

# Le rôle de l'utilisateur dans un processus d'extraction de règles d'association

Cyril Nortet\*, Ansaf Salleb\*\*  
Teddy Turmeaux\*, Christel Vrain\*

\* LIFO Rue Léonard de Vinci BP 6759 45067 Orléans Cedex 02  
{Cyril.Nortet, Teddy.Turmeaux, Christel.Vrain}@lifo.univ-orleans.fr

\*\* INRIA Projet Dream, Campus Universitaire de Beaulieu, 35042 Rennes  
Ansaf.Salleb@irisa.fr

**Résumé.** De nombreux travaux ont porté sur l'extraction de règles d'association. Cependant, cette tâche continue à intéresser les chercheurs en fouille de données car elle soulève encore plusieurs défis. En particulier, son utilisation en pratique reste difficile : d'une part, le nombre de règles apprises est souvent très grand, d'autre part, le traitement des valeurs numériques dans cette tâche est loin d'être maîtrisé. Nous nous intéressons dans cet article au rôle que peut jouer l'utilisateur pour pallier ces difficultés. Il s'agit d'impliquer l'utilisateur dans le processus de recherche de règles d'association qui est dans ce cas interactif et guidé par des schémas de règles qu'il aurait choisis. Nous illustrons notre propos avec QuantMiner qui est un outil convivial et interactif que nous avons développé. La présence de l'expert reste indispensable durant tout le processus d'extraction de règles.

**Mots clé.** Fouille de Données Interactive, Règle d'Association Quantitative, Optimisation.

## 1 Introduction

L'extraction de règles d'association est devenue aujourd'hui une tâche populaire en fouille de données. Elle a pour but de dégager des relations intelligibles entre des attributs dans une base de données. Une règle d'association (Agrawal et al. 1993) est une implication  $G \Rightarrow C_2$ , ou  $C_1$  et  $C_2$  expriment des conditions sur les attributs de la base de données. La qualité d'une règle est classiquement évaluée par un couple de mesures support et confiance, définies par :

- $\text{Support}(C)$ , ou  $C$  exprime des conditions sur les attributs, est le nombre de n-uplets (lignes de la base de données) qui satisfont  $C$ .
- $\text{Support}(C_1 \Rightarrow C_2) = \text{Support}(C_1 \wedge C_2)$
- $\text{Confiance}(C_1 \Rightarrow C_2) = \text{Support}(C_1 \wedge C_2) / \text{Support}(C_1)$

Une règle d'association est dite solide, si son support et sa confiance dépassent deux seuils fixés *a priori*, MinSupp et MinConf respectivement. De nombreux travaux se sont intéressés au problème crucial de performance que pose cette tâche (par ex. (Brin et al. 1997,