

Voici 30 questions à choix multiples pour la Note 1 (CSI 4506 - Automne 2024 : Introduction aux structures de données et algorithmes avancés) avec les réponses et explications :

---

**1. Quel est l'objectif principal du cours CSI 4506 ?**

- A) Apprendre les langages de programmation
- B) Comprendre les structures de données et les algorithmes avancés
- C) Étudier les systèmes d'exploitation
- D) Apprendre la création de sites web

Réponse : B

Explication : Le cours se concentre sur les structures de données avancées et les algorithmes.

**2. Quelle est une structure de données étudiée dans le cours ?**

- A) Tableaux
- B) Arbres
- C) Systèmes de fichiers
- D) Réseaux

Réponse : B

Explication : Les arbres sont des structures de données importantes étudiées dans le cours.

**3. Qu'est-ce que la complexité temporelle ?**

- A) Le temps nécessaire pour écrire un programme
- B) Le temps d'exécution d'un algorithme en fonction de la taille de l'entrée
- C) Le temps nécessaire pour compiler un programme
- D) Le temps de téléchargement d'un fichier

Réponse : B

Explication : La complexité temporelle évalue le temps que met un algorithme à s'exécuter en fonction de la taille de ses entrées.

**4. Quelle structure est utilisée pour représenter des relations hiérarchiques ?**

- A) Liste
- B) Graphe
- C) Arbre
- D) Tableau

Réponse : C

Explication : Les arbres sont utilisés pour représenter des relations hiérarchiques.

**5. Quel type d'algorithme est souvent analysé en termes de complexité spatiale ?**

- A) Algorithme de tri
- B) Algorithme de recherche
- C) Algorithme de compression
- D) Algorithme récursif

Réponse : D

Explication : Les algorithmes récursifs sont souvent analysés pour leur utilisation de la mémoire (complexité spatiale).

**6. Quel est un exemple de structure de données linéaire ?**

- A) Graphe
- B) Liste chaînée
- C) Arbre binaire
- D) Pile

Réponse : B

Explication : La liste chaînée est une structure de données linéaire où chaque élément pointe vers le suivant.

**7. Dans un arbre binaire, combien de sous-arbres chaque nœud peut-il avoir au maximum ?**

- A) 1
- B) 2
- C) 3
- D) 4

Réponse : B

Explication : Un arbre binaire est défini comme ayant au plus deux sous-arbres par nœud.

**8. Quel algorithme est utilisé pour trouver le plus court chemin dans un graphe ?**

- A) Algorithme de Kruskal
- B) Algorithme de Dijkstra
- C) Algorithme de tri rapide
- D) Algorithme de fusion

Réponse : B

Explication : L'algorithme de Dijkstra est couramment utilisé pour trouver le chemin le plus court dans un graphe.

**9. Quelle est la complexité temporelle du tri par insertion dans le pire des cas ?**

- A)  $O(n)$
- B)  $O(n^2)$
- C)  $O(\log n)$
- D)  $O(n \log n)$

Réponse : B

Explication : Le tri par insertion a une complexité temporelle de  $O(n^2)$  dans le pire des cas.

10. Quelle est la principale différence entre une pile et une file ?

- A) Une pile suit LIFO et une file suit FIFO
- B) Une pile stocke plus de données
- C) Une file est plus rapide
- D) Une pile utilise moins de mémoire

Réponse : A

Explication : Une pile fonctionne selon le principe LIFO (dernier entré, premier sorti) tandis qu'une file suit FIFO (premier entré, premier sorti).

11. Un graphe est un ensemble de :

- A) Nœuds et feuilles
- B) Sommets et arêtes
- C) Piles et files
- D) Entrées et sorties

Réponse : B

Explication : Un graphe est constitué de sommets (nœuds) et d'arêtes (liens entre les sommets).

12. Quel type de parcours est utilisé pour visiter tous les nœuds d'un arbre ?

- A) Parcours en largeur d'abord
- B) Recherche binaire
- C) Parcours en profondeur d'abord
- D) Tri rapide

Réponse : C

Explication : Le parcours en profondeur d'abord explore chaque branche de l'arbre aussi loin que possible avant de revenir en arrière.

