts of data mining (Barry & Linoff, 1997). Data mining is a computerized technology that uses complicated algorithms to find relationships and trends in large data bases, real or perceived, previously unknown to the retailer, to promote decision support. Currently being utilized by such retail giants as Federated Department Stores, Nordstrom, and Wal-Mart, Inc., data mining is touted to be one of the greatest technologies to hit the retailing industry this decade (Rabinovitch, 1999). The purpose of this study is to critique data mining technology in comparison with more familiar analytical tools for strategic decision making by small to medium size retailers. The context for this study includes current and future industry applications and practices for research performed in data mining applications within the retail sector.

KEYWORDS: data mining, retailers, decision making, marketing strategy, strategic planning, technology

Today's retail customer has changed from the typical customer of the past. Customers are becoming increasingly more informed and therefore more demanding. They are confused by too many choices and, for the sake of time, limit the businesses with which they deal (Cabena, Hadjinian, Stadler, Verhees & Zanasi, 1998). Market saturation has forced retailers to look for new niche markets such as different ethnic groups or those with special needs in order to compete. Traditional mass marketing is proving to be ineffective in today's competitive markets and customers are searching out more targeted channels for specific merchandise. Merchants are turning to the development of private label brands as a means of differentiation. Competition has made it more and more difficult to stand out from the crowd, so new markets are being explored through different channels, such as e-commerce. Product timing is of even greater concern. Getting new products on the market and on the sales floor is becoming increasingly more important. Product life cycles are getting shorter, (e.g. the life of a computer), leaving less time to realize a profit on a given product (Cabena et al, 1998).

With all of these changes in the business environment, retailers are beginning to focus more on preventing defection to the competition by creating better relationships to increase loyalty, although a 1996 study found that retailers spend five to ten times more money on obtaining new customers rather than trying to keep existing customers. By 2001, 52.5% of retailers surveyed where using data mining technology (Chain Store Age, Anonymous Author, 2001). They are also focusing on competitors' strategies with regard to new store location, product offerings and pricing.

Retailers are realizing that all the data they are collecting has value if properly stored and analyzed. Today the United States alone has 450 to 750 petabytes of electronically stored information. This is the equivalent of up to 375 trillion document pages and records. It is also reported that this is growing at a rate of 18% to 30% annually (Kempster, 1998).

Retailers struggle to find answers to questions such as:

- Who can I consider a loyal customer?
- What kind of marketing strategy is most likely to increase sales?
- What can customer-purchasing patterns reveal about improving inventory control?
- What is the most effective way to manage customer relations to increase revenues? (Rabinovitch, 1999).

With the exponential growth in the amount of data being collected, improvements in technology, and research in machine learning, retailers are now able to reduce the ever growing difficult and complex decision making process by recruiting the efforts of data mining (Barry & Linoff, 1997). Data mining is touted to be one of the greatest technologies to hit the retailing industry this decade (Rabinovitch, 1999).

The purpose of this study is to critique data mining technology in comparison with more familiar analytical tools for strategic decision making by small to medium size retailers. The context for this study includes current and future industry applications and practices for research performed in data mining applications within the retail sector.

What is Data Mining?

Data mining defined. There are abundant definitions of data mining in the literature. (Anderson, 2001; Marketing Computers, 1996; Ritzer, 1996; Thearling, 2002). For the purposes of this article, data mining is a computerized technology that uses complicated algorithms to find relationships and trends in large data bases, real or perceived, previously unknown to the retailer, to promote decision support. Data mining is currently in a state of growth. More products are being developed, more businesses are incorporating the efforts of data mining into their decision making processes, and more is being written about the benefits and burdens of data mining than ever before (Anderson, 2001). The market for commercial data mining software is expected to rise from $540 million in 2002 to $1.5 billion in 2005. However, according to Thearling (2002) it depends on what you call data mining. Vendors are focusing more on adding new tools to existing applications rather than developing new ones and there is a push toward standardization and integration. Data mining has been referred to as a statistical process of analyzing data stored in a data warehouse (Decker, 1998). A data warehouse is an extensive data repository consisting of information from all facilities of organizations' operations, including external sources, that is maintained to support decision-making. Data within the warehouse is manipulated to create easy access by data mining tools. Data mining is possible without a data warehouse, but more difficult. Smaller, downsized versions of data warehouses can be created. These are known as data marts and focus on one particular area of a database such as credit card users (Decker, 1998). Data marts are less expensive and can operate in a much smaller environment. The following example illustrates this concept:

Table 1. Comparison of Data Warehouse and Data Mart Databases.

