



Programme Bac+5 Ingénierie Informatique
Concours d'accès en Master 1
Spécialité Big Data et Aide à la décision
Epreuve écrite

Durée : 3 heures

Cette épreuve comprend 2 parties distinctes.

Partie 1 : Culture Managériale (10 points)

I- **Dans le cadre d'une réflexion structurée, argumentée et illustrée d'exemples, vous répondrez au choix à l'un des trois sujets suivants (4 points) :**

Sujet 1 : ChatGPT ami ou ennemi ?

Sujet 2 : Pegasus est-il encore d'actualité ?

Sujet 3 : Quels sont les avantages pour les entreprises marocaines de s'adapter à l'ère numérique ?

II- **Donnez votre réponse dans l'espace dédié (3 points) :**

Q1 : Quel est le prix actuel d'un litre de Gasoil ?
Q2 : Quel est le taux directeur Actuel au Maroc ?
Q3 : Dans quelle ville marocaine est organisé chaque année le festival de Gnaoua ?
Q4 : Donnez un pays frontalier de l'Italie ?
Q5 : Quel est le plus vieux continent du monde ?
Q6 : Quel est le plafond prévu pour l'abattement forfaitaire déduit du salaire brut imposable ?

III- Cochez la réponse juste (3 points) :

1- Qu'est-ce que le Cloud Computing ?

A.	Un processeur très rapide pour les applications informatiques.
B.	Un concept de stockage et de traitement de données sur des serveurs distants.
C.	Un système d'exploitation pour ordinateurs portables.
D.	Un logiciel de gestion de projets.

2- Qu'est-ce que la réalité augmentée ?

A.	Une technologie de pointe pour augmenter la durée de vie des ordinateurs
B.	Un logiciel de reconnaissance d'images
C.	Une technologie qui ajoute des informations numériques en temps réel sur des images de la vie réelle
D.	Une méthode de compression de fichiers

3- Qu'est-ce que le Machine Learning ?

A.	Une méthode de développement de logiciels pour les débutants
B.	Une technique d'apprentissage automatique utilisée en intelligence artificielle
C.	Une méthode de stockage de données en ligne
D.	Un outil pour écrire des scripts de code rapidement

4- Qu'est-ce que la blockchain ?

A.	Un réseau social pour les programmeurs informatiques
B.	Une technologie de stockage de données distribuée et sécurisée
C.	Un langage de programmation utilisé pour les applications mobiles
D.	Un outil pour surveiller la vitesse d'un ordinateur

5- Qu'est-ce que le Big Data ?

A.	Une base de données très complexe utilisée pour le stockage de données sensibles
B.	Une méthode d'analyse de données à grande échelle
C.	Un système d'exploitation pour les grands serveurs
D.	Un logiciel de gestion de projet en ligne

6- Qu'est-ce que l'Internet des objets (IoT) ?

A.	Un réseau de robots pour la surveillance de la ville
B.	Une technologie de communication pour les voitures autonomes
C.	Un système de contrôle pour les équipements industriels
D.	Une connectivité de différents objets physiques entre eux via Internet

7- Qu'est-ce que l'Intelligence Artificielle ?

A.	Un logiciel de traitement de texte avancé
B.	Une technologie capable de reproduire des comportements intelligents de l'être humain
C.	Un système de contrôle pour les équipements de sécurité
D.	Un système de reconnaissance vocale avancé

8- Qu'est-ce que le deep learning ?

A.	Une méthode d'apprentissage automatique qui utilise des réseaux de neurones artificiels
B.	Une technique pour optimiser les algorithmes de recherche
C.	Une méthode de stockage de données dans le Cloud
D.	Une méthode pour effectuer des analyses de données en temps réel

9- Qu'est-ce qu'un pare-feu (firewall) ?

A.	Un logiciel pour gérer les mots de passe sur les ordinateurs
B.	Un type de virus informatique pour empêcher l'accès aux données
C.	Un dispositif pour protéger les ordinateurs et les réseaux des attaques externes
D.	Un programme pour protéger les ordinateurs contre les programmes malveillants

10- Qu'est-ce qu'un VPN (Virtual Private Network) ?

A.	Un logiciel pour effectuer des appels téléphoniques gratuits sur Internet.
B.	Un type de réseau permettant de connecter des ordinateurs à un serveur distant.
C.	Une application permettant d'envoyer des messages cryptés en toute sécurité.
D.	Une solution pour établir une connexion sécurisée entre deux ordinateurs à distance

11- Qu'est-ce qu'un logiciel libre ?

A.	Un logiciel gratuit pour les consommateurs.
B.	Un logiciel open source dont le code source est disponible pour tous.
C.	Un programme pour accéder à des fichiers sur des ordinateurs distants.
D.	Un logiciel pour enregistrer des données sur un disque dur externe

12- Qu'est-ce qu'un "cookie" ?