13. Quelle est la différence entre un arbre binaire et un arbre binaire de recherche ?

- A) Un arbre binaire de recherche est ordonné
- B) Un arbre binaire a plus de nœuds
- C) Un arbre binaire de recherche a des cycles
- D) Un arbre binaire n'a pas de racine

Réponse : A

Explication : Un arbre binaire de recherche est un arbre dans lequel chaque nœud respecte un ordre : les nœuds à gauche sont inférieurs, et ceux à droite sont supérieurs.

14. Quel algorithme est souvent utilisé pour le tri des tableaux ?

- A) Tri à bulles
- B) Tri par fusion
- C) Recherche binaire
- D) Algorithme de Dijkstra

Réponse : B

Explication : Le tri par fusion est un algorithme de tri diviser-pour-régner efficace pour les tableaux.

15. Dans une liste chaînée, chaque élément contient :

- A) Une référence à l'élément suivant
- B) Une référence à l'élément précédent
- C) Une référence à l'élément suivant et précédent
- D) Aucune référence

Réponse : A

Explication : Chaque élément d'une liste chaînée contient une référence (ou un lien) vers l'élément suivant.

16. Quel est le but principal de l'analyse de la complexité d'un algorithme ?

- A) Calculer le temps d'exécution exact
- B) Estimer les ressources (temps et espace) nécessaires
- C) Écrire du code plus propre
- D) Tester la sécurité d'un programme

Réponse : B

Explication : L'analyse de la complexité vise à estimer les ressources nécessaires à un algorithme.

17. Quelle est la structure de données utilisée pour gérer les priorités dans un système ?

- A) File de priorité
- B) Graphe
- C) Liste chaînée
- D) Pile

Réponse : A

Explication : Une file de priorité permet de gérer les tâches selon leur priorité.

18. La complexité spatiale d'un algorithme mesure :

- A) La quantité de mémoire utilisée par l'algorithme
- B) La vitesse de calcul d'un programme
- C) Le temps nécessaire pour exécuter un programme
- D) La taille de l'entrée

Réponse : A

Explication : La complexité spatiale évalue la quantité de mémoire utilisée par l'algorithme.

19. Quelle méthode permet de réduire le temps d'exécution d'un algorithme ?

- A) Utilisation de la récursivité
- B) Réduction de la taille des données
- C) Optimisation de la complexité
- D) Multiplication des nœuds

Réponse : C

Explication : L'optimisation de la complexité permet de réduire le temps d'exécution.

20. Dans quel type de structure les éléments sont retirés dans l'ordre où ils ont été ajoutés ?

- A) Pile
- B) File
- C) Arbre
- D) Graphe

Réponse : B

Explication : Dans une file, les éléments sont retirés selon le principe FIFO (premier entré, premier sorti).

21. Quel est le principe d'une recherche binaire ?

- A) Diviser la recherche en deux parties égales
- B) Comparer chaque élément un par un
- C) Ajouter les éléments à la fin de la liste
- D) Comparer uniquement les nombres pairs

Réponse : A

Explication : La recherche binaire divise les données en deux à chaque étape pour accélérer la recherche.

22. Dans un graphe orienté, une arête pointe :

- A) Dans une direction spécifique
- B) Vers tous les nœuds
- C) Dans toutes les directions
- D) Aucun des choix

Réponse : A

Explication : Dans un graphe orienté, chaque arête a une direction spécifique entre deux nœuds.

23. Quel algorithme est efficace pour trouver un arbre couvrant minimal ?

- A) Algorithme de Kruskal
- B) Algorithme de tri rapide
- C) Algorithme de Dijkstra
- D) Recherche linéaire

Réponse : A

Explication : L'algorithme de Kruskal est utilisé pour trouver un arbre couvrant minimal dans un graphe pondéré.

24. Un arbre est appelé complet si :

- A) Tous ses niveaux sont complètement remplis sauf peut-être le dernier
- B) Chaque nœud a au moins trois enfants
- C) Il ne contient qu'un seul nœud
- D) Il ne contient aucune feuille

Réponse : A

Explication : Un arbre complet est un arbre où tous les niveaux sont remplis, sauf peut-être le dernier.