DATA WAREHOUSE	DATA MART
Total company sales information	Sales for a single location
Store sales/Hourly sales	Credit card sales only
Customer profiles	Credit card customer profiles
	Credit card purchasing history

Knowledge discovered through data mining can take on different forms. Existing research identified five general categories of information (Anonymous, 1996; Thearling, 2002):

1). Associations or affinities includes occurrences linked to a single event, such as evaluating the profitability of promotions or identifying patterns that exist within a single set of items, for example, product combinations.

2). Sequencing includes events that are linked together in some fashion. An example would be if a blazer were bought, then what percent of the time will a blouse be bought within how many days and accessories within how many days?

3) Classifications include rules for sorting characteristics into categories such as customer types. Classifications can be used to analyze the effectiveness of a promotion on a specified customer group or the necessary expenditure to obtain a certain patron.

4) Clustering or segmentation can be used to discover groups within a database. For example, creating customer segments based on a particular attribute such as risk, profit or response. It can help answer questions on retaining a high profit, high value, low risk customer or increase profit by cross selling to low profit, high value, low risk customers.

5) Forecasting or predicting involves making predictions of a certain dependant variable such as sales, response, behavior or lifetime value, based on patterns found within the data. The factors found within the data are independent variables; also know as predictors, such as purchase history and demographics. An accurate and complete historic database is necessary for accurate predictions.

Data mining techniques. In addition to different sizes of databases ands different forms that information can take, there are different data mining techniques. In a white paper published in 1995 by Data Intelligence Group, six common data mining techniques where discussed:

1) Artificial neural networks are a non-linear predictive decision making model which uses existing data with a known outcome to train a model which can then be used to make predictions on data with unknown outcomes.

2) Genetic algorithms are a technique that is based on the concepts of natural evolution using genetic combination, mutation, and natural selection as a form of optimization.

3) Decision trees are tree shaped decision models that utilize rules to classify a data set. Each model represents sets of decisions.

4) Nearest neighbor method is a data classification technique based on similar records within a historical database.

5) Rule induction is the use of if-then rules from data based on statistical significance.

6) Data visualization is the pictorial representation of relationships among data.

How Does Data Mining Work?

Gathering data. Historically, small and medium sized retailers have had the privilege of developing close and mutually beneficial relationships with their customers. These relationships were possible because peoples buying behavior did not change much. The availability of choices, which fostered changes in preferences for and perceptions of products, was limited. Price was less of an issue due to less competition (Cabena, et al., 1998). Customer bases were smaller and more loyal (Cabena, et al., 1998). Retailers benefited from these relationships in ways such as knowing whether or not customers would like a new product, what colors and sizes of apparel sold best and whether to discontinue a product or just reduce the stock because of certain customer preferences. Customers also benefited from these relationships. They had a store where they could count on products they needed at a price they expected and new products made available aligned with their needs and wants. With the onslaught of mass production and mass media, these relationships became increasingly difficult to maintain (Cabena et al. 1998). Today, with the presence of category killers, such as Toys R Us and hypermarkets, such as Biggs, these relationships are hard to find and difficult to create. Customer behavior has been altered exponentially with continuously changing demographic and pychographic behavior (Cabena et al. 1998).

In the past, retailers built relationships with customers by realizing their needs, remembering what they preferred, and learning over time how to serve them better (Berry & Linoff, 1997). Today's larger retailers attempt to replicate the intuition of retailers of past. They aim to create an image of remembering customers by name and knowing their purchasing behavior. Seem impossible? Technologies such as data mining available to retailers today make this easier than most think. With the overwhelming presence of the Internet, intranets, and data files in the giga and tera byte ranges, the ability to transform raw data into workable and predictable solutions to complex merchandising and retail problems has become a reality. Larger corporations have been data mining for years, but because of becoming cost prohibitive, smaller companies have been left behind. The fact that smaller companies could not afford or utilize data mining software has changed. According to the Data Intelligence Group (1995), with the now mature technologies of data collection and storage, powerful multiprocessor computers, and data mining algorithms, data mining is ready for application into the business community. More and more software aimed at this market is being developed (Marketing Computers, 1996, Thearling, 2002).

