Data Mining Applications in Steel Industry

Joaquín Ordieres-Meré
University of La Rioja, Spain

Manuel Castejón-Limas
University of León, Spain

Ana González-Marcos
University of León, Spain

INTRODUCTION

The industrial plants, beyond subsisting, pursue to be leaders in increasingly competitive and dynamic markets. In this environment, quality management and technological innovation is less a choice than a must. Quality principles, such as those comprised in ISO 9000 standards, recommend companies to make their decisions with a based on facts approach; a policy much easily followed thanks to the all-pervasive introduction of computers and databases.

With a view to improving the quality of their products, factory owners are becoming more and more interested in exploiting the benefits gained from better understanding their productive processes. Modern industries routinely measure the key variables that describe their productive processes while these are in operation, storing this raw information in databases for later analysis. Unfortunately, the distillation of useful information might prove problematic as the amount of stored data increases. Eventually, the use of specific tools capable of handling massive data sets becomes mandatory.

These tools come from what it is known as 'data mining', a discipline that plays a remarkable role at processing and analyzing massive databases such as those found in the industry. One of the most interesting applications of data mining in the industrial field is system modeling. The fact that most frequently the relationships amongst process variables are nonlinear and the consequent difficulty to obtain explicit models to describe their behavior leads to data-based modeling as an improvement over simplified linear models.

BACKGROUND

The iron and steel making sector, in spite of being a traditional and mature activity, strives to approach new manufacturing technologies and to improve the quality of its products. The analysis of process records, by means of efficient tools and methods, provides deeper knowledge about the manufacturing processes, therefore allowing the development of strategies to cut costs down, to improve the quality of the product, and to increase the production capability.

On account of their anticorrosive properties, the galvanized steel is a product experiencing an increasing demand in multiple sectors, ranging from the domestic appliances manufacturing to the construction or automotive industry. Steel making companies have established a continuous improvement strategy at each of the stages of the galvanizing process in order to lead the market as well as to satisfy the, every time greater, customer requirements (Kim, Cheol-Moon, Sam-Kang, Han, C. & Soo-Chang, 1998; Tian, Hou & Gao, 2000; Ordieres-Meré, González-Marcos, González & Lobato-Rubio, 2004; Martínez-de-Pisón, Alba, Castejón & González, 2006; Pernía-Espinoza, Castejón-Limas, González-Marcos & Lobato-Rubio, 2005).

The quality of the galvanized product can be mainly related to two fundamental aspects (Lu & Markward, 1997; Schiefer, Jörgl, Rubenzucker & Aberl, 1999; Tenner, Linkens, Morris & Bailey, 2001):

- As to the anticorrosive characteristics, the quality is determined by the thickness and uniformity of the zinc coat. These factors basically depend on

the base metal surface treatment, the temperature of its coating and homogenization, the bath composition, the air blades control and the speed of the band.
- As to the steel properties, they mainly depend on the steel composition and on the melting, rolling and heat treatment processes prior to the immersion of the band into the liquid zinc bath.

MAIN FOCUS

Data mining is, essentially, a process lead by a problem: the answer to a question or the solution to a problem is found through the analysis of the available records. In order to facilitate the practice of data mining, several standardization methods have been proposed. These divide the data mining activities in a certain number of sequential phases (Fayyad, Piatetsky-Shapiro, & Smyth, 1996; Pyle, 1999; Chapman, Clinton, Kerber, Khabaza, Reinartz, Shearer, & Wirth, 2000). Nevertheless, the names and contents of these phases slightly differ. The same general concept is present in every proposal: first the practitioner becomes familiar with the problem and the data set; later on, the data is adapted to suit the needs of the algorithms; then, models are built and tested; finally, newly acquired knowledge provides support for an enhanced decision making process.

Nowadays, nonlinear modeling comprises important techniques that have reached broad applicability thanks to the increasing speed and computing capability of computers. Genetic algorithms (Goldberg & Sastry, 2007) , fuzzy logic (Wang, Ruan & Kerre, 2007) and cluster analysis (Abonyi & Feil, 2007), to name a few, could be mentioned amongst these techniques. Nevertheless, if there is a simple and efficient technique that has made inroads into the industry, that definitely is the artificial neural network.

Amongst the currently existing neural networks, one of the most adequate for defining systems and processes by means of their inputs and outputs is the multilayer perceptron (MLP), which can be considered a universal function approximator (Funahashi, 1989; Hornik, Stinchcombe. & White, 1989); a MLP network with one hidden layer only needs a large enough number of nonlinear units in its hidden layer to mimic any type of function or continuous relationship amongst a group of input and output variables.

The following lines show the successful application of the MLP network, along with other data mining tools, in the modeling of a hot dip galvanizing line whose existing control systems were improved.

Estimation of The Mechanical Properties of Galvanized Steel Coils

The core of the control engineering strategies relies on the capability of sensing the process variables at a rate high enough so as to be able to act in response to the observed behavior. ‘Closed loop’ control strategies, as they are commonly referred to, are based on acting over the commanding actions to correct the differences amongst the observed and expected behavior. In that scenario, control engineers can choose amongst many effective and well known approaches.

Unfortunately, nowadays the mechanical properties of the galvanized steel coils (yield strength, tensile strength and elongation) cannot be directly measured. Instead, they have to be obtained through slow laboratory tests, using destructive methods after the galvanizing process. The lack of online (fast enough) measures compels the control engineers to adopt an ‘open loop’ control strategy: the estimation of the best commanding actions in order to provide the desired outputs.

As it has already been mentioned, the mechanical properties of the galvanized steel coils change during the manufacturing processes: from the steel production processes that determine their chemical composition, to the very last moment in which it becomes a finished product, either in the shape of galvanized coils or flat plates.

Ordieres-Meré, González-Marcos, González & Lobato-Rubio (2004) proposed a neural network model based on the records of the factory processes in operation. The model was able to predict the mechanical properties of the galvanized steel not only with great accuracy, which is mandatory to improve the quality of the products, but also fast enough, what is critical to close the loop and further improve the quality thanks to a wider range of more complex strategies.

This particular data-based approach to system modeling was developed to satisfy the needs of ARCELOR SPAIN. Following data mining practice standards, the 1731 samples of the data set were analyzed and prepared by means of visualization techniques (histograms, scatter plots, etc.) and projection techniques (Sammon projection (Sammon, 1969) and principal components

Related Content