

CHIC : traitement de données avec l'analyse implicative

Raphaël Couturier*, Régis Gras**

*LIFC - IUT Belfort

raphael.couturier@iut-bm.univ-fcomte.fr

**École polytechnique de l'université de Nantes

regisgra@club-internet.fr

Résumé. Cet article a pour but de montrer les possibilités offertes par le logiciel CHIC (Classification Hiérarchique Implicative et Cohésitive) pour effectuer certaines analyses de données. Il est basé sur la théorie de l'Analyse Statistique Implicative ou A.S.I. développée par Régis Gras et ses collaborateurs. Le principe premier de l'A.S.I. repose sur la problématique d'une mesure des règles d'association du type : «si a alors b » dans une population instanciant les variables a et b . CHIC enrichit sa réponse, établie sur des bases statistiques, en évaluant la responsabilité des sujets dans l'élection de la règle. L'article présent explique la démarche à suivre pour utiliser le logiciel ainsi que les possibilités offertes par celui-ci.

1 Introduction

CHIC est le fruit informatique des travaux sur l'analyse statistique implicative. La version actuelle de ce logiciel a été portée en C++ sous Windows il y a 10 ans environ à partir d'une version antérieure en Pascal, mais avec des développements importants et avec une plus grande convivialité Couturier (2000). Depuis elle a subi régulièrement de nombreuses modifications tant au niveau pratique que sur le plan théorique en intégrant de nombreux nouveaux modes de calculs et de nouveaux concepts. Dans cet article nous faisons le point sur les possibilités actuelles de CHIC pour l'usager des méthodes d'analyse de données. La théorie de l'analyse implicative est présentée dans Gras et al. (1996, 2001b, 2004). La documentation de CHIC peut apporter d'autres éclaircissements sur les concepts introduits dans cette présentation de CHIC. Rappelons-en brièvement les grandes lignes.

Partie d'une question de didactique (organiser les comportements de réponse d'élèves à des exercices de mathématiques et en valider une taxonomie de complexité), la problématique de l'A.S.I. s'est développée au fil des problèmes rencontrés dans d'autres situations où la recherche de règles –relations non symétriques, i.e. directionnelles– se pose et est résumée par l'énoncé : «si l'on observe a dans la population alors on observe généralement b ». La fonction première de la règle est donc d'être un prédicteur potentiel. L'extension de la problématique initiale s'est développée dans plusieurs directions :

- la nature des variables binaire, puis numérique, ordinale, intervalle Gras et al. (2001a), floue ;
- la mesure affectée à la règle : classique sur une base adaptée de l'algorithme de la vraisemblance du lien de Lerman (1981), puis entropique afin d'atténuer l'effet