Analyzing data. Statisticians have been manually data mining for years (Thearling, 2002). The first necessary step is to identify the database to be analyzed. The retailer then must decide if they have specific questions, or a hypothesis they are researching or if they would like the system to generate information for them. At this time, it is necessary to choose the appropriate data-mining tool for the database and type of information being sought. Data mining tools differ in the way they work and the type of problems they are designed to address. Once the tool is in place, the retailer either asks the system to generate the information or tests the predetermined hypothesis. Once the knowledge is discovered, it is necessary to validate. Once the retailer has valid information, action is taken and the outcome analyzed. The process is then adjusted and repeated.

Critical Analyses of Scenarios Illustrating Use of Data Mining versus Other Technology:

The type of information available to retailers with no computer for data analysis, retailers with computers using traditional data analysis capabilities, and retailers with computers using data mining capabilities will be compared to illustrate the value of data mining as input to strategic planning for current retailers. The following three types of questions are commonly encountered by retailers and merchandisers and thus serve to demonstrate realistic input and outcomes. Table 2 displays this comparison in summary format.

Scenario A – Tracking promotional effectiveness: Sales – "How well did I do in my last promotion?"

1: No computer – In tracking sales and promotional effectiveness with no computerized point of sale, a retailer would have to track information manually. Cash registers provide general information such as total sales and sales by department, however, discount dollars and coupon sales would have to be tracked manually. Sales and discount dollar information must be recorded in some sort of a ledger to make comparative data analysis with past time frames and promotions. There would most likely be no sales information by stock keeping unit (SKU) as this would be very time consuming. In addition, customer information would also have to be tracked in some manual fashion. Tracking the effectiveness of multiple promotions running simultaneously would prove very troublesome and any information with regards to add-on sales or selling up to a higher price point would not be possible without putting in a considerable amount of over time.

2: Using a computer with traditional data analysis capabilities. Computerized point of sale software programs can track sales information from very general to very detailed, depending on the software package being utilized and the manner in which the program has been setup. Numerous sales reports can be generated to facilitate data analysis. Retailers must formulate predetermined questions to analyze and all analyses are based on historical data that has been accumulated. Information available includes comparative sales information; discount dollars generated and customer information. The following are examples of some of the questions that retailers can answer through the use of reports.

a)"Did my sales increase during the time of my promotion?"

b)"Did I sell more of the SKU's than normal during promotion?"

c)"Which customers bought from me during the promotion?"

3: Using a computer with data mining capabilities. Databases are generated with similar sales information as the traditional data analysis software programs. The retailer does not necessarily need to formulate predetermined questions to analyze. Data mining automatically analyzes the information and can find relationships among the data to answer such questions as:

a) "Why are my discount coupons or promotions not attracting the sort of return I was expecting?"

b) "How do I get my other stores to match the sales figures of my northern stores?"

Using data mining in analyzing promotions can allow retailers to target the top 5% of customers, increasing the response rate by 2% and generating as much as 20% of the sales for a particular promotion (Murphy, 1998). The following are examples of the type of information that data mining can produce:

a) "If a store sold 100 more button down white collar shirts than any other white collar shirts in the past month, and if the retailer has been running a one month promotion, then sales will increase 25% when the promotion is repeated."

b) "Cola is purchased 65% of the time when customers buy corn chips at regular price and 85% of the time during promotions (Data Mining Information, Anonymous Author 1996)"

Scenario B – Inventory Control: "How do I avoid stock outs on fast moving items?"

1: No computer. When retailers do not use computers to keep track of inventories, all SKU counts must be done by hand. This is very time consuming and carries a great potential for human error. Counts must be done frequently and manually on fast moving items. "Eyeballing" (visually scanning inventory) becomes a frequent practice reducing accuracy. This type of inventory control reduces response time on reorders and increases the potential of missed sales due to stock outs. Commonly utilized figures such as stock to sales ratios, turnover, and gross margin return on investments must be calculated by hand with sales receipts, inventory counts and a calculator. This is time consuming, subject to human error and is most likely computed for longer periods of time. This may be acceptable for slower moving items, but can prove detrimental while attempting to maintain stock levels for fast moving SKUs.