25. La complexité temporelle moyenne de la recherche dans une liste chaînée est :

- A)  $O(1)$
- B)  $O(n)$
- C)  $O(\log n)$
- D)  $O(n^2)$

Réponse : B

Explication : En moyenne, la recherche dans une liste chaînée prend  $O(n)$  car on peut devoir parcourir toute la liste.

26. Quel algorithme est utilisé pour trier en temps  $O(n \log n)$  ?

- A) Tri rapide
- B) Tri à bulles
- C) Tri par insertion
- D) Recherche binaire

Réponse : A

Explication : Le tri rapide a une complexité moyenne de  $O(n \log n)$ .

27. Une liste doublement chaînée permet :

- A) De naviguer dans les deux sens
- B) De ne naviguer que vers l'avant
- C) De trier automatiquement les données
- D) D'économiser de la mémoire

Réponse : A

Explication : Une liste doublement chaînée contient des liens vers l'élément précédent et suivant.

28. Quel est un exemple de tri non-comparatif ?

- A) Tri par comptage
- B) Tri par insertion
- C) Tri à bulles
- D) Tri rapide

Réponse : A

Explication : Le tri par comptage est un tri non-comparatif qui trie les éléments en comptant le nombre d'occurrences.

29. Quel est le principal avantage de la récursivité dans certains algorithmes ?

- A) Simplicité et clarté de la solution
- B) Utilisation plus efficace de la mémoire
- C) Accélération des calculs
- D) Réduction du code source

Réponse : A

Explication : La récursivité permet souvent une solution plus simple et élégante, bien que ce ne soit pas toujours la plus efficace en termes de performance.

30. Quel est l'inconvénient principal de l'utilisation de la récursivité excessive ?

- A) Risque de débordement de pile
- B) Utilisation excessive de la CPU
- C) Trop rapide à exécuter
- D) Limitation du nombre d'éléments

Réponse : A

Explication : L'utilisation excessive de la récursivité peut entraîner un débordement de la pile (stack overflow) si l'appel de fonctions récursives est trop profond.

---

Voici 30 questions basées sur la Note 1 avec des réponses et explications pour chaque.

Voici 30 questions à choix multiples pour la Note 2 (Fondements de l'Intelligence Artificielle - CSI 4506) avec les réponses et explications :

---

### 1. Qu'est-ce que l'intelligence artificielle (IA) ?

- A) Un système qui répare des ordinateurs
- B) Une méthode pour automatiser des tâches complexes
- C) Un domaine de la science informatique qui vise à créer des machines capables de reproduire l'intelligence humaine
- D) Un logiciel utilisé pour analyser des photos

Réponse : C

Explication : L'IA vise à créer des systèmes capables d'exécuter des tâches qui nécessitent l'intelligence humaine.

### 2. Qu'est-ce que l'apprentissage automatique (machine learning) ?

- A) Une méthode de programmation manuelle
- B) Une sous-discipline de l'IA où les machines apprennent à partir de données
- C) Un processus d'optimisation des bases de données
- D) Un moyen d'améliorer la vitesse des ordinateurs

Réponse : B

Explication : L'apprentissage automatique permet aux machines d'apprendre à partir de données sans être explicitement programmées.

### 3. Quel type d'apprentissage utilise des données étiquetées ?

- A) Apprentissage non supervisé
- B) Apprentissage supervisé

- C) Apprentissage par renforcement
- D) Apprentissage non étiqueté

Réponse : B

Explication : L'apprentissage supervisé nécessite des données étiquetées pour entraîner un modèle.

### 4. Quel est un exemple d'apprentissage non supervisé ?

- A) Classification supervisée
- B) K-means
- C) Régression linéaire
- D) Arbre de décision

Réponse : B

Explication : K-means est un algorithme d'apprentissage non supervisé utilisé pour regrouper des données sans étiquettes.

### 5. Dans l'apprentissage supervisé, les données sont accompagnées de :

- A) Données brutes
- B) Étiquettes
- C) Modèles pré-entraînés
- D) Graphiques

Réponse : B

Explication : Les étiquettes sont les résultats attendus, utilisées pour entraîner le modèle.

