

INFORMATIQUE DECISIONNELLE

Exercice n°1 (8 points)

On considère le tableau X suivant croisant 6 individus et 3 variables :

	I	II	III
a	9	7	2
b	3	7	4
X = c	3	5	2
d	3	1	2
e	0	5	1
f	0	5	1

- 1) Calculer les moyennes et les variances des trois variables. En déduire le tableau Y centré et réduit (**1,5 points**)
- 2) Calculer la matrice des corrélations des trois variables (**1 point**)
- 3) Quelle est l'inertie totale du nuage des individus ? (**0,5 point**)
- 4) Ce tableau est soumis à l'analyse en composantes principales. Les deux premiers vecteurs axiaux issus de cette analyse sont :

$$u_1 = \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}, u_2 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{pmatrix}$$

- a) Calculer les coordonnées des 6 individus sur les deux axes définis par les vecteurs u_1 et u_2 . (**2 points**)
- b) Quelles sont les inerties projetées du nuage des 6 individus sur ces deux axes (**1 point**)
- c) Représenter ce nuage dans le plan engendré par ces deux axes (**0,5 point**)
- 5) Quelles sont les coordonnées des trois variables dans le plan engendré par les deux premières composantes principales (On ne demande pas le calcul de ces dernières) (**1 point**)
- 6) Représenter les trois variables dans ce plan (**0,5 point**)

Exercice n°2 (3 points)

On considère la matrice des distances entre les 5 éléments de l'ensemble I :

	a	b	c	d	e
a	0	9	5	8	13
b	9	0	4	1	14
c	5	4	0	6	10
d	8	1	6	0	7
e	13	14	10	7	0

Construire une hiérarchie indicée sur I en utilisant le critère d'agrégation du diamètre.

Exercice n°3 (3 points + 1 point)

1) Construire un arbre de décision à partir des données suivantes où V1, V2 V3 représentent les variables explicatives et « Classe » la variable à prédire, en utilisant l'algorithme C4.5

V1	V2	V3	Classe
True	90	A	Clas2
False	70	C	Clas1
False	80	C	Clas1
False	90	B	Clas2
False	80	C	Clas1
True	90	A	Clas2
True	70	B	Clas1
True	90	A	Clas2

2) Appliquer l'arbre ainsi construit pour déterminer la classe d'affectation des éléments supplémentaires :

V1	V2	V3	Classe
False	70	A	
True	70	C	
False	80	B	

Exercice n°4 (5 points)

On considère la base d'achats suivants :

Achats	Articles
A1	D E B
A2	D E B C
A3	E B A D
A4	D B C E
A5	D E
A6	B C D

On choisit les seuils suivants : support=50% et confiance=70%,

- 1) Déterminer les grands ensembles d'articles (2 points)
- 2) En déduire les règles fortes d'association. (2 points)
- 3) Y- a -t -il des règles erronées à rejeter ? (1 point)