2: Using a computer with traditional data analysis capabilities. When using a computerized inventory control system linked to a point of sale program, retailers can maintain a perpetual inventory. This is a much more accurate form of inventory control, allows for greater control of sizes and colors within styles and can free up the time spent counting by hand. Inventory reports can be generated to aid in answering predetermined questions such as:

a) "Which of my SKU's am I running low on?"

b) "Of my top ten sellers, what are my stock levels?"

Most traditional data analysis software programs can automatically generate stock to sales ratio, turnover, and gross margin return on investment figures for predetermined periods of time, such as one year or one day. This can be beneficial when monitoring stock levels on fast moving items.

If the retailer chooses, they can maintain on-line ordering with vendors to automatically monitor stock levels of certain items allowing for faster replenishing times and often avoid stock outs. This can lead to increases in sales and more accurate inventory control.

3: Using a computer with data mining capabilities. When maintaining inventory control using data mining, a retailer is not limited to sales volume and stock levels as a means to perform analysis. Data mining can go much deeper into the database to create information such as:

a) Forecast seasonal demands that match customer profiles.

b) Forecast an increase in demand due to promotion to help inventory correctly when something goes on promotion.

c) Analyze inventory that performs differently in different market segments and locations. This allows for better utility of floor space within multiple store chains.

d) Data mining can compile a sales profile for each SKU by analyzing all possible influences, such as sales trends, seasonal patterns, demographics, etc. This profile is modified over time and allows retailers to predict just-in-time purchasing way ahead of actual sales.

Scenario C – Facilitating buying decisions: "How can I determine which new products to carry?"

1: No computer. Determining new products and styles without the aid of computerized reports can be difficult. Experienced retailers may find this a less daunting task, however, figures such as turnover and gross margin, indicators of past product performance need to be generated manually. This can be a very time consuming process. Cash registers can maintain department sales information, which yield little useful information in deciding on new product purchases. Departmental sales information could be an indicator as to how general categories of merchandise are performing, but offer little help in determining specific product information.

Predicting what your customers may like without some sort of customer database can be a shot in the dark. Retailers in this situation may play a more reactive role to market demands, watching what is selling in other stores and markets and scrambling to get their hands on stock while customer demand is still high. All too often, merchandise is not available or the demand has dropped, retailers find themselves over stocked and markdowns are inevitable. Trade journals can provide buyers with forecasting information, which can prove helpful, but this type of information is often general and not targeted at their specific trading area.

2: Using a computer with traditional data analysis capabilities. Retailers who have the benefit of sales and inventory reports generated from traditional database queries have the advantage of detailed historical information as to what products have performed well in the past. This information coupled with forecasting reports from trade journals and other sources predicting customer-purchasing trends can improve the success rate of new product introduction. Traditional database analysis can also provide reports from customer databases, allowing buyers to see the type of products particular customers tend to purchase. Once again, this information is purely historical and predicting new product success is still somewhat a guessing game.

3: Using a computer with data mining capabilities. The use of data mining in making new product determinations eliminates a great deal of the guesswork. Data mining allows retailers to promote new products to the consumer rather than simply being reactive. It allows for greater understanding of various levels of customer sophistication to aid in determining new products or services (Stedman, 1997). This type of analysis can help the buyer understand whom the customer is, what they like and what to offer verses what the customer expects. In deciding which new products to stock, data mining can mine for:

a) customer complaints and pay attention to customer feedback on what product you currently sell to yield new product information

b) customer behavior with existing products, yielding new product information

c) customer aspirations rather than needs. An example of this is personal pair jeans by Levi Straus. Customizing a perfect fit is often more a hope than an expectation.

neglected or overlooked customer segments leading to new product decisions.

Table 2. Comparison of data mining versus no computer and traditional analyses.

Quality of Data Analysis

	No Computer	Traditional Analysis	Data Mining
A. Tracking promotional effectiveness:			
Total sales	moderate	high	high
Comparative sales information	low	moderate	high
Discount dollars	low	high	high
Coupon sales	low	high	high
Customer information	low	high	high
Tracking multiple promotions	low	moderate	high
Tracking add-on sales	low	moderate	high
Selling up to a higher price point	low	moderate	high
B. Inventory management:			
Tracking inventory counts	low	high	high
Maintaining accurate counts	low	high	high
Timely reorders	low	moderate	high
Avoiding stock outs	low	moderate	high
Calculating financial figures	low	high	high
Managing fast moving items	low	moderate	high
C. Buying decisions:			
Determine new products	low	low	moderate
Accessing historical data	moderate	high	high
Predicting customer preferences	low	moderate	high
Making timely purchases	low	moderate	high
Avoiding overstock	low	moderate	high
Reducing markdowns	low	moderate	high
Forecasting preferences/demand	low	moderate	high

Note:
High=high level of quality
Moderate=moderate quality
Low=low level of quality

Data Mining in Current Retail Decision Making and Strategic Planning:

A range of current uses for data mining within the retail industry demonstrates growing applications and utility within strategic planning processes.

