ANNALS of the ORADEA UNIVERSITY.
Fascicle of Management and Technological Engineering, Volume VI (XVI), 2007

FROM DATA TO KNOWLEADGE. AN EXAMPLE OF FUZZY SYSTEM APPLICATION IN MANUFACTURING

Alexandru-Viorel PELE, Sandu Florin BLAGA,
Călin Florin BĂBAN, Mihai GROZA
Oradea University, apele@uoradea.ro

1. Introduction

Progress in digital data acquisition and storage technology has resulted in the growth of huge databases [3]. This has occurred in all areas of human endeavor, such as supermarket transaction data, credit card usage records, telephone call details and government statistics. Also to the unexpected fields such as images of astronomical bodies, molecular databases, and medical records and so on.

The interest has grown in the possibility of tapping these data, of extracting from them information that might be of value to the owner of the database. The discipline concerned with this task has become known as Data Mining.

2. An overview over Data Mining

Data mining (DM) and *Knowledge Discovery in Databases* (KDD) are recently recognized as a distinct area [6] and are often set in the borders context of Statistics and Artificial Intelligence (AI).

2.1. Data Mining and Statistics

Data Mining is often considered to be "a blend of statistics, AI [artificial intelligence], and data base research" [3], which until very recently was not commonly recognized as a field of interest for statisticians, and was even considered by some "a dirty word in Statistics" [3].

An important general difference between Data Mining and the traditional Exploratory Data Analysis (EDA) is that Data Mining is more oriented towards applications and less concerned with identifying the specific relations between the involved variables.

The focus is on producing a solution that can generate useful predictions and not on data preprocessing issues such as data cleaning, data verification, and defining variables. Instead on focus on the basic principles for modeling data and for constructing algorithmic processes to fit these.

The KDD process involves several stages: selecting the target data, preprocessing the data, transforming them if necessary, performing data mining to extract patterns and relationships, and then interpreting and assessing the discovered structures, Figure1.

In practice, the two primary goals of data mining tend to be *prediction* and *description*.

Prediction involves using some variables or fields in the data set to predict unknown or future values of other variables of interest.

Description, on the other hand, focuses on finding patterns describing the data that can be interpreted by humans.

Therefore, it is possible to put data-mining activities into one of two categories:

- Predictive data mining, which *produces the model* of the system described by the given data set, or
- Descriptive data mining, with *produces new, nontrivial information* based on the available data set.

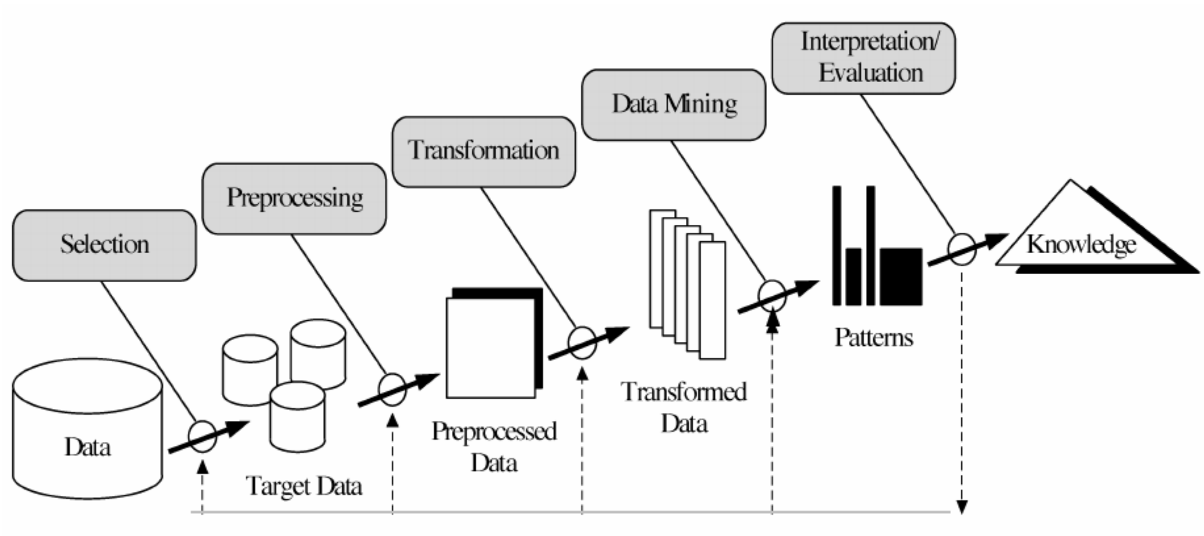

Fig. 1. Knowledge discovery from data [4]

On the predictive end of the spectrum, the goal of data mining is to produce a model, expressed as an executable code, which can be used to perform classification, prediction, estimation, or other similar tasks.