### 6. Qu'est-ce qu'un réseau de neurones artificiels ?

- A) Un système biologique basé sur le cerveau humain
- B) Un modèle informatique qui imite le fonctionnement des neurones biologiques
- C) Un programme qui calcule des probabilités
- D) Un algorithme de tri rapide

Réponse : B

Explication : Les réseaux de neurones artificiels imitent la structure et le fonctionnement des neurones du cerveau humain.

### 7. Quel est l'objectif principal d'un réseau de neurones ?

- A) Automatiser la gestion des fichiers
- B) Optimiser la vitesse de calcul
- C) Apprendre à partir des données pour faire des prédictions ou des classifications
- D) Minimiser la consommation d'énergie

Réponse : C

Explication : Les réseaux de neurones sont utilisés pour reconnaître des modèles, faire des prédictions et des classifications.

### 8. Qu'est-ce qu'une couche cachée dans un réseau de neurones ?

- A) Une couche qui ne fait pas partie du modèle
- B) Une couche intermédiaire entre les entrées et les sorties



- C) Une couche de stockage de données
- D) Une couche qui prédit les résultats

Réponse : B

Explication : Les couches cachées sont situées entre la couche d'entrée et la couche de sortie et sont essentielles pour l'apprentissage.

### 9. Quel type d'apprentissage repose sur un système de récompenses et punitions ?

- A) Apprentissage supervisé
- B) Apprentissage non supervisé
- C) Apprentissage par renforcement
- D) Apprentissage en ligne

Réponse : C

Explication : L'apprentissage par renforcement utilise des récompenses et des punitions pour apprendre à l'agent à prendre les meilleures décisions.

### 10. Quelle est la principale utilisation des arbres de décision en IA ?

- A) Analyser des graphes
- B) Classifier des données
- C) Compiler des programmes
- D) Trier des données

Réponse : B

Explication : Les arbres de décision sont utilisés pour la classification et la prédiction.

### 11. Quel algorithme est utilisé pour optimiser les réseaux de neurones ?

- A) Algorithme génétique
- B) Descente de gradient
- C) Algorithme de tri rapide
- D) Recherche binaire

Réponse : B

Explication : La descente de gradient est un algorithme qui minimise l'erreur dans les réseaux de neurones.

### 12. Quelle méthode est utilisée pour éviter le surapprentissage (overfitting) ?

- A) Utiliser un modèle plus complexe
- B) Utiliser la validation croisée
- C) Augmenter le nombre d'itérations
- D) Réduire la taille des données

Réponse : B

Explication : La validation croisée permet de tester le modèle sur des ensembles de données différents pour éviter le surapprentissage.

### 13. Qu'est-ce que le surapprentissage (overfitting) ?

- A) Un modèle qui fonctionne trop rapidement
- B) Un modèle qui apprend trop bien les données d'entraînement et échoue à généraliser

- C) Un modèle qui ne fonctionne pas
- D) Un modèle qui utilise trop de mémoire

Réponse : B

Explication : Le surapprentissage se produit lorsque le modèle apprend trop bien les détails des données d'entraînement, rendant la généralisation difficile.

### 14. Qu'est-ce qu'une fonction de coût dans l'apprentissage automatique ?

- A) Une mesure de la complexité du modèle
- B) Une fonction qui mesure l'erreur d'un modèle
- C) Une méthode pour optimiser les performances
- D) Un algorithme d'optimisation

Réponse : B

Explication : La fonction de coût mesure l'écart entre les prédictions du modèle et les résultats attendus.

### 15. Quelle technique est utilisée pour entraîner un réseau de neurones profond ?

- A) Backpropagation
- B) Tri rapide
- C) Tri par sélection
- D) Analyse descendante

Réponse : A

Explication : La rétropropagation (backpropagation) est utilisée pour ajuster les poids dans un réseau de neurones.

### 16. Quelle est la principale différence entre l'apprentissage supervisé et non supervisé ?

- A) L'apprentissage supervisé utilise des étiquettes
- B) L'apprentissage non supervisé est plus rapide
- C) L'apprentissage supervisé ne nécessite pas de données
- D) L'apprentissage non supervisé fonctionne uniquement sur des données numériques

Réponse : A

Explication : L'apprentissage supervisé utilise des données étiquetées pour entraîner un modèle, contrairement à l'apprentissage non supervisé.

