Mining Your Own Business

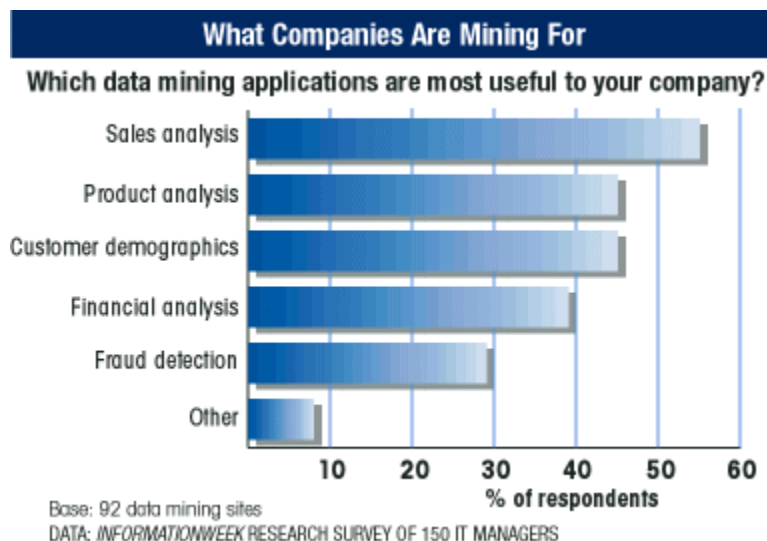**Vendors seek to ease deployment as more companies look to data mining to turn data into profits**

By John Foley and Joy D. Russell

With transaction volumes rising and databases getting bigger, companies are turning to data mining as a way to gain insights into the vast amounts of data they're generating and collecting. ITT Sheraton, the National Association of Securities Dealers, Safeco, and Wells Fargo are among the organizations launching applications that range from identifying which hotel guests might want a cigar in their rooms to detecting stock market fraud.

Data mining, which uses mathematical algorithms to find patterns in data, is in the midst of a growth spurt. New products and customer demand will boost the market to $8.4 billion in 2000 from $3.3 billion in 1996, according to the Meta Group.

Leading suppliers, including DataMind, IBM, NeoVista, SAS Institute, and Silicon Graphics, have recently introduced or are about to unveil packages that address some long-standing barriers to deployment. Increasingly, these products are being tuned for specific applications--such as fraud detection or customer-relationship management--and for vertical industries.

What Companies Are Mining For

Which data mining applications are most useful to your company?

- Sales analysis
- Product analysis
- Customer demographics
- Financial analysis
- Fraud detection
- Other

% of respondents

Base: 92 data mining sites
DATA: *INFORMATIONWEEK* RESEARCH SURVEY OF 150 IT MANAGERS

Also, two leading software vendors--Oracle and Microsoft--are about to put their stamp on the market. Next week, Oracle will unveil a marketing and technology pact with seven data mining suppliers, including plans to integ rate their products with the next release of its database, Oracle8.1. That release, due later this year, will include capabilities such as data sampling that are commonly used in data mining operations. "The enhancements will further reduce the barriers for making data mining a basic part of a decision support system," says George Lumpkin, product manager for data warehousing at Oracle. Oracle's data mining partners include Angoss, DataMind, Datasage, Information Discovery, SPSS, SRA International, and Thinking Machines.

Microsoft is incorporating data mining technology into its electronic-commerce system. Microsoft's Site Server 3.0 Commerce edition, due next quarter, will come with a feature called Intelligent Cross Sell, a data mining algorithm that can be used to analyze the activity of shoppers on a Web site and automatically adapt the site to that user's preferences. It's the first time Microsoft has made data mining a standard feature in one of its products.

"My vision is that this ki nd of information can be used to completely optimize a Web store, to reorganize the appearance of a store to serve you better," says Usama Fayyad, a senior researcher with Microsoft Research, which developed the feature. A Microsoft Research project called Socrates is studying the potential use of data mining in very large databases running on Microsoft platforms.

These vendors are reacting to a push by companies to exploit their data to gain competitive advantages. ITT Sheraton uses data mining software from Datasage, in Reading, Mass., as part of a marketing push to better understand its customers. "We're not heavily into data mining yet, but we're setting the stage," says David Van Kalsbeek, ITT Sheraton's senior VP and director of marketing and strategic planning.

As a first step, the hospitality chain is running Datasage against data in Oracle's Express multidimensional database. Next month, the results of that pilot will be presented to senior management to request approval for a full implementation. If Van Kalsbeek gets a thumbs-up, the next step will be to build a data warehouse for full-fledged mining of customer information. For example, Van Kalsbeek says Sheraton is interested in evaluating "acquisition-versus-retention expenditures" to determine how much the hotel should spend to grow its customer base versus how much it should spend to keep current customers coming back. "As a marketing person, I'm licking my lips," says Van Kalsbeek. "Data mining can actually show us who's inclined to do what and when."

Some users see data mining as more than a marketing option. The National Association of Securities Dealers' regulatory division, NASD Regulation Inc. in Washington, is deploying data mining to monitor NASDAQ and over-the-counter stock markets for what it calls unfair trading practices. NASD's first two data mining applications were deployed last year; a third is scheduled for production this month. Already, the watchdog group reports a 400% increase in the number of useful "breaks," or irregular activities, it uncovers.

