Chapter 54

DATA MINING FOR SELECTION OF MANUFACTURING PROCESSES

Bruno Agard
École Polytechnique de Montréal

Andrew Kusiak
The University of Iowa

Abstract Data Mining tools extract knowledge from large databases. The data generated in manufacturing has not been entirely exploited. This chapter discusses applications of Data Mining in a manufacturing environment. A methodology for selection of manufacturing processes is proposed and illustrated with an industrial scenario.

Keywords: manufacturing, process selection

1. Introduction

Enterprise Resources Planning (ERP) systems generate large volumes of data. The data collection efforts are often driven by productivity improvements. This large quantity of data makes it almost impossible for a person to develop a complete understanding of the entire process without using any tools (see Figure 54.1).

Examples of data sources in manufacturing include:

- Schedules.
- Production capacity, efficiency, failures, etc.
- Manufacturing parameters.
- Process quality.
- Process plans.

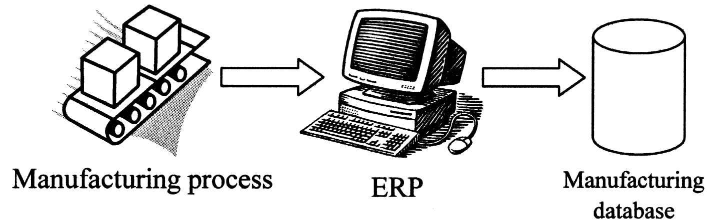

Figure 54.1. Data Generated in an ERP Environment.

The knowledge discovered from the ERP data may benefit a company. The focus of this chapter is on extracting knowledge from industrial databases in support of selection of manufacturing processes.

Some of the data-mining applications of interest to the methodology presented in this chapter are reviewed in Section 2. Section 3 presents a methodology for the selection of manufacturing processes with Data Mining, an example scenario is included. Section 4 concludes the chapter.

2. Data Mining in Engineering

Anand and Büchner (1998) defined Data Mining as the discovery of non-trivial, implicit, previously unknown, and potentially useful and understandable patterns from large data sets.

Westphal and Blaxton (1998) categorized Data Mining functions as classification, segmentation, description and estimation.

- Classification involves assigning labels to new data based on the knowledge extracted from historical data.

- Segmentation (called also clustering) divides a population into smaller sub-populations with similar behaviour according to a predefined metric. It maximizes homogeneity within a group and maximizes heterogeneity between the groups.

- Description and visualization are used to explain the relationships among the data. Frequent patterns may be extracted in the form of $A \Longrightarrow B$ rules with two measures of quality: the support which represents the number of times A occurs as a fraction of the total number of examples and confidence which expresses the number of times B exists in the data when A is present.

- Estimation focuses on filing in a missing value in a particular field of an incoming record as a function of other fields in the record (usual statistical regression techniques and neural network are most often employed).

Büchner *et al.* (1997) focused on the applications of Data Mining in manufacturing. Leu *et al.* (2001) applied Data Mining in support of the tunnel design. Prediction of the quality products in the semiconductor industry was discussed in (Kusiak, 2001). Gertosio and Dussauchoy (2004) applied Data Mining to reduce the processing time for checking and adjusting electronically controlled truck diesel engines. Agard and Kusiak proposed a methodology based on multiple Data Mining techniques for the design of product families (Agard and Kusiak, 2004) and applications of Data Mining in a Computer Integrated Manufacturing Environment (Agard and Kusiak, 2005).

Many algorithms are available for extracting knowledge from databases (Fayyad *et al.*, 1997). The efficiency of these algorithms may be improved with decomposition methods (Kusiak, 2000).

The following section focuses on a methodology for the selection of manufacturing processes based on a Data Mining approach. The methodology is illustrated with an example.

3. Selection of Manufacturing Process with a Data Mining Approach

A company manufacturing diverse components could establish a database describing all previously manufactured parts, their characteristics, and the processes used for their manufacturing (see Table 54.1).