Direct mail marketing. The Body Shop International is testing data mining to increase the efficiency of its mail order business. They are interested in cutting back on the number of catalogs they mail out while at the same time improve response rates for those customers whom already receive the catalog (Whiting, 2002). Macy's West found by analyzing promotional response data targeted to specific customer, they received a 44% response rate. The same offer sent to a general audience received an 8% response rate (Chain Store Age, Anonymous Author, 1999).

Customer relationship management (CRM) and customer profiles. Federated Department stores is combining customer and transaction data to identify best customers and offer exclusive extras. For example, Macys West by identifying customer-spending habits can predict life-changing events such as marriage and children within one year (Chain Store Age, Anonymous Author, 1999). ZCMI , a department store chain based in Utah is using data mining to integrate customer data with multiple other merchandising systems to identify popular products and specific customer categories (Rabinovitch, 1999). Sears Canada discovered through data mining technology that their best customers shop in multiple different channels and as a result are redeveloping their marketing strategies to more aggressively promote the various channels (Kroll, 2001). Neiman Marcus is using data mining technology to build and enhance customer relationships. According to Billy Payton, VP of membership programs "Our data is extremely important to us…our business is relationship based. We've found we reach our customers best through our data" (Chain Store Age, Anonymous Author, 2000).

Category management and inventory control. Rubbermaid is using data mining technology for category management and merchandising optimization (Discount Store News, Anonymous Author, 1999). Wal-Mart, Proctor and Gamble, Coke, Pepsi, and American Greetings have been using data mining for category management for several years (Parks, 2000).

Market basket analysis. J Crew Group Inc. is combining click stream analysis from its web site along with point-of-sale (POS) data from retail store locations to perform product affinity analysis. They want to determine what clothes, shoes, and accessories customers most often purchase together. The data will then be used to make complementary product suggestions for on-line shoppers (Whiting, 2001).

Web site analysis and personalization. Nordstrom is using click stream analysis to personalize the web shopping experience (Maselli, 2001). In addition, Macy's has incorporated data mining technologies to help manage and fine-tune Macys.com (Fonseca, 2000). Amazon uses web site personalization to suggest new items when certain items are purchased.

Additional possible retail applications . All of the above areas are very customer driven and on-line shopping driven. Other business objectives that data mining can address in a more effective and efficient manner than traditional data analysis techniques include:

a) analyzing store hours and staffing needs within store management

b) assortment planning in single and multi-store organizations including inventory levels/control within merchandising

c) markdown schedules and promotional timing within promotion

d) replenishment and reorder analysis with distribution

e) effective sizing, grading, cutting, sourcing, and trend analysis within product development

Other business applications. Retail is not the only business that is utilizing data mining. Credit card companies use data mining to determine which customers are most likely to respond to travel tie-ins like hotel and airfare discounts (Wheaton, 1998) and identifying individual purchasing habits to customize coupon offerings in individual statements (Peacock, 1998). A food club warehouse used data mining to discover that coupon offerings during Passover and Easter boosted sales. Data mining also helped identify that travel size toiletries sold better during Thanksgiving and Christmas, the two heaviest travel times of the year (GlobalIntel.com, Anonymous Author, 1998). Telemarketing companies are utilizing data mining to fine tune the calls they make resulting in a substantial drop in overhead costs. They target prospects predicted by demographics, lifestyle, pyschographics and purchasing history (Hollander, 1998). Brokerage and investment firms are also utilizing data mining to increase their customer base and determine who is more likely to invest in what (Press, 1998).