A.	Un petit fichier stocké sur un ordinateur par un site web
B.	Un logiciel de surveillance
C.	Un dispositif de stockage pour les données sensibles
D.	Un système de sécurité pour les transactions en ligne

Partie 2 : Spécialité Big Data et Aide à la décision

I. Etude de cas : (4 points)

Une entreprise de vente en ligne dispose d'une base de données de ventes contenant des informations sur les achats des clients, telles que le nom du client, le produit acheté, la date d'achat, le montant de l'achat, etc. L'entreprise souhaite utiliser ces données pour améliorer ses ventes et sa rentabilité. Pour ce faire, elle a décidé de mener une analyse de données de ventes.

Répondez aux questions suivantes en justifiant votre réponse :

Q1/ Quelle est la méthodologie d'analyse de données la plus appropriée pour cette étude de cas ?
Q2/ Quel est l'objectif de cette étude de cas ?
Q3/ Quels types de données sont utilisés dans cette étude de cas ?
Q4/ Quel est l'outil de Big Data le plus approprié pour cette étude de cas ?

II – Répondez dans l'espace dédiée (3 points) :

Q1/ Expliquer la différence entre Managed Table et External Table
Q2/ Donner la commande pour créer sur Hive une table Personne (num , nom , prenom, ville)
Q3/ Donner la commande pour charger les données du bureau de la machine locale vers la table personne
Q4/ Donner la commande pour créer une table PersonnePart contenant les mêmes champs partitionnés par ville
Q5/ Expliquer le rôle des Buckets sur Hive

Q6/ Donner la commande pour activer la Partition dynamique sur Hive
--	----------------

III – Cochez la réponse juste (3 points) :

1- Qu'est-ce que le big data ?

E.	A. La collecte et l'analyse de grandes quantités de données
F.	B. L'analyse de données provenant de sources multiples
G.	C. L'analyse de données structurées et non structurées
H.	D. Toutes les réponses ci-dessus

2- Qu'est-ce que l'apprentissage profond ?

A.	Une technique d'apprentissage supervisé
B.	Une technique d'apprentissage non supervisé
C.	Une technique d'apprentissage renforcé
D.	Une technique d'apprentissage semi-supervisé

3- Qu'est-ce que le prétraitement des données ?

A.	Une technique pour convertir les données brutes en données exploitables
B.	Une technique pour réduire la taille des données
C.	Une technique pour augmenter la taille des données
D.	Une technique pour supprimer les données inutiles

4- Qu'est-ce qu'un réseau de neurones ?

A.	Un ensemble de processus informatiques interconnectés qui simulent le fonctionnement du cerveau humain
B.	Un ensemble de processus informatiques interconnectés qui simulent le fonctionnement du système nerveux central
C.	Un ensemble de processus informatiques interconnectés qui simulent le fonctionnement du système endocrinien
D.	Un ensemble de processus informatiques interconnectés qui simulent le fonctionnement du système lymphatique

5- Qu'est-ce que Hadoop ?

A.	Une base de données relationnelle
B.	Un framework pour le traitement distribué de données
C.	Un outil d'analyse de données
D.	Un logiciel de visualisation de données

6- Qu'est-ce que Spark ?

A.	Un outil de traitement de données en batch
B.	Un framework de traitement de données en temps réel
C.	Un outil d'analyse de données

D.	Un logiciel de visualisation de données
----	---

7- Qu'est-ce que le clustering de données ?

A.	Une méthode pour classer les données en fonction de leur contenu
B.	Une méthode pour visualiser les données dans un espace multidimensionnel
C.	Une méthode pour partitionner les données en groupes homogènes
D.	Une méthode pour réduire la dimensionnalité des données

8- Qu'est-ce que le réseau neuronal artificiel ?

A.	Une méthode de traitement de données non supervisée
B.	Une méthode de traitement de données supervisée
C.	Un algorithme de machine learning inspiré du cerveau humain
D.	Un algorithme de clustering

9- Qu'est-ce que le deep learning ?

A.	Une méthode de traitement de données non supervisée
B.	Une méthode de traitement de données supervisée
C.	Un algorithme de machine learning inspiré du cerveau humain
D.	Un algorithme de clustering

10- Qu'est-ce que la régression linéaire ?

A.	Une méthode de classification de données
B.	Une méthode de clustering de données
C.	Une méthode de prédiction de données
D.	Une méthode d'analyse de données

11- Qu'est-ce que la classification de données ?

A.	Une méthode de partitionnement des données
B.	Une méthode de prédiction de données
C.	Une méthode de visualisation de données
D.	Une méthode de réduction de dimensionnalité

12- Qu'est-ce que la visualisation de données ?

A.	Une méthode pour réduire la dimensionnalité des données
B.	Une méthode pour partitionner les données en groupes homogènes
C.	Une méthode pour représenter les données de manière graphique
D.	Une méthode pour nettoyer les données