

Nom et Prénom :

Code

Date de naissance :

Code

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Aboubekr BELKAID – TLEMCEM
Faculté des Sciences – Tidjani Haddam



الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي
جامعة أبو بكر بلقايد - تلمسان
كلية العلوم - تيجاني هدام

Concours d'accès à la formation de 3^{ème} cycle LMD 2022/2023

Samedi 11 février 2023

Domaine : Mathématique Informatique, Filière : Informatique

SPECIALITE : INTELLIGENCE ARTIFICIELLE ET TRAITEMENT D'IMAGE

Epreuve de spécialité : INTELLIGENCE ARTIFICIELLE ET ANALYSE DE DONNEES (Durée : 02h)

Partie : ANALYSE DE DONNEES

Questions (Réponses à choix multiples) : Cochez la (ou les) bonne(s) réponse(s).

Attention : Une mauvaise réponse élimine une réponse correcte. Une question qui a toutes les réponses cochées est considérée comme fausse.

1. Dans une ACP (1pt)

- ☐ Si deux individus sont très séparés sur le premier axe alors ils prennent des valeurs très différentes sur beaucoup de variables.
- ☐ Si un individu est proche du centre de gravité du nuage alors il prend des valeurs moyennes pour beaucoup de variables.
- ☐ Si deux individus sont superposés sur le plan principal à deux dimensions alors ils prennent les mêmes valeurs sur toutes les variables.
- ☐ Deux individus sont proches s'ils prennent des valeurs proches sur au moins une variable.
- ☐ Deux individus sont proches s'ils prennent des valeurs proches sur toutes les variables.

2. Parmi les indicateurs suivants, quel(s) est (sont) (les) l'indicateur(s) de liaison entre variables quantitatives : (1pt)

- ☐ Le coefficient de corrélation
- ☐ La moyenne
- ☐ La variance
- ☐ L'écart type

3. On considère l'échantillon de 50 clients d'un hypermarché dont on connaît l'âge [18,70], le revenu [20000, 150000] DA, le montant des achats [10, 25000] DA et le nombre d'enfants [0, 10]. On veut déterminer les caractéristiques principales de la clientèle. Nous souhaitons appliquer pour cela une analyse en composante principale (ACP). Cela nécessite un prétraitement, lequel ? (1pt)

Normalisation ou Standardisation ou Réduction

Ou $X = (X - \text{Moy}) / \text{ecart type}$

X : tableau de données

4. A l'aide de la représentation graphique figure 1, quelles sont les affirmations justes? **(1pt)**

- ☐ Les variables V1 et V3 contribuent dans des proportions équivalentes à l'apparition de l'axe 1.
- ☒ Les variables V2 et V3 contribuent dans des proportions équivalentes à l'apparition de l'axe 1
- ☒ La variable V1 contribue quasi-exclusivement à l'apparition de l'axe 2
- ☐ Les variables V1 et V3 sont très corrélées négativement entre elles
- ☒ Les variables V2 et V3 sont très corrélées négativement entre elles
- ☒ Toutes les variables sont correctement représentées dans le repère

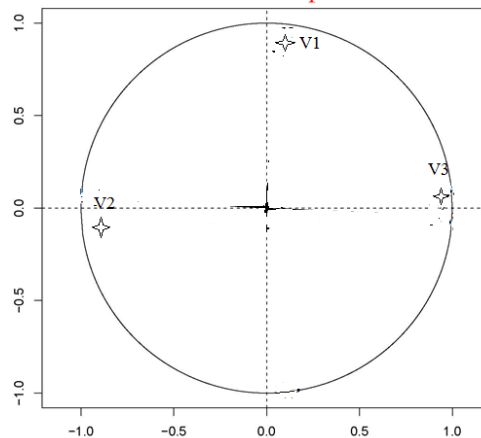


Figure1

5. Les éléments d'une table de contingence, à l'intersection de la ligne i et de la colonne j , sont des : **(1pt)**

- ☐ Couples de valeurs des deux modalités i et j des deux variables
- ☒ Effectifs conjoints relatifs aux deux modalités i et j des deux variables
- ☐ Fréquences relatives conjointes pour les deux modalités i et j des deux variables
- ☐ Fréquences conditionnelles en pourcentage ligne (ou pourcentage colonne) de la modalité i selon j (ou l'inverse)

6. En AFD (Analyse Factorielle Discriminante) à p variables explicatives et une variable à expliquer pour q modalités ($p > q$). Indiquer les affirmations justes **(1pt)**

- ☒ L'AFD est un modèle décisionnel.
- ☐ Le nombre maximum de facteurs discriminants est égal à q .
- ☒ Plus l'inertie inter-classes est grande plus le modèle est discriminant.

7. Parmi les méthodes de classification, il existe : **(1pt)**

- ☐ La méthode de classification hiérarchique des nuées dynamiques.
- ☐ La classification ascendante hiérarchique (CAH) dite classification supervisée.
- ☒ Aucune de ces réponses n'est correcte.

Nom et Prénom :

Code

Date de naissance :

Code

Exercice 1: Soit la classification des k-plus proches voisins K-PPV (1pt)

La classe C1 $\Rightarrow w_1(1 ; 1), w_2(2 ; 2), w_3(2,5 ; 2), w_4(1 ; 3)$

La classe C2 $\Rightarrow w_5(5,5 ; 7), w_6(6 ; 5), w_7(5 ; 5)$

La classe C3 $\Rightarrow w_8(-4 ; 4), w_9(-5 ; 4), w_{10}(-5 ; 5)$

- Déterminer le groupe d'appartenance (la classe) de l'individu $\omega^*(4,4)$ pour $k = 3$ ensuite pour $k = 5$ en utilisant la distance de Manhattan ($d(I, J) = |X1(I) - X1(J)| + |X2(I) - X2(J)|$)

K= 3 alors $W^* \in C2$


K= 5 alors $W^* \in C1$

Exercice 2 : Dans le cadre d'une analyse discriminante, cinq individus du même poids décrits par deux variables X et Y sont partitionnés en deux groupes G1 et G2 selon le Tableau 1. (2pts)

Ind	X	Y	Groupe
1	1	5	G1
2	3	6	G1
3	2	4	G1
4	3	3	G2
5	6	2	G2

Tableau 1 – cinq individus partitionnés en deux groupes

- Déterminer la matrice de variance - covariance Totale « T », la matrice de variance - covariance Inter « B » et la matrice de variance - covariance Intra « W ».


$$T = \begin{pmatrix} \frac{1}{5} & 14 & -8 \\ -8 & 10 \end{pmatrix}$$

$$B = \begin{pmatrix} \frac{1}{5} & 7.5 & -7.5 \\ -7.5 & 7.5 \end{pmatrix}$$

$$W = \begin{pmatrix} \frac{1}{5} & 6.5 & -0.5 \\ -0.5 & 2.5 \end{pmatrix}$$