Existing Research:

Empirical research on data mining applications in the retail and product development industry is limited. Studies have been focusing on the on-line or e-commerce sector. Lee, Podlaseck, Schonberg, and Hock (2001) analyzed click stream data to study on-line shopping behavior as well as visualization and data mining analysis techniques to study the movement of customers through web sites as a means to better understand on-line merchandising. Path analysis has been used to study web traffic (Berkhin, Beche, and Randall 2001). Data mining has been explored in optimizing inventory levels for electronic commerce, to analyze product performance of online stores and to analyze web-based shopping systems (Dhond, Gutpa, and Vadhavkar 2000, Lee, Podlaseck, 2000, Arlitt, Krishnamurthy, and Rolia, 2001).

Research in non on-line sector studies was much more limited. Two studies identified were focused on product selection and assortment (Brijs, Goethals, Swinnen, Vanhoof, and Wets, 2000a and 2000b). Clearly, there is much more to be accomplished with this very powerful research tool.

Recommendations for Future Research:

Work towards determining a cost/benefit model would benefit the industry. The more powerful the information obtained from data mining technology spurs greater utility which in turn drives the need for more data collection which adds to the hardware/software needs which creates more powerful information and the cycle continues. All of this costs money. Where is the point of diminishing returns? How can we truly measure return on investment? Developing guidelines for users on when the cost of input (money and time) begins to exceed output (money saved, time saved, and the utility of the information obtained) could aid in effective implementation of data mining technologies and application of the results.

There needs to be a better understanding of the ethical and legal issues surrounding the use of information. Individual privacy is a key social issue in data mining technology today. It is now possible to glean a significant amount of information about a particular customer by analyzing purchasing behavior (www.anderson.ucla.edu, 2002). Additional concerns center around the issue of data ownership and use (Thearling, 2002). Government regulations such as the FDA impact data integrity and traceability and there are situations where the data being analyzed cannot be legally used in decision-making such as race, gender, and age (Anderson, 2001; Thearling, 2002).

Research needs to be done toward helping develop better models and applications that focus specifically on the retail/merchandising/product development industry. Rabinovitch (1999) stated that retailers are more concerned with the application of data mining results and the impact on the bottom line rather than the technical jargon of how it works. One way to accomplish this is to analyze who is using it, why, and look closely at successes and failures.

Finally, more discussion is warranted on the entire process of technology and decision-making. What part does technology really play in decision-making? There exists a fallacy that data equals information that equals knowledge, which results in action (Keen, 1997). Much research has been done on the supply side of information: gathering (point of sale scanners and coupon cards, automatic teller machines, etc.), sharing (internet, intranet, email, etc.), and analyzing information (traditional statistical techniques, data mining). More time and emphasis needs to be placed on the action piece rather than on the data collection and analysis piece. Perhaps the focus of research needs to shift to the other end of the spectrum.

“Companies whose data mining efforts are guided by “mythology” will find themselves at a serious competitive disadvantage to those organizations taking a measured, rational approached based on facts” (Small, 2002, pg. 3). As electronic databases continue to grow, data mining will continue to make advances, producing predictive information and helping businesses leverage their data to facilitate more effective and efficient proactive decision-making.

References

Anderson, J.L. (2001). Profiles of successful graduate students using artificial neural networking and simultaneous multiple regression. Unpublished Dissertation, Colorado State University.

Arlitt, M., Krishnamurthy, D., & Rolia, J. (2001, August). Characterizing the scalability of large web-based shopping systems. *ACM Transactions on Internet Technology, (1),* 1, 44-69.

Berkin, P., Beche, J.D., & Randall, D.J. (2001). Interactive path analysis of web site traffic. *American Association for Artificial Intelligence, Proceedings*. 414-419.

Berry, M.J.A., & Linoff, G. (1997). *Data mining techniques for marketing, sales and customer support.* New York: John Wiley & Sons, Inc.

Brijs, T., Swinnen, G., Vanhoof, K., & Wets, G. (1999). Using association rules for product assortment decisions: A case study. *Data Mining and Knowledge Discovery*, 254.

Brijs, T., Goethals, B., Swinnen, G., Vanhoof, K., & Wets, G. (2000). A data mining framework for optimal product selection in a retail supermarket: The generalized PROFSET model. *Data Mining and Knowledge Discovery*, 300.

Cabena, P., Hadjinian, P., Stadler, R., Verhees, J & Zanasi, A. (1998). *Discovering data mining: From concept to implementation.* Upper Saddle River, N.J.: Prentice Hall.

Data mining/CRM: Searching for an ROI. (2001, October). *Chain Store Age*, 24-26.

