

Examen de Rattrapage (Classification et Analyse de données)

Le : 18/09/2011

durée : 1H : 30

Question de cours : (6 points)

1. Quelle est la différence entre des techniques de classification et des techniques de réduction de dimensions ?
2. Avant l'utilisation de la méthode de régression linéaire simple, quelle est la condition de travail nécessaire pour mettre en œuvre cette méthode ?
3. Expliquez le principe de l'algorithme de classification « *k*-means ».

Exercice 1 (6 points) :

Effectuer une classification hiérarchique ascendante sur *A* et tracer le dendrogramme correspondant.

$$A = \begin{pmatrix} 1 & 1 \\ 1.5 & 0.6 \\ 0.9 & 1.2 \\ -1 & -0.8 \\ -0.3 & 0.6 \\ 0.3 & -1 \end{pmatrix}$$

Exercice 2 (8 points) :

On observe les séries statistiques (accédants routière) suivantes :

Année	1985	1990	1995	2000	2005
Routes nationales	648	793	839	886	907
Autoroutes	450	676	794	957	1126

1. Calculer les droites d'ajustement par la méthode des moindres carrés pour les deux séries considérées.
2. Calculer et commenter les coefficients R^2 pour chacune de ces régressions linéaires.
3. Estimer le nombre d'accédants sur routes nationales en 2012. Même question pour les autoroutes.
4. En quelle année peut-on penser que les accédants sur les autoroutes ont dépassé ceux sur les routes nationales ?

Remarque : présentez vos résultats en millièmes.
La clarté des réponses et de la présentation sera prise en considération.