

MIXMOD : un logiciel de classification supervisée et non supervisée pour données quantitatives et qualitatives

Florent Langrogné

Laboratoire de Mathématiques de Besançon
UMR6623, CNRS & Université de Franche-Comté
16, route de Gray, 25030 Besançon, France

Résumé MIXMOD, logiciel de classification supervisée et non supervisée, traite des données multidimensionnelles en utilisant la richesse des modèles de mélanges. Il propose de nombreuses fonctionnalités (large palette de modèles gaussiens et multinomiaux, d’algorithmes, de critères de sélection, ...) et repose sur une architecture particulièrement adaptée pour atteindre un haut niveau de performance (temps de calcul et robustesse). MIXMOD est téléchargé environ 250 fois par mois, et il est utilisé dans des situations très diverses à la fois dans le milieu de la recherche et, pour une part croissante, par des utilisateurs novices. Des évolutions sont régulièrement proposées pour répondre aux demandes des utilisateurs et intégrer les résultats de la recherche.

Keywords : classification des données, logiciel, modèles de mélanges.

1 Introduction

Par leur flexibilité, les modèles finis de distributions de probabilité sont très utiles pour modéliser une grande variété de phénomènes aléatoires et sont naturellement considérés comme un outil de choix pour traiter des problématiques de classification supervisée et non supervisée. Utilisant ce cadre de travail, le logiciel MIXMOD est un logiciel adapté à de nombreuses situations, y compris dans des situations complexes. Il peut être utilisé dans les environnements Scilab et Matlab et est disponible, sous licence GNU GPL, pour les systèmes d’exploitation Linux et Windows.

Ce article ne vise pas à remplacer le guide de l’utilisateur ni la documentation statistique que l’on peut trouver sur le site web dédié à MIXMOD. Il a pour but de présenter ce logiciel (son historique, son architecture, les différentes façons de l’utiliser) et de montrer, sur des exemples concrets, ses principales fonctionnalités et l’intérêt d’utiliser un tel logiciel dans des situations diverses (données quantitatives ou qualitatives par exemple).

2 Fiche d’identité

2.1 Historique

Le projet MIXMOD est né en 2001 avec pour objectif de développer un outil efficace, rapide et robuste pour traiter des problématiques de classification de données en utilisant les modèles de mélange (voir [11]). Sous l’impulsion des quatre auteurs principaux