
VAHID PARTOVI NIA, McGill University
Classification double agglomérative bayésienne

Dans plusieurs domaines biologiques tels que la métabolomique, la protéomique et la génétique, une matrice de données dont les sujets correspondent aux lignes et les variables aux colonnes, est produite. Pour ce type de données, la classification simultanée des sujets et des variables, appelée classification double, est intéressante. Par exemple, avec des données d'expression génétique, la classification des sujets groupe les sujets ayant des similarités génétiques, tandis que la classification des gènes groupe les gènes ayant des fonctions similaires. Nous suggérons un algorithme de classification double complètement automatisé en utilisant un modèle bayésien dont il n'est pas nécessaire de connaître le nombre de partitions. De plus, la méthode agglomérative produit une représentation graphique des classes doubles à l'aide d'un dendrogramme.