### 17. Quel algorithme est souvent utilisé pour le regroupement de données en apprentissage non supervisé ?

- A) Algorithme de Dijkstra
- B) Réseaux de neurones
- C) K-means
- D) Descente de gradient

Réponse : C

Explication : K-means est couramment utilisé pour le regroupement en apprentissage non supervisé.

### 18. Quel est l'objectif de l'IA symbolique ?

- A) Apprendre à partir de données brutes
- B) Représenter la connaissance à l'aide de symboles et de règles logiques
- C) Maximiser les ressources d'un système
- D) Automatiser les tâches répétitives

Réponse : B

Explication : L'IA symbolique se concentre sur la représentation des connaissances à l'aide de symboles et de règles logiques.

### 19. Quelle technique permet de prédire des résultats continus ?

- A) Classification
- B) Régression
- C) Regroupement
- D) Apprentissage par renforcement

Réponse : B

Explication : La régression est utilisée pour prédire des valeurs continues.

### 20. Dans l'apprentissage par renforcement, le but de l'agent est de :

- A) Maximiser une fonction de récompense
- B) Minimiser le coût
- C) Accumuler des données
- D) Suivre les instructions humaines

Réponse : A

Explication : Dans l'apprentissage par renforcement, l'agent apprend à maximiser une fonction de récompense au fil du temps.

### 21. Quel concept désigne la capacité d'un modèle à bien fonctionner sur des données inédites ?

- A) Surapprentissage

B) Généralisation

- C) Régression
- D) Apprentissage incrémental

Réponse : B

Explication : La généralisation fait référence à la capacité d'un modèle à bien s'adapter à de nouvelles données non vues lors de l'entraînement.

### 22. Quels sont les deux types principaux de biais dans l'IA ?

- A) Biais de l'échantillon et biais des données
- B) Biais algorithmique et biais cognitif
- C) Biais de conception et biais de calcul
- D) Biais de sélection et biais statistique

Réponse : B

Explication : Les biais en IA incluent le biais algorithmique et le biais cognitif, qui affectent la prise de décision des systèmes.

### 23. Quel est le rôle des hyperparamètres dans l'entraînement d'un modèle ?

- A) Ils contrôlent la complexité du modèle
- B) Ils sont automatiquement optimisés
- C) Ils définissent les paramètres internes du modèle
- D) Ils ajustent directement les prédictions

Réponse : A

Explication : Les hyperparamètres sont utilisés pour ajuster la complexité et le comportement général d'un modèle.

### 24. Qu'est-ce que l'optimisation dans le contexte de l'IA ?

- A) Minimiser l'erreur du modèle
- B) Maximiser la taille des données
- C) Réduire le nombre d'étiquettes
- D) Augmenter la mémoire de calcul

Réponse : A

Explication : L'optimisation vise à minimiser l'erreur entre les prédictions et les résultats réels.

### 25. Quel est l'inconvénient principal des réseaux de neurones profonds ?

- A) Leur vitesse d'exécution
- B) Le temps de calcul élevé
- C) Leur capacité à gérer des petits jeux de données
- D) La consommation d'énergie

Réponse : B

Explication : Les réseaux de neurones profonds nécessitent beaucoup de temps et de puissance de calcul pour l'entraînement.

### 26. Quelle technique est utilisée pour réduire la dimensionnalité des données ?

- A) Descente de gradient
- B) Régression logistique
- C) Analyse en composantes principales (PCA)
- D) Tri à bulles

Réponse : C

Explication : L'analyse en composantes principales (PCA) est une méthode utilisée pour réduire la dimensionnalité des données.

### 27. Quel est le principal défi dans l'implémentation des systèmes d'IA dans la vie réelle ?

- A) Manque de données
- B) Coût de l'infrastructure
- C) Biais des algorithmes et des données
- D) Vitesse des algorithmes

Réponse : C

Explication : Un défi majeur pour l'IA est de minimiser le biais présent dans les algorithmes et les données utilisées.

