Chapter VIII
Introduction to Data Mining and its Applications to Manufacturing

Jose D. Montero
Management Information System Department, Utah State University, USA

ABSTRACT

This chapter provides a brief introduction to data mining, the data mining process, and its applications to manufacturing. Several examples are provided to illustrate how data mining, a key area of computational intelligence, offers a great promise to manufacturing companies. It also covers a brief overview of data warehousing as a strategic resource for quality improvement and as a major enabler for data mining applications. Although data mining has been used extensively in several industries, in manufacturing its use is more limited and new. The examples published in the literature of using data mining in manufacturing promise a bright future for a broader expansion of data mining and business intelligence in general into manufacturing. The author believes that data mining will become a main stream application in manufacturing and it will enhance the analytical capabilities in the organization beyond what is offered and used today from statistical methods.

INTRODUCTION

There are many driven factors leading manufacturing companies to embrace manufacturing intelligent solutions. Among them are the rigorous customer specifications for designing, developing, and testing of products, the complexity of products, the need for rapid market response, the pressures of global competition, and the strict regulatory requirements. Competitiveness, productivity, and efficiency in the global economy will be affected by how manufacturing companies utilize their computational resources for decision support processes beyond the traditional uses of those resources for just operation processes.

Manufacturing companies are notorious for accumulating mountains of operational data from design, laboratory testing, engineering and research, manufacturing processes, and testing of final products. Some data bases collect operational and historical data of many years. How can manufacturing companies overcome the data barriers presented from these large and in many situations disparate and nonintegrated data bases? How can they leverage the large amounts of data to extract knowledge to support the decision making process? The answers to these questions include the adoption of data integration and data mining. This last subject is the main focus of this chapter.

It has been said as a cliché that "today's high-tech world is drowning in data but is starved for knowledge". That could also be applicable to manufacturing, as well as to other areas of science and engineering. The gap between data and their analysis is growing, and opportunities for extracting valuable knowledge from those data are being lost. This is well expressed by NASA Ames, in the Web site of their Intelligent Data Understanding project.

Many scientists are deluged with data, and the gap between data collection and analysis is growing. Data can be archived for later use, but at a cost (including lost knowledge about the data). There are still LANDSAT data sets from the 70's and 80's that have not been analyzed. (NASA AMES, 2005)

To launch the new millennium, the January/February 2001 issue of *Technology Review*, MIT's magazine of innovation, hailed data mining as one of the top 10 technologies that will change the world in the 21st century (MIT, 2001; Van der Werff, T. J., 2001). These 10 emerging areas of technology have been predicted to have a profound impact on the economy and how we live and work. Data mining, since its most prominent emergence in late 1990s, has revolutionized everything from how companies monitor customers' online purchasing habits and how supermarkets place products in their shelves, to how the federal government practices counterterrorism in the post 9/11 world.

Data mining has been applied successfully in a wide range of businesses in the last decade. It is primarily used in retail, insurance, finance, banking, communications, and direct marketing (Braha, 2001; Han & Kamber, 2001). However, in manufacturing, the application of data mining has started only a few years ago. Factors attributed for slow start of data mining in manufacturing include academic researchers in data mining not being familiar with manufacturing, the majority of researchers in manufacturing are not familiar with data mining, and the restricted access or limitations to publish proprietary and sensitive enterprise data.

The available volumes of data in manufacturing provide many opportunities for knowledge extraction with data mining. Some areas of engineering and manufacturing where data mining has been used with great success, as reported in the literature, are fault diagnosis, process and quality control, process analysis, maintenance interval prediction, production and research process optimization, resource management, and process modeling.

In just a few short years the application of data mining methods in manufacturing has proven very effective in improving both the quality of products and the quality of decision making. It has also provided substantial savings in time and money. One example of this is the Motorola Laboratories' experience in using data mining (Gardener & Bieker, 2000). On the other hand, there is no free lunch. Data mining projects may require a substantial investment. This is particularly true in organizations without well integrated data, or

Related Content

Simulation and Visualization of Concept Algebra in MATLAB
Xiaoyu Linand Yingxu Wang (2014). *International Journal of Software Science and Computational Intelligence (pp. 30-55).*
www.irma-international.org/article/simulation-and-visualization-of-concept-algebra-in-matlab/114095

A Fig-Based Method for Prediction Alumina Concentration
Jun Yi, Jun Pengand Taifu Li (2012). *International Journal of Software Science and Computational Intelligence (pp. 41-50).*
www.irma-international.org/article/a-fig-based-method-for-prediction-alumina-concentration/88926

Quotient Space-Based Boundary Condition for Particle Swarm Optimization Algorithm
Yuhong Chi, Fuchun Sun, Langfan Jiang, Chunyang Yuand Chunli Chen (2013). *Advances in Abstract Intelligence and Soft Computing (pp. 31-42).*
www.irma-international.org/chapter/quotient-space-based-boundary-condition/72771

Experience-Based Approach for Cognitive Vehicle Research
Hironori Hiraishi (2020). *International Journal of Software Science and Computational Intelligence (pp. 60-70).*
www.irma-international.org/article/experience-based-approach-for-cognitive-vehicle-research/262588

Delay-Range-Dependent Robust Stability for Uncertain Stochastic Neural Networks with Time-Varying Delays
Wei Fengand Haixia Wu (2012). *Breakthroughs in Software Science and Computational Intelligence (pp. 418-432).*
www.irma-international.org/chapter/delay-range-dependent-robust-stability/64622