

: *What is data mining and how is it related to DSS?*

by Daniel Power

Editor, DSSResources.com

Data mining is a data analysis innovation first discussed in the 1990s. "Big data" and analytics has led to a renewed and expanded interest in data mining technologies. Academics tend to use the related terms Knowledge Discovery and Intelligent Decision Support Methods (Dhar and Stein, 1997) or more derogatory terms like data surfing or data dredging. In general, data mining is a group of analytical methods like neural networks, genetic algorithms, and decision trees, that help people conduct computerized searches for patterns in a data set. Data mining is both a process and a set of tools.

The goal of data mining is to find interesting and relevant patterns in a set of data. Many data miners hope to find "hidden" patterns or relationships that can be used to predict future behavior. DSS developers sometimes incorporate results from data mining into operational decision support systems (DSS).

The data mining process involves identifying an appropriate data set to "mine" or sift through to discover data content relationships. Additional data mining tools include techniques like case-based reasoning, cluster analysis, data visualization, and fuzzy query analysis. Data mining sometimes resembles the traditional scientific method of identifying a hypothesis and then testing it using an appropriate data set. Sometimes however data mining is reminiscent of what happens when data has been collected and no significant results were found and hence an ad hoc, exploratory analysis is conducted to find a significant relationship.

Data mining has helped identify meaningful relationships and when it is done well the results should be useful in business decision making. In particular, data mining can conceivably be a major part of an ad hoc, decision support special study. Special studies are usually prepared to support decision-making in situations that are especially important and one-time (novel or infrequent). Data can be mined with a specific purpose in mind and statistically significant results can be reported to managers. What is not always clear is how data mining is related to building decision support systems. Some commentators imagine we should provide managers with a data mining tool and let them mine data until they have thoroughly understood the relationships that are "hidden" in the data set. This vision doesn't seem too fruitful or too desirable. It is appropriate for a trained decision support or data analyst to work with data mining tools to prepare decision support special studies, but most managers won't have the interest or skills to participate in such an activity.

So is data mining relevant to building DSS? Yes, I think it is if we are realistic about what is possible. First, data mining can help identify relations and rules that can be incorporated in knowledge-driven

: *What is data mining and how is it related to DSS?*

DSS. Second, case-based reasoning can be used to create a specific Knowledge-driven DSS that can be used by a manager or a knowledge worker who is trying to diagnosis problems in that "case" environment. Third, data visualization tools can be incorporated with a structured data set to assist managers in making a recurring decision where the data set is routinely updated. For example, a stock portfolio manager may find that a data-driven DSS with visualization tools may help understand the composition of the portfolio and help identify what changes need to be made in its component stocks. Fourth, other tools like neural networks may also have a place in creating capabilities in specific DSS. For example, rather than using only a heuristic scoring model and possibly a risk analysis model for supporting commercial loan decision making there may be some situations where a neural network model from a database of prior loans could also inform and support the decision maker. One can also identify DSS applications that use data mining tools in fraud detection, category management and direct marketing.

Some specific uses of data mining include: 1) customer churn - develop a model to predict which customers are likely to leave your company and go to a competitor; 2) fraud detection - routinely screen transactions to identify those most likely to be fraudulent; 3) direct marketing - develop a model to predict prospects to target in email, sales calls, and direct mailing to obtain the highest response rate; 4) web and interactive marketing - use a model in real-time to predict what each individual accessing a Web site is most likely interested in viewing and purchasing; 5) market basket analysis - determine what products or services are commonly purchased together; 6) market segmentation - determine characteristics of customer groups and their purchase behavior; and 7) trend analysis - forecast customer purchase behavior.

Data mining tools are relevant to building DSS when the decision situation warrants the use of such tools and when the DSS Builder understands there uses and limitations.

For more information on data mining visit Gregory Piatetsky-Shapiro's KDnuggets website (<http://kdnuggets.com/>) or the ACM Special Interest Group on Knowledge Discovery in Data and Data Mining web site (<http://www.kdd.org/>).

References

Dhar, V. and R. Stein, *Intelligent Decision Support Methods: The Science of Knowledge*, Upper Saddle River, NJ: Prentice-Hall, 1997.

The above response is based upon Power, D., *What is data mining and how is it related to DSS?* DSS News, Vol. 2, No. 25, December 2, 2001 (updated November 29, 2015).

: *What is data mining and how is it related to DSS?*

Author: Daniel Power

Last update: 2015-11-30 11:02