**Examen Machine Learning Part 1 :**

**Théorie et pratiques**

Aucune question ne pourra être posée durant l'examen.

En cas de doute concernant le sujet, vous poursuivrez votre réponse en expliquant vos hypothèses.

Durée : 1h30 (10h15- > 11h45)

Épreuve du 08/11/2024

# Modalités du travail

* Durée : 2h ;
* Aucun document autorisé, calculatrice non autorisée ;
* Ecrire vos réponses sur la copie, dans les cases réservées à cet effet ;
* Mettre vos noms et prénoms sur chaque feuille ;
* ***Toute réponse donnée sans explications sera considérée comme incorrecte***
* ***Tout texte indéchiffrable sera considéré comme une absence de réponse***

**Part 1 : Questions générales (9 points)**

1. Quels sont les principaux facteurs ayant favorisé la popularité de l’apprentissage automatique ? **(0.25 pt)**

|  |
| --- |
|  |

1. Quelle est la différence entre l'apprentissage automatique (ML) et l'apprentissage profond (DL) ? **(0.25 pt)**

|  |
| --- |
|  |

1. Pourquoi l'apprentissage profond est-il souvent préféré à l'apprentissage automatique classique dans certaines applications ? **(0.5 pt)**

|  |
| --- |
|  |

1. En quoi consiste l'algorithme de descente de gradient et pourquoi est-il essentiel en apprentissage automatique ? **(0.5 pt)**

|  |
| --- |
|  |

1. Définissez brièvement la rétropropagation (backpropagation) et expliquez son rôle dans l'entraînement des réseaux neuronaux. **(0.5 pt)**

|  |
| --- |
|  |

1. Quelles sont les différences principales entre la descente de gradient et la rétropropagation ? **(0.25 pt)**

|  |
| --- |
|  |

1. Expliquez la différence entre des fonctions convexes et non convexes. **(0.25 pt)**

|  |
| --- |
|  |

1. La régression logistique en tant qu'algorithme de classification :
   1. Pourquoi est-elle appelée « régression logistique » et non « classification logistique » ? **(0.25 pt)**
   2. Donnez l’équation de base de la régression logistique. **(0.25 pt)**
   3. Comment estimer les paramètres dans une régression logistique ? **(0.5 pt)**
   4. Quelle est la fonction de coût de la régression logistique ? Donnez sa formule. **(0.5 pt)**

|  |
| --- |
|  |

1. Quelles sont les limites de l’accuracy comme mesure d’évaluation ? Quelles alternatives proposez-vous pour évaluer un modèle de machine learning ? **(0.5 pt)**

|  |
| --- |
|  |

1. Pourquoi utilise-t-on une forêt aléatoire au lieu de simples arbres de décision ? Quelles sont les étapes nécessaires pour construire une forêt aléatoire ? **(0.75 pt)**

|  |
| --- |
|  |

1. Citez deux méthodes d’évaluation utilisées pour les modèles de régression. **(0.25 pt)**

|  |
| --- |
|  |

1. Qu'est-ce que la validation croisée en K-fold et pourquoi l’utilise-t-on ? **(0.25 pt)**

|  |
| --- |
|  |

1. Donnez des exemples d’hyperparamètres et expliquez comment les optimiser. **(0.5 pt)**

|  |
| --- |
|  |

1. Qu'est-ce que le ROC-AUC et comment le construire ? **(0.75 pt)**

|  |
| --- |
|  |

1. Expliquez le concept de "mini-batch stochastique gradient descent". **(0.25 pt)**

|  |
| --- |
|  |

1. Décrivez le compromis biais-variance (bias-variance tradeoff) et sa relation avec le surapprentissage (overfitting) et le sous-apprentissage (underfitting). **(0.5 pt)**

|  |
| --- |
|  |

1. Qu’est-ce que le "dropout" en deep learning et pourquoi l’utilise-t-on ? **(0.25 pt)**

|  |
| --- |
|  |

1. Donnez les principaux composants d’un réseau de neurones convolutif (CNN) classique et expliquez l'importance de chacun. **(1 pt)**

|  |
| --- |
|  |

**Part 2 : KNN (3 points)**

Supposons que vous disposez d’un petit ensemble de données avec trois caractéristiques et une classe cible binaire (classe 0 et classe 1). Voici les données :

|  |  |  |  |
| --- | --- | --- | --- |
| **Feature 1** | **Feature 2** | **Feature 3** | **Class** |
| 1.2 | 0.8 | 2.5 | 0 |
| 2.0 | 1.5 | 2.9 | 0 |
| 1.8 | 1.0 | 3.1 | 0 |
| 2.5 | 2.2 | 3.5 | 1 |
| 3.0 | 2.7 | 3.8 | 1 |
| 2.8 | 3.0 | 3.3 | 1 |

