

Evaluation et analyse multicritère des mesures de qualité des règles d'association

Philippe Lenca*, Patrick Meyer**, Benoît Vaillant*,
Philippe Picouet*, Stéphane Lallich***

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

<http://www-iasc.enst-bretagne.fr/~nom/>

**Département de Mathématique - Université de Liège
Bât. B37 Grande Traverse, 12, B-4000 Liège, Belgique
patrick.meyer@internet.lu

***Laboratoire ERIC - Université Lumière - Lyon 2
5 avenue Pierre Mendès-France, 69676 Bron Cedex, France
lallich@univ-lyon2.fr,
<http://eric.univ-lyon2.fr/~lallich/>

Résumé. Le nombre particulièrement important de règles générées par les algorithmes utilisés en extraction de connaissances à partir de données (ECD) ne permet pas aux utilisateurs de faire eux-mêmes la sélection des règles pertinentes. Un des problèmes centraux de l'ECD est le développement de mesures évaluant l'intérêt des règles découvertes. Ainsi de nombreuses mesures ont été proposées, parmi lesquelles l'expert est censé choisir celle qui est la plus appropriée à ses besoins. Mais le nombre élevé de mesures est lui-même un frein à la capacité de choix d'un expert. Pour y remédier, de nombreuses publications dressent des panoramas partiels de ces mesures, s'attachant tantôt à leur qualité algorithmique, tantôt à la formalisation de leurs propriétés, etc. Le résultat n'est malheureusement pas à la hauteur des espérances. A la multiplicité des mesures s'ajoute maintenant la diversité des caractéristiques, lesquelles ne reflètent généralement pas les objectifs de l'expert. Enfin, l'hétérogénéité des valeurs prises par ces caractéristiques n'est pas la moindre des difficultés que rencontre l'utilisateur quand il entreprend de comparer les mesures. Dans ce contexte, et malgré les efforts entrepris, le choix, par un utilisateur métier, de mesures répondant à son objectif reste un défi majeur de la recherche en ECD. La recherche des meilleures règles parmi le vaste ensemble de règles produit, passe aussi par la recherche et l'utilisation des bonnes mesures. On se trouve donc dans une problématique d'aide multicritère à la décision (AMD). Nous fondons notre approche sur l'évaluation de 20 mesures à partir de 8 critères.

Mots-clés : mesures d'intérêt, règles d'association, aide à la décision.