### 28. Qu'est-ce que l'IA faible ?

- A) Un système qui peut surpasser l'intelligence humaine
- B) Une IA conçue pour une tâche spécifique
- C) Une IA qui ne fonctionne pas correctement
- D) Un programme basé sur des règles simples

Réponse : B

Explication : L'IA faible se réfère à des systèmes spécialisés dans des tâches spécifiques, comme Siri ou Alexa.

### 29. Dans quel domaine l'IA est-elle couramment utilisée aujourd'hui ?

- A) Conception de circuits électroniques
- B) Recherche d'information et assistants vocaux
- C) Gestion des fichiers informatiques
- D) Compilation de programmes

Réponse : B

Explication : L'IA est largement utilisée dans la recherche d'informations, les assistants vocaux et d'autres applications quotidiennes.

### 30. Quelle est une méthode courante pour améliorer la performance d'un modèle d'IA ?

- A) Réduire le nombre de paramètres
- B) Utiliser plus de données d'entraînement
- C) Diminuer la taille des couches neuronales
- D) Éviter la rétropropagation

Réponse : B

Explication : En augmentant la quantité de données d'entraînement, le modèle peut mieux généraliser et améliorer ses performances.

---

Ces 30 questions couvrent les concepts fondamentaux liés à l'intelligence artificielle et à l'apprentissage automatique.

Voici 30 questions à choix multiples pour la Note 3 (Introduction à l'Apprentissage Automatique - CSI 4506) , avec réponses et explications :

---

### 1. Qu'est-ce que l'apprentissage automatique (machine learning) ?

- A) Un moyen de calculer des données

B) Un processus où les machines apprennent à partir de données sans programmation explicite

C) Un système de stockage d'informations

D) Une méthode pour organiser des fichiers

Réponse : B

Explication : L'apprentissage automatique permet aux machines d'apprendre à partir de données sans être explicitement programmées.

### 2. Quel est l'objectif principal de l'apprentissage automatique ?

A) Résoudre des équations mathématiques

B) Faire des prédictions ou des classifications basées sur des données

C) Créer des images

D) Améliorer la vitesse des processeurs

Réponse : B

Explication : L'objectif principal de l'apprentissage automatique est de créer des modèles qui permettent de faire des prédictions ou des classifications à partir de données.

### 3. Quel type d'apprentissage utilise des données étiquetées ?

A) Apprentissage non supervisé

B) Apprentissage supervisé

C) Apprentissage par renforcement

D) Apprentissage en profondeur

Réponse : B

Explication : L'apprentissage supervisé utilise des données étiquetées pour entraîner un modèle à prédire des résultats.

### 4. Quel algorithme est un exemple d'apprentissage supervisé ?

A) K-means

B) Régression linéaire

C) Algorithme de clustering

D) Réseaux de neurones

Réponse : B

Explication : La régression linéaire est un algorithme d'apprentissage supervisé qui prédit une valeur continue.

### 5. Quel type d'apprentissage est utilisé lorsque les données ne sont pas étiquetées ?

A) Apprentissage supervisé

B) Apprentissage non supervisé

C) Apprentissage semi-supervisé

D) Apprentissage par renforcement

Réponse : B

Explication : L'apprentissage non supervisé est utilisé pour trouver des structures cachées dans des données non étiquetées.

### 6. Qu'est-ce que le surapprentissage (overfitting) ?

A) Un modèle qui s'adapte trop bien aux données d'entraînement, mais échoue sur des données nouvelles

B) Un modèle qui est trop simple pour les données

C) Un modèle qui ne parvient pas à prédire avec précision

D) Un modèle qui fonctionne trop rapidement

Réponse : A

Explication : Le surapprentissage se produit lorsque le modèle mémorise les détails des données d'entraînement, réduisant sa capacité à généraliser sur des données nouvelles.

### 7. Quelle technique peut être utilisée pour éviter le surapprentissage ?

A) Utiliser un modèle plus complexe

B) Collecter plus de données d'entraînement

C) Diminuer la taille des couches neuronales

D) Éviter les données non étiquetées

Réponse : B

Explication : Utiliser un plus grand ensemble de données d'entraînement aide à éviter le surapprentissage en permettant au modèle de mieux généraliser.