* **Calculer la distance :** utiliser la distance euclidienne pour calculer la distance entre un nouveau point de données et tous les points de l'ensemble de données. (2 pts)

Nouveau point de données : [Feature 1 = 2.3, Feature 2 = 1.9, Feature 3 = 3.0]

* **Utiliser K = 3** (les 3 voisins les plus proches) pour classer ce nouveau point de données. (1 pt)

|  |
| --- |
|  |

**Part 3 : Decision Tree (4 points)**

Les arbres de décision (également appelés arbres de classification et de régression « CART ») sont un type d'algorithme d'apprentissage automatique qui prend des décisions en divisant les données en sous-ensembles en fonction de certaines caractéristiques. Ils sont utilisés à la fois pour la classification (attribution d'une étiquette à un élément) et la régression (prédiction d'une valeur numérique).

**Question 1 :** Quelles sont les composantes (l’architecture) d’un arbre de décision ? et quelles sont les étapes à suivre pour créer un arbre de décision ? (1 pt)

**Problème :** Considérons l'ensemble de données ci-dessous, en utilisant un critère de division, quel est le nœud racine de cet arbre (Donnez les étapes avec les formules que vous avez suivies pour déterminer le nœud racine) ? (3 pts)

|  |  |  |  |
| --- | --- | --- | --- |
| **Outlook** | **Temp** | **Humidity** | **Play** |
| Sunny | Hot | High | 0 |
| Sunny | Hot | High | 0 |
| Overcast | Hot | High | 1 |
| Rainy | Mild | High | 1 |
| Rainy | Cool | Normal | 1 |
| Rainy | Cool | Normal | 0 |
| Overcast | Cool | Normal | 1 |
| Sunny | Mild | High | 0 |
| Sunny | Cool | Normal | 1 |
| Rainy | Mild | Normal | 1 |
| Sunny | Mild | Normal | 1 |
| Overcast | Mild | High | 1 |
| Overcast | Hot | Normal | 1 |
| Rainy | Mild | High | 0 |

|  |
| --- |
|  |

**Part 4 : MLP-ANN (4 points)**

Supposons que vous disposez d'un ensemble de données simple avec deux caractéristiques et que votre objectif est de le classer dans l'une des deux classes (Classe 0 ou Classe 1). Vous allez créer un réseau neuronal monocouche (un réseau neuronal avec une seule couche cachée) avec deux neurones dans la couche cachée et un neurone de sortie.

Voici un petit ensemble de données avec 4 points pour l'entraînement :

|  |  |  |
| --- | --- | --- |
| **Feature 1** | **Feature 2** | **Feature 3** |
| 0.1 | 0.3 | 0 |
| 0.4 | 0.5 | 1 |
| 0.7 | 0.8 | 1 |
| 0.6 | 0.1 | 0 |

**Structure du ANN**

* Couche d'entrée : 2 nœuds d'entrée (un pour chaque fonction).
* Couche cachée : 2 neurones, chacun avec une fonction d'activation ReLU.
* Couche de sortie : 1 neurone avec une fonction d'activation sigmoïde, produisant une sortie de probabilité (classe 0 ou classe 1).

**Étapes :**

* **Initialiser les poids et les biais :**
  1. **Hidden layer :**
     1. Les poids du neurone 1 : w11 = 0.2, w12 = 0.4
     2. Biais du neurone 1 : b1 = 0.1
     3. Les poids du neurone 2 : w21 = 0.5, w22 = 0.3
     4. Biais du neurone 2 : b2 = 0.2
  2. **Output layer :**
     1. Les poids : wo1 = 0.6 et wo2 = 0.7
     2. Biais : bo = -0.3
* **Calcul du passage en avant (Forward propagation) :**
  1. Calculez la sortie de chaque neurone de la couche cachée à l'aide de la fonction d'activation **ReLU**
  2. Calculez la sortie du neurone final (couche de sortie) à l'aide de la fonction d'activation **sigmoïde**
* **Classer un nouveau point de données :**
  + Étant donné un nouveau point de données avec la Feature 1 = 0.5 et la Feature 2 = 0.4, calculez les sorties via le réseau et classez-le en classe 0 ou en classe 1 en fonction d'un seuil de 0.5.

**Questions :**

1. Calcul de la couche cachée : Calculez la sortie de chaque neurone de la couche cachée pour le nouveau point de données (0.5, 0.4). (1 pt)
2. Calcul de la couche de sortie : Utilisez les sorties de la couche cachée pour calculer la valeur du neurone de sortie. Appliquez la fonction sigmoïde et déterminez la classe prédite finale (utilisez 0.5 comme seuil de décision). (1 pt)
3. Interprétation :
   1. Expliquez le rôle des pondérations et des biais dans l'influence de la sortie. (1 pt)
   2. Discutez de l'importance des fonctions d'activation telles que ReLU et sigmoïde dans ce réseau. (1 pt)

|  |
| --- |
|  |