

Data Mining -- Exploiting the Data Warehouse

by Joe Bigus

The changes in the business computing environment over the past three decades have been dramatic. Computer processing and storage technology advances provide businesses with the ability to keep hundreds of gigabytes or even terabytes of data on-line. However, this is a good news, bad news story. The good news is that now we can have years of historical business data available for decision support applications. The bad news is that traditional data query and analysis methods are not capable of dealing with that much data. The consequence is that businesses are drowning in data.

Data mining or knowledge discovery offers a solution to this problem. With the emphasis on the discovery of valuable information from large databases, data mining provides added value to the investment in the corporate data warehouse. The data mining process comprises three basic steps, data preparation, information discovery by a data mining algorithm, and analysis of the mining algorithm output.

The benefits of data mining are evident in two major business activities, decision support and application development. In decision support systems, data mining transforms the data to reveal hidden information in the form of facts, rules, and graphical representations of the data. The extremely large amounts of data are compressed to reveal the inner relationships among the data elements. When used in the application development cycle, data mining with neural networks provides automated construction of transaction processing systems and forecasting models.

Applications of data mining span all industries. Businesses of all types use data mining to target marketing messages to specific customer sets, both to satisfy their customer's needs and to increase revenues. Retailers use data mining to find associations between products purchased at the same time and to forecast sales and corresponding inventory requirements. The Finance industry uses data mining techniques to manage risks and to detect trends in the markets. Manufacturers use neural networks in the design, production scheduling, process control and quality inspections of their products. Hospitals and insurance companies mine their data to detect fraudulent claims by health care providers and patients, and physicians use advanced pattern recognition capabilities of neural networks to automate laboratory tests. Utilities use neural networks to forecast demand and quickly respond to equipment outages and changes in the weather.

Any business with data about their customers, suppliers, products, and sales can benefit from data mining. When businesses are looking for the slightest edge over their competition, they are willing to travel far and wide, and spend millions of dollars to buy information about their markets. Often this information is sitting right in their offices, hidden away in their data warehouses.

[From *Data Mining with Neural Networks*, Joseph P. Bigus, McGraw-Hill Inc., 1996]



IEEE Computer Society Meeting

Featuring Joe Bigus: Data Mining

Monday, October 6, 7:00 pm
Mayo Medical Sciences Building
(321 3rd Avenue SW, Rochester)

🍕 Pizza & socializing at 6:30 pm 🍷

In this presentation, Joe will discuss the business environment and the information technology trends which have made data mining both necessary and achievable. He will provide a formal definition of data mining, and describe the major steps in the data mining process. Finally, he will describe some of the many data mining applications which have been developed.

Joe Bigus is the software architecture responsible for the neural network components of the IBM Intelligent Miner data-mining products. He holds 16 U.S. patents dealing with neural networks and with artificial intelligence applications. He is the author of two books: **Data Mining with Neural Networks**, and the forthcoming **Constructing Intelligent Agents with Java**. Joe holds a B.S. in Computer Science from Villanova University, and an M.S. and Ph.D. in Computer Science from Lehigh University.

IEEE Section Meeting

Featuring George O'Clock

Monday, November 17, 7:00 pm
Mayo Medical Sciences Building

See the next (November) newsletter for info about the topic of George O'Clock's presentation at our November meeting. There will be no Southern Minnesota section meeting in October (just the Computer Society meeting on October 6).

IEEE Southern Minnesota Section Board Members

<i>Chair</i>	Ron Jensen	253-3887	<i>ron@rchvmw2.vnet.ibm.com</i>
<i>Vice Chair</i>	Rob Harveland	253-0780	<i>rharvela@vnet.ibm.com</i>
<i>Secretary</i>	Bill Gorder	253-1409	<i>w.gorder@ieee.org</i>
<i>Treasurer</i>	Steve Kerchberger	253-4047	<i>s.kerschberger@ieee.org</i>
<i>Awards</i>	Tom Roettger	253-4796	<i>t.roettger@ieee.org</i>
<i>Computer Society</i>	Duane Wenzel	253-1035	<i>d.j.wenzel@ieee.org</i>
<i>Computer Society</i>	Pat Wong		<i>p.c.wong@ieee.org</i>
<i>Home Page</i>	Diane Manlove	253-7613	<i>dmanlove@prodigy.com</i>
<i>Membership</i>	Tim Wolf	455-4304	<i>t.wolf@ieee.org</i>
<i>Newsletter</i>	Chris Kimble	253-7571	<i>c.kimble@ieee.org</i>
<i>PACE</i>	Dick Hall	282-7085	<i>dick.hall@ieee.org</i>
<i>Power Society</i>	Neil Stiller	280-1647	<i>nstiller@rpu.org</i>
<i>Pre-College</i>	Vince Lynch	433-0456	<i>vmlynch@wolf.co.net</i>
<i>Student Activities</i>	Paul Dahlinger	253-1769	<i>dahlinger@aol.com</i>

**The Institute of Electrical
and Electronics Engineers, Inc.
Southern Minnesota Section**

Dated Material -- *Please Deliver Immediately*

Non-Profit U.S. Postage Paid Permit #511 Rochester, MN
--

IEEE Power Society Chapter Possible

Neil Stiller, Rochester Public Utilities, has joined the board of the IEEE Southern Minnesota Section to work towards the formation of a Power Society chapter within the section.

See the next (November) newsletter for details about the first Power Society event, tentatively scheduled for November 13.

Minnesota Power Systems Conference

The 33rd annual Minnesota Power Systems Conference will be held October 7-9 at the University of Minnesota. The conference offers an opportunity for power engineers to keep abreast of today's power system technology. Topics include planning, substation, operations, distribution, communications, relaying, and utility industry futures. Call 612-626-2259 for more information.