On the descriptive end of the spectrum, the goal is to gain an understanding of the analyzed system by uncovering patterns and relationships in large data sets. The relative importance of prediction and description for particular data-mining applications can vary considerably.

The goals of prediction and description are achieved by using data-mining techniques, explained later in this book, for the following *primary data-mining tasks:*

1. *Classification* – discovery of a predictive learning function that classifies a data item into one of several predefined classes.
2. *Regression* – discovery of a predictive learning function, which maps a data item to a real-value prediction variable.
3. *Clustering* – a common descriptive task in which one seeks to identify a finite set of categories or clusters to describe the data.
4. *Summarization* – an additional descriptive task that involves methods for finding a compact description for a set (or subset) of data.

5. *Dependency Modeling* – finding a local model that describes significant dependencies between variables or between the values of a feature in a data set or in a part of a data set.
6. *Change and Deviation Detection* – discovering the most significant changes in the data set.

2.2. Data Mining and Artificial Intelligence

In the previous paragraphs, a number of different methodologies for the analysis of large data sets have been mentioned. Most of the approaches presented assume that the data is precise and on deals with exact measurements for further analysis.

Historically, as reflected in classical mathematics, we commonly seek a precise and crisp description of things or events by expressing phenomena in numeric or categorical values.

In most of real-world scenarios, on never have totally precise values. There is always going to be a degree of uncertainty or fuzziness. The polarity between fuzziness and precision is quite a striking contradiction in the development of modern information-processing systems.

The knowledge-based information systems include *artificial neural networks*, *evolutionary computing*, *fuzzy logic* and their fusion. Knowledge-based systems are designed to mimic the performance of biological systems.

Artificial neural networks can mimic the biological information processing mechanism in a very limited sense. Evolutionary computing algorithms are used for optimization applications, and fuzzy logic provides a basis for representing uncertain and imprecise knowledge.

Fuzzy logic techniques have been successfully applied in a number of applications: computer vision, decision making, and system design etc.

The most extensive use of fuzzy logic is in the area of control, where examples include controllers for cement kilns, braking systems, elevators, washing machines, hot water heaters, air-conditioners, video cameras, rice cookers, and photocopiers.

3. Example of fuzzy decisional system in manufacturing

Fuzzy decisional systems may be utilized to establish operational priorities in manufacturing systems [1,2,5]. The decisional system has a *generic objective*: to respect delivery terms. The output consists in *priorities array of processing* of eight workpieces that form manufacturing task. The decisional system, named *CE*1a is presented in Figure 2.

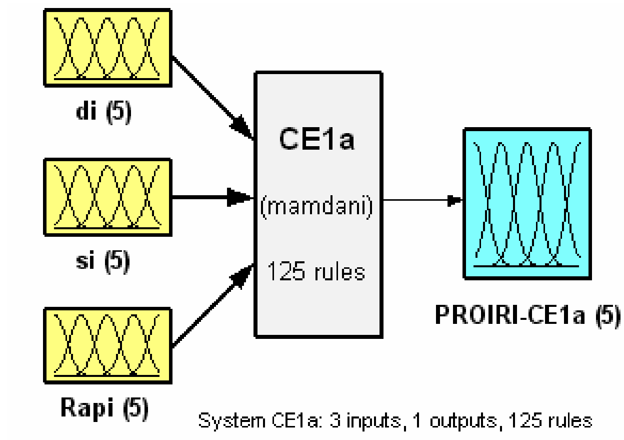

Fig. 2. Decisional fuzzy system CE1a

The system has three input variable:

- *Delivery time* of workpiece *i,* d_i; $i = \overline{1,8}$, with the value domain D_{di} =[1;120] min;
- *Initial time reserve* associated for workpiece i, s_i ; $i = \overline{1,8}$, with the value domain D_{si} = [1;100] min., defined by

$$s_i = d_i - R_{1,i} - TT_i , \tag{1}$$

where d_i – delivery time of workpiece i, $R_{1,i}$ – the moment of arrival in system of workpiece i,; TT_i – total time need to processing workpiece i;

- *Ratio between maximum time to stay in system and the processing time* (*RAP_i* $i = \overline{1,8}$), with the values domain D_{RAPi} = [0;30] min., defined by :

$$RAP_i = \frac{d_i - R_{1,i}}{TT_i} \quad ; \quad i = \overline{1,8} \tag{2}$$

For each input value it is associated a *linguistic variable* with associated linguistic degrees: *fm* – very small; *m*- small; *Md* – medium; *M* – big; *FM* – very big.

The membership functions associated of each linguistic degree, for input values, are Gaussian, Figure 3.

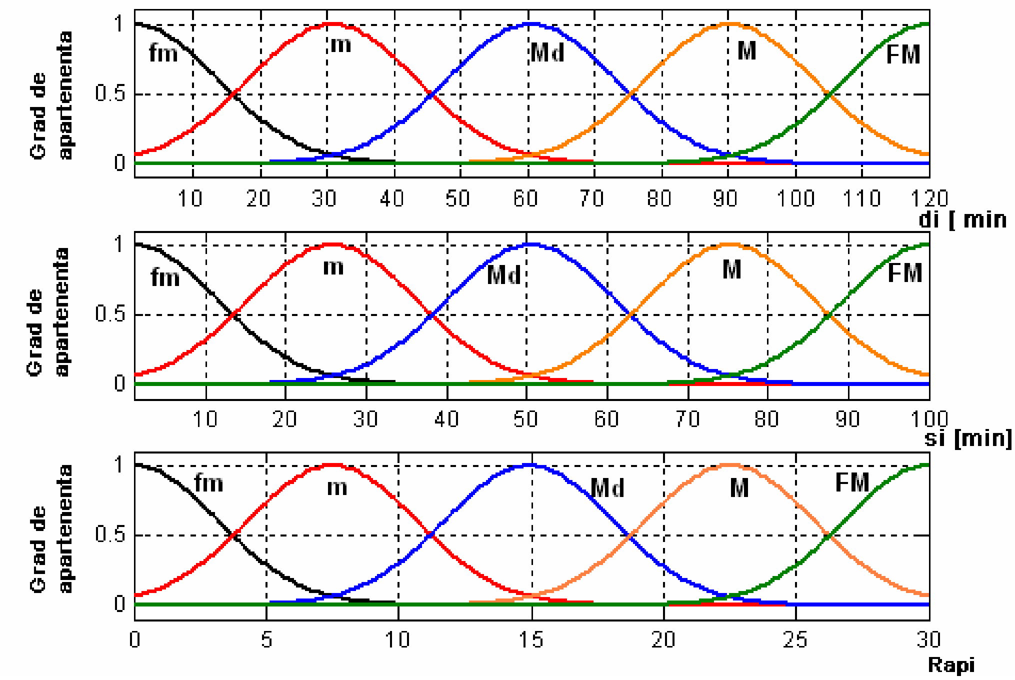

Fig. 3. Membership functions, input variable of CE1a.

The output value from decisional system means *processing priority* of each workpiece of manufacturing task, value determined according with input values.

The output value, named *PRIORI*-CE1a$_i$ has the values domain:

$$D_{PRIORI-CE1a_i} = [0,1]; \quad i = \overline{1,8} \tag{3}$$

and become a Gaussian output linguistic value, with following degrees: *Pfm* – very small; *Pm* – small; *PMd* – medium; *PM* – big; *PFM* – very big, Figure 4.

