Business effiScience
Science Driven Strategic Advisory

Subject : Introduction to Business EffiScience Algorithms

Introduction

For a long time behaviour or process analysis and description approaches have had recourse to more or less sophisticated tools enabling either considerable populations to be described synthetically, or making it possible to attempt to qualify or understand specific behaviours. New approaches and tools now enable the technical difficulties previously encountered to be overcome, so that quantitative and qualitative analyses enabling the “why” of phenomena to be understood can be carried out.

History

The best and most widely used tool is clearly statistics, either in its basic information synthesis forms (mean, minimum, maximum, standard deviation, linear regression), or in its more mathematically evolved forms (multivariate analysis, factorial analysis, logistical regression).

More sophisticated tools appeared in the 60s and began to be used extensively through industry from the 80s thanks to the increasing calculation power of computers. Two families can be distinguished: learning systems and descriptive or visual systems. Visual systems are based on a multidimensional spatial representation of data. This graphic database analysis software enables assisted navigation in space. It is passive and represents what users request so as to enable them to focus on certain regions or eliminate areas of space. The best-known learning systems are neural networks, which reconstitute an association process through learning in terms of their architecture and enable it to be reproduced. The decision trees are less refined but easier to parameterise. Finally, recent developments in vector support machines are leading to high-quality predictions.

Issue

These tools essentially make it possible to analyse, describe and predict. Yet the aim of many studies or analyses is increasingly to endeavour to understand the “why”, to understand the nature of the more or less complex interactions and links that relate parameters, descriptors or variables to results or behaviours, to simulate new interactions or new scenarios and optimise them in terms of the initial data.

Business effiScience algorithms contribute a focused approach and focused tools to the **search for “why”** and the formalisation of configurations in the shape of **formal rules.** They make it possible to carry out **explicative analysis of phenomena**.

Technology

Recent developments in discrete mathematics, genetic algorithmics and induction (inductive?) transitivity are behind the development of Business effiScience algorithms. **Discrete mathematics** is the basis of scientific IT and is used here more particularly in the form of logical sequences on multidimensional matrices. **Genetic algorithmics** comprises all the automatic optimisation techniques of an algorithm, by imitation of genetic crossing-over processes. The software that uses this technique is very flexible and multifunctional due to its ability to adapt and improve automatically with the context. **Induction transitivity**, for its part, is currently used in the interpretation of language by machines and involves the interpretation of partial data (or how to perform a synthesis from partial or local information).

Around the principle of convex subspaces parallel to the axes, the ascending learning algorithms resulting from the bringing together of these three technologies have been the subject of several years' study and have been particularly difficult to develop. The first works date from the 90s at the Ecole Polytechnique in the Applied Mathematics laboratory.

From real digital or textual input data describing a process whose result is expressed in families or classes, Business effiScience algorithms learn to recognise complex interactions between some descriptors that explain the differentiation between one output class and another. They run through the classes and extract information from them regarding local consistencies detected within the group of entities of the same nature and then incorporate them using this local or partial information to access the most global level of knowledge possible (as long as there is no information loss).

Results

Business effiScience algorithms thus succeed in **characterising entities of different classes** (for example satisfactory classes or risk classes) through a common signature within each class. This signature is formalised in the shape of a rule that accurately quantifies the values taken by some data that make it possible to identify or diagnose the relevant class.

In addition, once the rules have been generated, they can be modified or completed in terms of constraints relating to reliability, opportunity or the ability of specialists to see and understand their impacts on a result straight away.

Assets

The 6 major assets that lead Business effiScience algorithms to stand out from classical analysis techniques are:

1. Possibility of taking digital and non-digital descriptors into account (description by attributes),
2. Resistance to noise and missing data,
3. Resistance to irrelevant descriptors,
4. Identification of local behaviours or particular cases,
5. Exhaustiveness of analysis (100% of the entities are “explained”),
6. Formalisation of results in the form of directly explicit formal rules,
7. Simulations and scenario games applicable from the rules learnt or ex-nihilo.

Thus, a **structured application** of these algorithms will make it possible to **cross a major qualitative and quantitative threshold in issues involving the comprehension and simulation of complex phenomena.**