The Inductive Database framework: a long-term perspective on query languages for data mining.

Jean-François Boulicaut

Institut National des Sciences Appliquées de Lyon
LISI Bâtiment Blaise Pascal
F-69621 Villeurbanne cedex, France
Jean-François.Boulicaut@insa-lyon.fr

We consider the challenge of supporting KDD processes by means of queries. We introduce a research perspective, the so-called inductive database framework [4, 1, 3, 2]. An inductive database is a database that in addition to data contains intentionally defined generalizations about the data. Specifying and implementing inductive databases is a long-term perspective. However, more or less specialized data mining query languages have been proposed, e.g., MSQIL [7] or MINE RULE [5]. It is interesting to discuss whether they are good candidates for inductive database querying. Furthermore, the optimization of inductive queries is clearly related to constraint-based mining (see e.g., [6]). Considering the association rule mining task, we can figure out what should be done in a near future in order to implement inductive databases for that simple but useful kind of pattern. However, applying the framework and studying the robustness of the identified concepts for other data mining tasks remain open. Considering these issues is the main goal of the European project chnQ IST-2000-26469 (May 2001-May 2004), funded by the Future and Emerging Technologies arm of the IST Programme.

References