

Classification automatique: Projet

3 juin 2021

Le fichier *camera.csv* contient des caractéristiques d'appareil photos. il sera utilisé afin de mettre en oeuvre les différentes méthodes implémentées lors du projet. Tous les traitements et choix effectués devront être présentés et justifiés. Vous proposerez à chaque fois une description et une interprétation des classes obtenues. En plus de la version numérique, une version papier du rapport devra être rendu le 14 Juin 2021.

1. Analyser et traiter au préalable les données.
2. Implémenter l'algorithme k-means vu en cours.
3. Appliquer la méthode implémentée sur le jeu de données *camera.csv*, et proposer une segmentation adéquate.
4. Proposer une segmentation à partir de la méthode des *k-means* de R(ou python) et comparer la solution obtenue à celle de la question précédente.
5. Il existe des méthodes de segmentation dites *k-medoids*, effectuer une recherche documentaire et comparative de ces méthodes (PAM, CLARA) :
 - (a) Présenter ces méthodes, leurs avantages et leurs inconvénients.
 - (b) Appliquer ces méthodes aux données *camera.csv* et comparer aux résultats précédents.
6. Effectuer une classification mixte des données, ceci améliore-t-il les résultats ?