Data mining: What is data mining? (no date). [online]. Available: http://www.anderson.ucla.edu/faculty [2002, June 11].

Data Intelligence Group. (1995, September). An overview of data mining at Dun and Bradstreet. [online]. Available: http://www3.shore.net [2002, June 11].

Decker, P. (1998, Mar/Apr). Data mining's hidden dangers. *Banking Strategies, 74*(2), 6-14.

Dhond, A., Gupta, A. & Vadhavkar, S. (2000). Data mining techniques for optimizing inventories for electronic commerce. *Data Mining and Knowledge Discovery*, 480-486.

Do-it-yourself data mining. (1996, December). *Marketing Computers, 16,* (11), 1-2.

Foley J. & Russell, J.D. (1998, March 16). Mining your own business. *Informationweek, 673,* 18-20.

Fonseca, B. (2000, July 7). Personalization and web analysis sink in. *Infoworld, (22),* 29, 39.

Growing lifetime value. (1999, May). *Chain Store Age, (75),* 5, 143.

Hollander, G. (1998, May 25). Model 1 deftly pares customer characteristics. *Infoworld, 20,*(21), 148.

Kempster, M. (1998, June). Digging for data. *Inform, 12*, (6), 22-30.

Keen, P.G.W. (1997, November 17). Let's focus on action not info. *Computerworld* [online]. Available: http://www.computerworld.com. [2002, June 11].

Kroll, K.M. (2001, April). Sears Canada gains improved customer knowledge with data mining and analysis. *Stores, (83),* 4, 128-129.

Lee, J. & Podlaseck, M. (2000, October). Using a starfield visualization of ranalyzing product performance of online stores. *Data Mining and Knowledge Discovery,* 168-175.

Lee, J., Podlaseck, M., Schonberg, E., & Hoch, R. (2001, January-April). Visualization and analysis of click stream data of online stores for understanding web merchandising. *Data Mining and Knowledge Discovery, (5),* 1-2, 59-84.

Maselli, J. (2001, February). Business-intelligence pays off for Nordstrom. *Information Week, (825),* 91.

Mining the store. (2000, January). *Chain Store Age, 76*, (1), 4-5.

Murphy, P. A. (1998, March). Focus on existing customers seen driving revenue growth. *Stores*, 32-34.

Parks, L. (2000, June). Information resources inc. *Drug Store News, (22),* 9, 219-222.

Peacock, P.R. (1998, Winter). Data mining in marketing: Part 1. *Marketing Management*, 9-18.

Press, S. (1998, June). Fool's gold. *Sales and Marketing Management, 150,* (6), 58-62.

Rabinovitch, L. (1999, December). America's "first" department store mines customer data. *Direct marketing, (62),*8, 42-44+.

Ritzer, J.R. (1996, August). New rule-based systems tackle employee shrink. *Stores, (78),* 8, 71-72.

Rubbermaid mines for profits. (1999, March). *Discount Store News, (38),* 6, S3-S5.

Small, R.D. (1997, January). Debunking data mining myths. *Informationweek*, [online]. Available: http://www.informationweek.com. [2002, June 11].

Stedman, C. (1997, December). Data mining despite the dangers. *Computerworld, (32),* 1, 61-62.

Thearling, K. (2002). More than algorithms: Data mining in the real world. [online]. Available: http://www.thearling.com. [2002, June 11].

Thearling, K. (2002). An introduction to data mining. [online]. Available: http://www.thearlin g.com. [2002, June 11].

What is data mining? (No date). [online]. Available: http://www.globalintel.com/data.html [1998, Sept. 13].

Wheaton, J. (1998, May). Data detectives. *Catalog Age, 15,* (6), 94-100.

Whiting, R. (2002, February). Companies boost sales efforts with predictive analysis. *Information Week, (877)*, 62-64.

Whiting, R. (2001, November). Retailer seeks success on how customer dress. *Information Week, (865)*, 65.

Whiting, R. (2002, March). Business intelligence: Feasible or feeble? [online]. Available: http://www.informationweek.com. [2002, June 11].

AUTHORS' INFORMATION

Joan Anderson, Ph.D.
Department of Apparel, Merchandising, and Interior Design
Washington State University
Pullman, Washington 99164
509-335-8399
joana@wsu.edu

Antigone Kotsiopulos, Ph.D.
College of Applied Human Sciences
Colorado State University
Fort Collins, Colorado 80523
970-491-7740