

Détermination du nombre optimal de classes présentant un fort degré de chevauchement

Ammor.O*- Raiss.N- Slaoui.K*****

**Laboratoire de modélisation et calcul scientifique. Faculté des sciences et techniques Fès Université Sidi Mohammed Ben Abdellah*

***Laboratoire ISQ. Faculté des sciences Dhar Mehraz, Université Sidi Mohammed Ben Abdellah , Fès, Maroc*

****Laboratoire LESSI . Faculté des sciences Dhar Mehraz, Université Sidi Mohammed Ben Abdellah , Fès, Maroc*

Adresse de correspondance : w_ammor@yahoo.fr

RÉSUMÉ : Dans cet article, nous présentons un nouvel indice pour la détermination du nombre optimal et correct de classes nommé V_{MEP} basé sur le Principe du Maximum d'Entropie. Les performances de ce nouvel indice déduit d'une combinaison originale entre des méthodes d'analyse des données et le critère du maximum d'entropie, sont montrées à travers un ensemble d'exemples simulés et réels. La procédure est complètement automatique dans le sens qu'elle ne nécessite aucun paramètre de réglage. V_{MEP} montre une grande robustesse, et une supériorité par rapport à d'autres indices déjà existants et assez récents, particulièrement dans le cas du chevauchement spatial entre classes.

ABSTRACT: In this paper, we propose a new and efficient clusters validity measure named V_{MEP} for determination of the optimal and correct number of clusters based on the maximum entropy principle. The performance of this new index which has been shown in by many simulated and real examples is deducted from original combination of data analysis methods and the maximum entropy principle criterion. The method does not require any parameter adjustment, it is then completely automatic. Our new index V_{MEP} shows high robustness and superiority to the existing and recent ones, especially in overlapping clusters case.

MOTS – CLÉS : Classification non supervisée, Principe du Maximum d'Entropie, chevauchement de classes, nombre optimal de clusters.

KEY WORDS: unsupervised classification, the maximum entropy principle, overlapping clusters, optimal number of clusters.
