

Arbres de décision crédibilistes

Patrick VANNOORENBERGHE
Université Paul Sabatier Toulouse3
Laboratoire de Télédétection à Haute Résolution
Centre de Télédétection Spatiale

Une grande majorité des travaux sur la reconnaissance des formes et l'apprentissage se sont focalisés sur l'induction d'arbres de décision à partir d'un ensemble d'apprentissage constitué d'exemples dont l'étiquetage est connu avec précision. Dans certaines applications cependant, on est loin d'avoir une information aussi parfaite sur l'ensemble d'apprentissage et notamment sur l'étiquetage des exemples. C'est le cas notamment d'un groupe d'experts dont les opinions divergent concernant l'appartenance d'un exemple à une classe précise ou lorsque l'étiquetage devient fastidieux dû à un trop grand nombre d'exemples. Pour résoudre ce problème, une méthode d'induction d'arbre de décision basée sur la théorie des fonctions de croyance est présentée. La mesure d'entropie classiquement utilisée pour quantifier l'impureté de chaque nœud dans l'arbre est remplacée par une mesure d'incertitude (dérivée de la fonction de croyance en chaque nœud). Celle-ci permet de prendre en compte non seulement la proportion des classes mais aussi le nombre d'exemples en chaque nœud. Par ce biais, elle permet d'éviter la mise en oeuvre d'une procédure d'élagage de l'arbre. L'algorithme qui en résulte permet de traiter les ensembles de données dont l'appartenance aux classes n'est connue que partiellement. Quelques essais sur des données simulées ou réelles seront présentés.