

Mesurer l'intérêt des règles d'association

Benoît Vaillant*, Patrick Meyer***, Elie Prudhomme**,
Stéphane Lallich**, Philippe Lenca*, Sébastien Bigaret*

*GET ENST Bretagne / Département LUSSI – CNRS UMR 2872
Technopôle de Brest Iroise - CS 83818, 29238 Brest Cedex, France
{*prenom.nom*}@enst-bretagne.fr

**Laboratoire ERIC - Université Lumière - Lyon 2
5 avenue Pierre Mendès-France, 69676 Bron Cedex, France
lallich@univ-lyon2.fr

***Service de Mathématiques Appliquées, Faculté de Droit,
d'Economie et de Finance, Université du Luxembourg,
162a, avenue de la Faïencerie, L-1511 Luxembourg
patrick.meyer@uni.lu

Résumé. Cet article présente nos travaux sur la mesure de l'intérêt des règles d'association. Une vingtaine de mesures ont été retenues, sur la base d'un critère d'éligibilité. Différentes propriétés sont d'abord proposées qui fondent une étude formelle des mesures. Cette étude formelle se double d'une étude de comportement, grâce à HERBS, une plate-forme développée pour expérimenter les mesures sur des bases de règles. Il est alors possible de confronter la typologie formelle des mesures et la typologie expérimentale. Une fois transformées en critères, ces propriétés fondent une méthode d'assistance au choix de l'utilisateur. Le problème de la validation est enfin abordé, où l'on présente une méthode de contrôle du risque multiple adaptée au problème.

1 Introduction

Nous nous intéressons aux mesures relatives à l'intérêt des règles d'association $A \rightarrow B$ telles que définies dans (Agrawal *et al.*, 1993) : dans une base de données transactionnelles, $A \rightarrow B$ signifie que si les articles qui constituent A sont dans *le panier d'une ménagère*, alors le plus souvent les articles qui constituent B le sont aussi. Les algorithmes de type APRIORI (fondé sur le support et la confiance) ont tendance à produire un grand nombre de règles pas toujours intéressantes du point de vue de l'utilisateur. Les mesures d'intérêt jouent alors un rôle essentiel en permettant de pré-filtrer les règles extraites.

Cet article présente une synthèse de l'ensemble de nos travaux sur la qualité des règles d'association. C'est un résumé de (Vaillant *et al.*, 2005).

Différentes voies ont été explorées. Ainsi, nous définissons des mesures et proposons des propriétés souhaitables section 2. La section 3 concerne le développement de la plate-forme expérimentale HERBS. La section 4 est relative au développement d'une aide à la sélection de bonnes mesures. Les deux typologies des mesures, l'une fondée sur une approche expérimentale, l'autre sur une approche formelle sont mises en regard section 5. Enfin, la section 6 s'intéresse à la validation des règles.