The data in Table 54.1 shows that Part 1 has characteristics C1, C2, C4 and C6. The manufacturing process of this part is Machine M1 with Setting S_{11}, Machine M2 with Setting S_{12}, and Machine M4 with Setting S_{13}. Part 2 shares characteristics C1, C2, C3, C5 and C7 and involves Machine M1 with Setting S_{21}, Machine M2 with Setting S_{22}, Machine M3 with Setting S_{23} and Machine M5 with Setting S_{24}.

Each Setting S_{ij} may describe all information about the manufacturing settings of Part i in Process j. This information may concern NC programs, cutting parameters, tools, speed, temperature, voltage, duration, etc.

Consider the descriptive characteristics pertinent to the manufacturing process of a part. Characteristics (C1 to C7) may be features (e.g., a hole), it may contain a more specific meaning (the diameter of the hole, and the expected quality) that impacts the process.

For example, the type of drilling machine used, impacts the quality of a hole. In other cases, the material may be important, e.g., metallic vs plastic parts.

The group technology approach (Hyer, 1984) involves grouping parts, products, and processes to take advantage of their similarities driven by the economy of scale. Numerous algorithms are available to define these groups ac-

Table 54.1. Description of Parts and the Corresponding Manufacturing Processes.

	Part Description							Manufacturing process							
Part	C1	C2	C3	C4	C5	C6	C7	Process 1		Process 2		Process 3		Process 4	
								Machine 1	Setting 1	Machine 2	Setting 2	Machine 3	Setting 3	Machine 4	Setting 4
1	1	1		1		1		M1	S_{11}	M2	S_{12}	M4	S_{13}		
2	1	1	1		1		1	M1	S_{21}	M2	S_{22}	M3	S_{23}	M5	S_{24}
3	1				1		1	M1	S_{31}	M3	S_{32}	M5	S_{33}		
4	1	1		1	1	1		M1	S_{41}	M2	S_{42}	M4	S_{43}		
5	1	1	1		1			M1	S_{51}	M2	S_{52}	M3	S_{53}		
6		1			1	1	1	M1	S_{61}	M3	S_{62}	M4	S_{63}	M5	S_{64}
7	1		1		1		1	M1	S_{71}	M2	S_{72}	M3	S_{73}	M5	S_{74}
8		1			1		1	M1	S_{81}	M3	S_{82}	M5	S_{83}		
9	1	1	1	1		1		M1	S_{91}	M2	S_{92}	M3	S_{93}	M4	S_{94}
10	1	1		1		1		M1	S_{101}	M2	S_{102}	M4	S_{103}		
11	1		1	1	1			M1	S_{111}	M2	S_{112}	M3	S_{113}		
12			1		1		1	M2	S_{121}	M3	S_{122}	M5	S_{123}		
13	1	1		1		1	1	M1	S_{131}	M2	S_{132}	M4	S_{133}	M5	S_{134}
14	1		1		1	1		M1	S_{141}	M2	S_{142}	M3	S_{143}	M4	S_{144}
15			1		1	1	1	M2	S_{151}	M3	S_{152}	M4	S_{153}	M5	S_{154}

cording to various criteria, e.g., production cost, setup time, and in-process inventory.

A data-mining approach may link the description of the parts to the manufacturing processes to construct such groups. The goal is to generate relevant processes for manufacturing new parts. The proposed methodology (see Table 54.2 and Figure 54.2) consists of three stages:

Table 54.2. Methodology for the Selection of Manufacturing Processes with Data Mining.

The ***learning stage*** focuses on discovering knowledge from manufacturing processes:
Step 1: Similar parts and processes are grouped into clusters.
Step 2: Relevant processes are associated with each cluster.
The ***exploitation stage*** takes advantage of the clusters to improve the efficiency of generation of process plans for new parts:
Step 3: A new part to be manufactured is matched with a suitable cluster.
Step 4: The new part is assigned the relevant process plan.
The ***specialization stage*** adapts the relevant process for the new part:
Step 5: The relevant process is adapted to the new part.
Step 6: The new process plan data is incorporated into the database.

