

Data Mining Technologies Inc.

Knowledge Discovery From Databases

Data Mining and Process Improvement and Control

White Paper

Data Mining Technologies Inc.
Melville, NY 11714
(631) 692-4400 x100
e-mail mgilman@data-mine.com
www.data-mine.com

February 2006

Data Mining and Process Improvement and Control

Keywords: Manufacturing, Process Control, Power Generation, Predictive Analytics

Summary

Data Mining can often be successfully employed in improving manufacturing processes. Among the areas where improvement is possible include cost savings of raw materials, higher first quality percentage, lower equipment maintenance costs, compliance with toxin and pollutant requirements and avoidance or postponement of capital expenditures due to increases in productivity.

What is the Key Issue In Process Control?

What settings of our process will make the results “better”?

Many manufacturing processes are well suited to Data Mining because they are controlled using sophisticated process control systems. Sensors keep track of and store process parameters such as pressures, flow rates temperatures, efficiencies and other variables that are needed to monitor the process. Thus huge amounts of data are collected and stored.

Computers monitor the output from the sensors and order adjustments to keep the process within the proper levels. The question is – what are the proper levels? There are so many variables, each with a range of operational values that change with time and have complex non-linear interaction with the other variables that even experienced and expert human *operators are often unaware of what the desirable levels of the parameters are at any given time.*

The consequence of an improperly controlled manufacturing process can be costly because of some or all of the following factors: lowering of efficiency, generation of unsafe levels of pollutants or contaminants, increase in levels of equipment failure and uneconomical operation.

Even with the richness of historical operating data not all Data Mining Software is suitable for Process Control and Management.

Many of the process variables are correlated with each other and statistical methods are not well suited to this fact. For example temperature and pressure are often correlated, as are flow rates. The large number of predictor variables is also difficult for most methods to handle. Thus “variable reduction methods” are required to “prune” the number of variables to a limit of 10 or 15 which introduces noise and inaccuracies to the modeled results.

What Type of Data Mining Tool is Appropriate?

The desirable characteristics of an appropriate data mining tool are:

- Produces accurate results
- Not a “black box”
- As few limitations on the data as possible
- Handles large amounts of data (up to 50,000 variables and “unlimited” rows)
- Provides Validation Capabilities
- Handles Constraints on Variables
- Predicts
- Generalizes
- Generates Business Rules
- Handles Missing Data
- Handles Large Problems
- Deals with non-linear interactions among variables
- Easy for non-statisticians to use

Data Mining Technologies Modeling and operational toolkits handle the above requirements. Nuggets is a robust well tested technology employing Artificial Intelligence methods to uncover useful and hidden patterns in data.

Benefits of Using Nuggets

- Can Defer Large Capital Expenditures by Optimizing Current Assets
- Can Develop Models Which Respect Cost or Other Operating Constraints
- Can Operate In Real Time or Offline Mode
- Can Generate Results Quickly
- Can Generate Business Rules Understood by Analysts and Management
- Can Help Cleanse Data
- Can Help Determine Key Operating Variables