EXAMEN

Module : Machine Learning

Enseignant(s): Dr. Anis CHERFI **anis.cherfi@yncrea.fr**

Date: 27/11/2026 **Durée :** 1h

A - Questions aux choix multiples

Veuillez choisir une ou plusieurs réponses :

1. Tous les exemples suivants sont des applications de Machine Learning, sauf :
    a. Personnaliser les campagnes de marketing en fonction de la démographie des clients et de leur historique d'achat.
    b. Analyser les IdO (internet des objets) afin de prédire les problèmes d'équipements avant qu'ils se produisent.
    c. Détecter les activités frauduleuses dans des transactions financières.
    d. Analyser les revenus passés pour déterminer la cause de la chute des ventes.
2. Quelles sont les principales catégories de Machine Learning ? Cochez les réponses qui sont des catégories de Machine Learning :
    a. Apprentissage supervisé
    b. Toutes les réponses sont des catégories de machine learning.
    c. Les K plus proches voisins (KNN)
    d. Apprentissage par renforcement
    e. Apprentissage non supervisé
3. Qu'est-ce que la gestion des valeurs manquantes ?
    a. Le processus de traitement des données manquantes dans un ensemble de données, soit en les imputant, soit en les supprimant.
    b. Le processus de transformation des données en un format différent.
    c. Le processus de sélection des variables pertinentes pour l'analyse en éliminant les variables non pertinentes ou redondantes.

4. Qu'est-ce que la Feature Engineering ?
    a. Le processus de sélection des variables pertinentes pour l'analyse en éliminant les variables non pertinentes ou redondantes.
    b. Le processus de conversion de variables continue en variables catégorielles pour permettre une analyse plus facile.
    c. Le processus de création de nouvelles variables à partir des données existantes pour améliorer l'analyse.
5. Dans une société future, une machine est utilisée pour prédire un crime avant qu'il ne se produise. Si vous étiez responsable du réglage de cette machine, quelle métrique d'évaluation voudriez-vous maximiser pour vous assurer qu'aucune personne innocente (personnes qui ne sont pas sur le point de commettre un crime) ne soit emprisonnée (où le crime est l'étiquette positive) ?
    a. Accuracy
    b. Precision
    c. Recall
    d. F1

6. Quelles mesures d'évaluation sont couramment utilisées pour évaluer les performances d'un modèle de classification lorsqu'il y a un problème de déséquilibre de classe ?
    a. Précision et rappel
    b. Précision et exactitude
    c. Exactitude et erreur
    d. Précision et erreur

B - Régression

Soit le modèle de régression linéaire de la figure 1. Le modèle mesure la variable cible Prix en fonction de l'attribut Taille

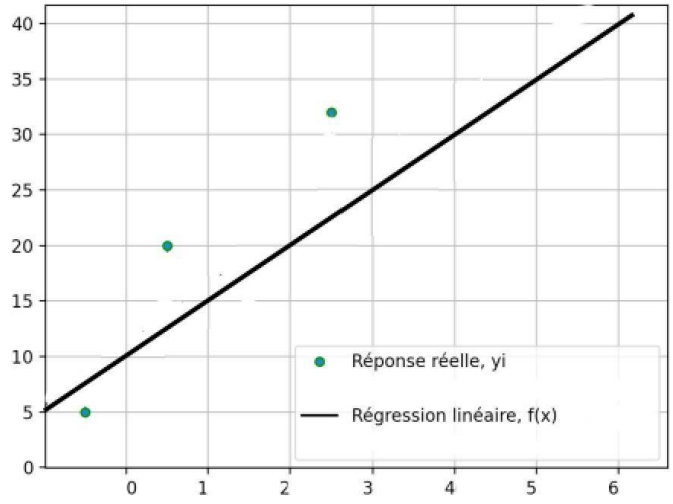


Figure 1

1. Donner la valeur à l'origine de ce modèle
2. Donner la valeur de la pente de ce modèle
3. Donner la formule de régression du modèle de la figure 1
4. Donner la prédiction de la valeur de prix lorsque (Taille = 2 ; et Taille = 5)

C - Classification

L'entreprise DQ-Helth cherche à étudier le degré de satisfaction de ses employés. Dans ce cadre, le service ressources humaines de l'entreprise a collecté un ensemble de données pour les aider à prendre des décisions concernant les salariés non satisfait et donc susceptible de démissionner.

Vous pouvez voir l'ensemble de données dans la feuille de calcul de la figure 2.

Poste	Education	Age	État civil	Heures supplémentaires	Formation	Prime	Satisfaction
1		1	1	1	2	1	insatisfait
1		2	2	2	1	1	insatisfait
1		1	1	2	2	1	insatisfait
2		1	2	2	2	1	satisfait
2		1	2	1	1	2	insatisfait

.....

Figure 2

Vous avez entrainé deux algorithmes d'apprentissage automatique pour résoudre ce problème :

- Algorithme 1 : un modèle des k plus proches voisin avec k = 13. Ce modèle produit 100% de précision dans l'échantillon d'apprentissage et 80% dans l'échantillon test.

- Algorithme 2 : un modèle d'arbre de décision en utilisant les paramètres par défaut.

1. Quel est le type d'apprentissage automatique qui convient à notre étude de cas ? Préciser la variable cible.

2. Interpréter les résultats du modèle de l'algorithme 1. Décider si vous allez le garder et argumenter votre réponse.

3. Quelles sont les manipulations possibles pour améliorer le modèle de l'algorithme 1.