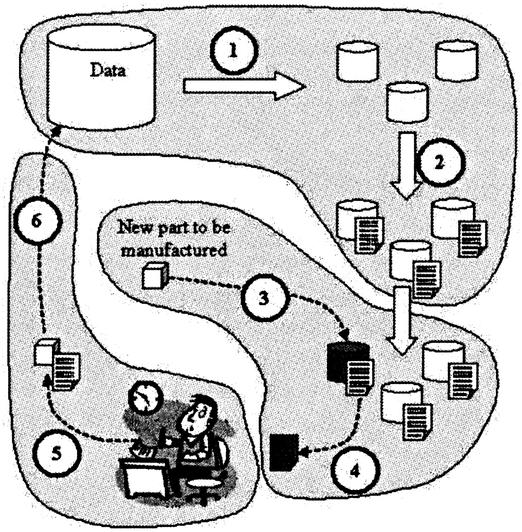

Figure 54.2. Process Selection with a Data Mining Approach.

Step 1 and 2 represent the learning stage. Step 3 and 4 exploit the knowledge extracted. Step 5 adjusts the knowledge provided for a new part and Step 6 provides more data to the database.

The process generates new data that enriches the database.

To illustrate this methodology, consider the data in Table 54.1 representing characteristics of former parts with the corresponding manufacturing processes.

A new part P* with characteristics {C1; C2; C3; C5; C6} is to be manufactured. For the selection of the manufacturing process of the new part P*, the following results are obtained:

3.1 Learning Stage

The initial stage takes advantage of the data collected in Table 54.1 to construct reusable knowledge.

This step differentiates categories of parts. Different number of clusters could be tested to obtain most desirable model. Segmentation algorithms are suitable for this task.

Based on features C1...C7, three clusters are built that distinguish 3 families of parts (see Table 54.3).

Based on the manufacturing processes of the parts contained in each cluster, relevant processes are constructed for each group.

Table 54.3. Clusters Extracted from Table 54.1

Cluster	Parts
Cl_1	{3; 6; 8; 12; 15}
Cl_2	{1; 4; 9; 10; 13}
Cl_3	{2; 5; 7; 11; 14}

Different techniques are available. Decision tree algorithms, decision rules algorithms and neural networks are competent to extract frequent patterns representing the manufacturing processes of each cluster.

From Table 54.1, with clusters described in Table 54.3, the following results in Table 54.4 are obtained.

Table 54.4. Relevant Manufacturing Processes for Clusters.

Cluster	Relevant manufacturing process
Cl_1	{M1; M3; M5} or {M2; M3; M5}
Cl_2	{M1; M2; M4}
Cl_3	{M1; M2; M3}

3.2 Exploitation Stage

The new part to manufacture (P*) is now considered. Knowledge extracted in the previous stage is employed to provide automatically a relevant manufacturing process for P*.

Consider the new part to be manufactured P*. The characteristics of P* are compared to those of the different clusters. Once the closest cluster is identified, P* is affected to that cluster. Classification methods could be applied.

The new part P* considered in this sample has features {C1; C2; C3; C5; C6}. According to its characteristics, this part is classified in cluster Cl_3.

P* belongs to a cluster (Cl_3), that means P* has similar characteristics to the part contained in this cluster. As a result, P* may have a similar manufacturing process for its production.

In actual scenario, P* is affected to the process {M1; M2; M3} which is the pertinent process for Cl_3.

3.3 Specialization Stage

The process affected in the previous step may necessitate some adjustments. An expert from the company adapts it to better fit the specificities of P*.

Suppose the expert obtains the following process {M1; M2; M3; M4}.

The new data completes the database of the company.

The characteristics and manufacturing process of P* are included in the manufacturing database. P* becomes Part 16 and a new entry completes the database like in Table 54.5.