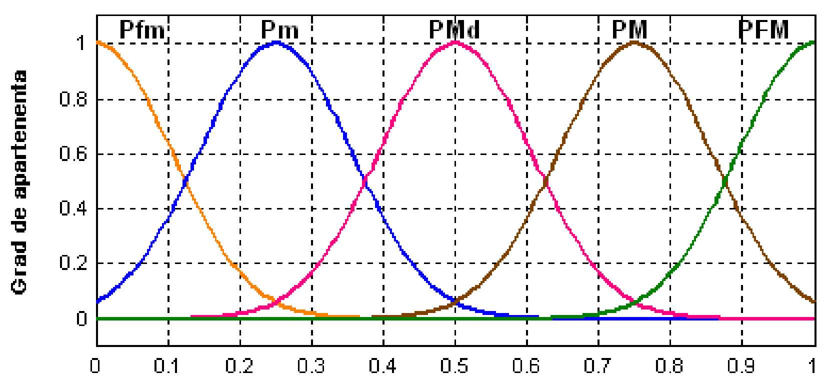

Fig. 3. Gaussian output variable PRIORI-CE1a$_i$

The linguistic variable with their linguistic degrees and membership functions, fuzzy characterize the firmed values of input and output values. The connection between two categories of variable, fuzzy inference, are made by MIN-MAX method, consisting in 125 inference rules, like followings:

$$\left.\begin{array}{l}
1.\text{If (di is fm) and (si is fm) and (Rapi is fm) then (PROIRI - CE1ai is PFM)}\\
\vdots\\
64.\text{If (di is Md) and (si is Md) and (Rapi is M) then (PROIRI - CE1ai is PMd)}\\
\vdots\\
125.\text{If (di is FM) and (si is FM) and (Rapi is FM) then (PROIRI - CE1ai is Pfm)}
\end{array}\right\} \quad (4)$$

For example, considering a manufacturing task composed by eight worpiece, the input parameters are presented in Table 1.

Table 1. Input values

PIESA	R_1 [min]	d [min]	TT [min]
P_1	10	40	5.0
P_2	5	50	9.2
P_3	15	50	8.2
P_4	12	30	4.0
P_5	8	80	4.4
P_6	5	100	7.1
P_7	9	70	8.5
P_8	4	65	7.5

The manufacturing order of worpieces is presented in Table 2. The processing order is : **P_3, P_1, P_2, P_4, P_7, P_8, P_5, P_6**.

Table 2. Processing order

$Piesa_i$	**P_1**	**P_2**	**P_3**	**P_4**	**P_5**	**P_6**	**P_7**	**P_8**
PRIORI-$CE1_i$	0.7727	0.6601	0.7951	0.6063	0.3364	0.3097	0.4956	0.4775

4.Conclusions

Data mining has recently attracted a lot of attention in different fields and has a great potential for valuable commercial and scientific discoveries. Many businesses and scientific communities are currently employing data-mining technology [3].

There are many collections of real-life examples of data-mining implementations from the business, marketing, industry and scientific world.

In this paper, a fuzzy decisional system is used to establish operational priorities in manufacturing systems.

5.Bibliography

1. Blaga Florin, Tarcă, R., Mihăilă, I., Ţarcă, I., *About the manufacturing scheduling in FMC-2R-2002 flexible manufacturing cell*, ROBOTEP 2004 7th International Conference on Automation/Robotics in theory and practice Kosice, ISBN 80-8073-134-9, Slovak Republic, pp 160-167, pg. 8, 2004
2. Blaga Florin, *The use of generalized Petri Nets in Automatic Packing Line Modeling*, International Conferences TMCR 2005, ISBN 9975-9748-0-5, ISBN 9975-9748-4-8, Chisinau, Republica Moldova, pp. 55-54, pg. 4, 2005.
3. David Hand, Heikki Mannila and Padhraic Smyth, *Principles of Data, MIT Press*, 2001.
4. Fayyad, U.M., G. Piatetsky-Shapiro, P. Smyth: *From Data Mining to Knowledge Discovery: An Overview.* In : „Fayyad, U.M. et al: Advances in Knowledge Discovery and Data Mining. AAAI Press/The MIT Press. Menlo Park, California 1996“, pp. 1-36
5. Florin Blaga, A.V., Pele, *Decision Procedure Based on Fuzzy Sets Integration in a Specifically Flexible Manufacturing Systems Scheduling Function*, Proceedings of International Scientific Conference SI 2005 Mechanical Engineering, Bratislava, Slovak Republic, 2005.
6. Irina Neaga, Oana Dodun, *The Extended Manufacturing Enterprise. An Approach based on Knowledge and WEB Technology*, Politehnium, Iaşi, 2005.
7. Paolo Giudici, *Applied Data Mining Statistical Methods for Business and Industry*, John Wiley & Sons Ltd, England, 2003.