NASD Regulation uses its data mining system, called the Advanced Detection System, to monitor 900,000 stock trades and 1.5 million quotes each day. "If you had to eyeball 900,000 trades, you just can't do it as well," says Randeen Klarin, director in the office of technology services with NASD Regulation. "The system tells you which ones to look at." NASD's Advanced Detection System uses data mining software from SRA International Inc., in Arlington, Va., running on an Oracle8 database and a Sequent NUMA-Q computer.

The project has required close collaboration between NASD's technology and business managers, and maintaining optimal system performance continues to be a challenge. "It has been more challenging than I ever expected," says Lowell Cooper, a senior technical specialist with NASD Regulation. Still, data mining has

been invaluable in helping to detect fraud, Cooper says, adding: "The answer is in the data."

Too Difficult?
Nearly two-thirds of IT managers surveyed by InformationWeek say data mining products are still difficult to use and administer. "In its current generation, data mining tools are a little premature and require technical expertise to manipulate correctly," says Stephen Cole, director of analysis at American Century Investments, a mutual fund company in Kansas City, Mo., that uses data mining software from SPSS Inc. to understand customer behavior for marketing its financial products.

Data mining software providers are attempting to ease some of the deployment pain right now. Last month, IBM introduced data mining software and service packages, called DecisionEdge, for financial and insurance companies. This week, DataMind, in San Mateo, Calif., will unveil a data mining software bundle designed to automate marketing efforts by comparing proposed campaigns against existing customer data.

Scott Toborg, a data mining specialist with SBC Technology Resources, the advanced technologies unit of Bell company SBC Communications, says small businesses and departments of large companies can get the benefits of data mining without hiring specialists by using some of the new packaged products. "If the data is clean and you have nicely focused vertical applications, then it's pretty easy to use," Toborg says. "If you step outside of that, you're going to need some pretty beefy help."

Microsoft is targeting that low end of the market, at least to start. The company's upcoming data mining capability, Intelligent Cross Sell, can be installed from a wizard in Site Server with the click of a mouse, says Fayyad. While Microsoft hasn't revealed plans to add data mining to its other products, its researchers are already thinking along those lines.

Seeking Integration
Would your organization be interested in data mining products that are integrated with packaged applications or databases?

Don't know 5%
No 20%
Yes 75%

Base: 92 data mining sites
DATA: INFORMATIONWEEK RESEARCH SURVEY OF 150 IT MANAGERS

At the other end of the spectrum are mainframe-class data mining platforms such as Wells Fargo's Sun Microsystems Enterprise 10000 server running Oracle databases and NeoVista Software Inc.'s data mining algorithms. David Holvey, a senior VP and director of marketing information services with Wells Fargo, in San Francisco, admits that such an environment can be complex to manage and maintain. But Holvey says Wells Fargo wants to get under the hood of its data mining applications so it can fully understand the results it's getting. "Ease of use is almost the antithesis of understanding what processing is taking place," he says.

Data mining products range in price from $1,000 for software that runs on a PC to more than $100,000 for algorithms that run on mainframes or massively parallel computers. But price doesn't appear to be an issue at most companies. Only 5% of companies that have yet to deploy data mining software listed price as a reason in the *InformationWeek* Research survey.

For those companies that do make the move, what's the return on investment? Surprisingly, many companies report only moderate returns from their data m ining efforts. Technology managers rated ROI 5.6 on a scale of 10 in the InformationWeek Research survey. Another explanation: Some of data mining's best successes are well-kept secrets. Wells Fargo and Mobil Corp. both declined to discuss their data mining applications for this story.

More than half the revenue generated by the data mining market goes to what Meta Group calls data-service providers--companies like Acxiom Corp. in Conway, Ark., that mine data in their own proprietary databases and sell the results to corporate customers. As more companies bring data mining functions in house, the service providers' market share is shrinking.

Still, for companies such as Safeco Corp., a $6 billion insurer in Seattle, the option of letting outside experts manage the data mining application is appealing. "It takes highly specialized personnel to run the software, build the model, and read the results," says Doug Gillette, a senior business analyst with Safeco. "In none of those cases do we have the experience."

Safeco has turned to NeoVista for help. Safeco is three months into a pilot to determine whether data mining can improve customer retention in its auto insurance business. NeoVista, in Cupertino, Calif., provides the technical experts to model the data and the data mining platform. Gillette says it's too early to know how Safeco will proceed with data mining--but the results so far look promising. And the word is getting around. "In the process, we've heard from other areas of the company that might be interested," he says.

It's the same story at SBC. The telecom carrier is in the early stages of deploying data mining in a number of business units, says Toborg. SBC's telephone operating company is evaluating data mining for market segmentation and sales opportunities, while its wireless and Internet subsidiaries want to use it to reduce customer churn. SBC's Yellow Pages unit is interested in the technology to understand customer life cycles. In its research lab, SBC uses d ata mining to analyze traffic patterns in its network and predict problems. Says Toborg, "Interest keeps popping up all over the place."