

Un critère Bayésien pour évaluer la robustesse des règles de classification

Dominique Gay*, Marc Boullé*

*Orange Labs
2, avenue Pierre Marzin
F-22307, Lannion Cédex
{prenom.nom}@orange-ftgroup.com

Résumé. L'utilisation de règles de classification dans les modèles prédictifs a été très étudiée ces dernières années. La forme simple et interprétable des règles en font des motifs très populaires. Les classifieurs combinant des règles de classification intéressantes (selon une mesure d'intérêt) offrent de bonnes performances de prédictions. Cependant, les performances de ces classifieurs dépendent de la mesure d'intérêt (e.g., confiance, taux d'accroissement, ...) et du seuillage (non-trivial) de cette mesure pour déterminer les règles pertinentes. De plus, il est facile de montrer que les règles extraites ne sont pas individuellement robustes. Dans cet article, nous proposons un nouveau critère pour évaluer la robustesse des règles de classification dans les données Booléennes. Notre critère est issu d'une approche Bayésienne : nous proposons une expression analytique de la probabilité d'une règle connaissant les données. Ainsi, les règles les plus probables sont robustes. Le critère Bayésien nous permet alors d'identifier (sans paramètre) les règles robustes parmi un ensemble de règles données.

1 Introduction

Les règles d'association (Agrawal et al., 1993) font certainement partie des motifs les plus étudiés en fouille de données. Dans les données binaires, une règle d'association est une expression de la forme $\pi : X \rightarrow Y$, où X (le corps) et Y (le conséquent) sont des sous-ensembles d'attributs Booléens. Intuitivement, la sémantique de π est : "*lorsqu'on a observé les attributs de X , alors souvent on a observé aussi les attributs de Y* ". Le principal intérêt d'une règle est son pouvoir d'inférence inductive. En effet, si maintenant on observe les attributs de X alors on va probablement aussi observer les attributs de Y . Lorsque Y est un attribut classe, on parle alors de règle de classification ($X \rightarrow c$). Les règles de classification semblent propices aux tâches de prédiction ; puisque si un objet est décrit par les attributs de X alors il est probablement de classe c . Les récentes avancées en extraction de motifs ont donné naissance à de nombreux classifieurs à base de règles (e.g. Liu et al. (1998) pour les pionniers ou Bringmann et al. (2009) pour une vue d'ensemble). Ces méthodes sont connues pour leur interprétabilité et sont performantes en terme de prédiction dans les tâches de classification supervisée. Toutefois, on peut identifier au moins deux verrous :