

Apprendre les relations de préférence et de co-occurrence entre les labels en classification multi-labels

Khalil Laghmari^{*,**}
Christophe Marsala^{**}
Mohammed Ramdani^{*}

^{*}Laboratoire Informatique de Mohammedia,
FSTM, Hassan II University of Casablanca,
BP 146 Mohammedia 20650 Maroc.
laghmari.khalil@gmail.com
ramdani@fstm.ac.ma
^{**}Sorbonne Universités,
UPMC Univ Paris 06,
CNRS, LIP6 UMR 7606,
4 place Jussieu 75005 Paris, France.
christophe.marsala@lip6.fr

Résumé. En classification multi-labels, chaque instance est associée à un ou plusieurs labels. Par exemple, un morceau de musique peut être associé aux labels 'heureux' et 'relaxant'. Des relations de co-occurrence peuvent exister entre les labels : par exemple, les labels 'heureux' et 'triste' ne peuvent pas être associés au même morceau de musique. Les labels peuvent aussi avoir des relations de préférence : par exemple, pour un morceau de musique contenant plusieurs piques, le label 'heureux' est préféré par rapport au label 'relaxant'. Les relations entre les labels peuvent aider à mieux prédire les labels associés aux instances. Les approches existantes peuvent apprendre soit les relations de co-occurrence, soit les relations de préférence. Ce travail introduit une approche permettant de combiner l'apprentissage des deux types de relations. Les expérimentations menées montrent que la nouvelle approche introduite offre les meilleurs résultats de prédiction par rapports à cinq approches de l'état de l'art.

1 Introduction

La reconnaissance de formes (pattern recognition) constitue une compétence d'intelligence fondamentale. Par exemple, en considérant un ensemble de labels disponibles : {'heureux', 'relaxant', 'triste', 'nerveux'}, l'intelligence humaine est capable d'associer ces labels à des morceaux de musique selon les émotions qu'ils expriment.

Le défi de transférer la compétence de reconnaissance de formes aux ordinateurs fait partie du domaine de l'apprentissage artificiel. L'apprentissage est dit supervisé