Table 54.5. Updated Database with the New Part Included.

	Part Description							Manufacturing process							
								Process 1		Process 2		Process 3		Process 4	
Part	C1	C2	C3	C4	C5	C6	C7	Machine 1	Setting 1	Machine 2	Setting 2	Machine 3	Setting 3	Machine 4	Setting 4
1	1	1		1		1		M1	S_{11}	M2	S_{12}	M4	S_{13}		
2	1	1	1		1		1	M1	S_{21}	M2	S_{22}	M3	S_{23}	M5	S_{24}
..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
15			1		1	1	1	M2	S_{151}	M3	S_{152}	M4	S_{153}	M5	S_{154}
16	1	1	1		1	1		M1	S_{161}	M2	S_{162}	M3	S_{163}	M4	S_{164}

This new part will enrich the database, and the modifications provided in Step 5 will improve the solution for the next new part.

The more complete is the database, the more exhaustive is the relevant process suggested in Step 4. For example, the database may contain the sequence of machines as well as the manufacturing parameters (speed, temperatures, times, settings) that is the role of the setting parameters S_{ij} that have not been exploited for the demonstration.

4. Conclusion

This chapter illustrates application of Data Mining in manufacturing. The proposed methodology uses the data generated from manufacturing processes to improve efficiency in the manufacturing processes selected for a new part.

Data Mining groups parts in clusters according to their characteristics and manufacturing processes. Standardized processes are generated to represent each cluster.

A standardized process is adapted to the new part and the new data is incorporated into the database to enrich it.

This chapter has illustrated that Data Mining is a useful tool for selection of manufacturing processes. The existing industrial databases should be explored more systematically to extract new knowledge. The latter is key to the improvement of products, processes, and systems.

References

Agard, B., Kusiak, A., Computer Integrated Manufacturing: A Data Mining Approach. in The Engineering Handbook, Second Edition, R.C. Dorf (Ed.), CRC - Taylor and Francis, Boca Raton, FL, pp. 192.1 – 192.11, 2005.

Agard, B. Kusiak, A., Data-Mining Based Methodology for the Design of Product Families, International Journal of Production Research, 42(15): 2955-2969, 2004.

Anand, S.S., Büchner, A.G., Decision Support Using Data Mining. London: Financial Times Pitman Publishers, 1998.

Berry, M.J.A., Linoff, G., Data Mining Techniques: For Marketing, Sales, and Customer Support. New York: Wiley, 2004.

Büchner, A.G., Anand, S.S., Hugues, J.G., Data Mining in Manufacturing Environments: Goals, Techniques and Applications. Studies in Informatics and Control 1997; 6(4):319-328.

Fayyad, U.M., Piatetsky-Shapiro, G., Smyth, P., Uthurusamy, R., Advances in Knowledge Discovery and Data Mining. Cambridge, MA: AAAI Press/The MIT Press, 1996.

Gertosio, C., Dussauchoy, A., Knowledge discovery from industrial databases. Journal of Intelligent Manufacturing 2004; 15:29-37.

Hyer, N., Wemmerlöv, U., Group technology and productivity. Harvard Business Review 1984; July-August:140-149.

Kusiak, A., Decomposition in Data Mining: An industrial case study. IEEE Transactions on Electronics Packaging Manufacturing 2000; 23(4):345-353.

Kusiak, A., Rough set theory: A Data Mining tool for semiconductor manufacturing. IEEE Transactions on Electronics Packaging Manufacturing 2001; 24(1):44-50.

Leu, S.S., Chen, C.N., Chang, S.L., Data Mining for tunnel support stability: neural network approach. Automation in Construction 2001; 10(4):429-441.

Pahl, G., Beitz, W., Engineering Design. Berlin: Springer-Verlag, 1988.

Westphal, C., Blaxton, T., Data Mining Solutions. New York: John Wiley, 1998.