### 8. Quelle méthode est couramment utilisée pour évaluer les performances d'un modèle d'apprentissage automatique ?

A) Validation croisée

B) Tri rapide

C) Algorithme de Dijkstra

D) Augmentation de données

Réponse : A

Explication : La validation croisée est une méthode utilisée pour évaluer la performance d'un modèle en le testant sur différents sous-ensembles des données.

### 9. Qu'est-ce qu'un algorithme de clustering ?

A) Un algorithme qui trie les données

B) Un algorithme qui regroupe les données non étiquetées en clusters basés sur des similarités

C) Un algorithme de recherche

D) Un algorithme de tri par fusion

Réponse : B

Explication : Le clustering est une méthode non supervisée qui regroupe des données en fonction de leurs similarités.

### 10. Quel algorithme est un exemple d'apprentissage non supervisé ?

A) Régression logistique

B) K-means

C) Réseaux neuronaux supervisés

D) Régression linéaire

Réponse : B

Explication : K-means est un algorithme de clustering utilisé dans l'apprentissage non supervisé.

#### 11. Quelle est la première étape dans un projet d'apprentissage automatique ?

- A) Entraînement du modèle
- B) Collecte de données
- C) Validation du modèle
- D) Analyse des résultats

Réponse : B

Explication : La première étape dans un projet d'apprentissage automatique est la collecte des données qui seront utilisées pour entraîner et tester le modèle.

#### 12. Qu'est-ce que la sélection de caractéristiques ?

- A) Une méthode de sélection des meilleurs algorithmes
- B) Une technique pour identifier les caractéristiques les plus importantes des données
- C) Une méthode pour ajuster les hyperparamètres
- D) Une méthode pour normaliser les données

Réponse : B

Explication : La sélection de caractéristiques consiste à choisir les variables les plus pertinentes pour améliorer la précision du modèle.

#### 13. Quel type d'apprentissage combine des données étiquetées et non étiquetées ?

- A) Apprentissage supervisé
- B) Apprentissage non supervisé
- C) Apprentissage semi-supervisé
- D) Apprentissage par renforcement

Réponse : C

Explication : L'apprentissage semi-supervisé utilise à la fois des données étiquetées et non étiquetées pour entraîner le modèle.

#### 14. Quel est le but de la validation croisée ?

- A) Accélérer l'entraînement du modèle
- B) Évaluer la performance du modèle sur plusieurs sous-ensembles de données
- C) Réduire la complexité du modèle
- D) Maximiser la taille de l'ensemble d'entraînement

Réponse : B

Explication : La validation croisée teste le modèle sur différents sous-ensembles de données pour évaluer sa capacité à généraliser.

#### 15. Qu'est-ce qu'un modèle de régression ?

- A) Un modèle qui prédit des valeurs continues
- B) Un modèle qui prédit des valeurs discrètes
- C) Un modèle qui trie les données



D) Un modèle qui regroupe des données

Réponse : A

Explication : La régression est utilisée pour prédire des valeurs continues, comme les prix ou les températures.

### 16. Quel est un exemple de modèle de classification ?

A) Régression linéaire

B) K-means

C) Réseaux de neurones

D) Régression logistique

Réponse : D

Explication : La régression logistique est un modèle de classification utilisé pour prédire des catégories.

### 17. Quelle méthode est utilisée pour ajuster les paramètres d'un réseau de neurones pendant l'entraînement ?

A) Algorithme génétique

B) Descente de gradient

C) Tri rapide

D) Algorithme de Bellman-Ford

Réponse : B

Explication : La descente de gradient est utilisée pour minimiser la fonction de coût et ajuster les poids dans un réseau de neurones.

### 18. Quel est un exemple d'algorithme d'apprentissage par renforcement ?

A) Réseaux de neurones

B) K-means

C) Q-learning

D) Régression logistique

Réponse : C

Explication : Q-learning est un algorithme d'apprentissage par renforcement utilisé pour entraîner un agent à maximiser une récompense cumulée.

### 19. Dans quel type de problème l'apprentissage semi-supervisé est-il utile ?

A) Lorsque toutes les données sont étiquetées

B) Lorsque certaines données sont étiquetées et d'autres ne le sont pas

C) Lorsque toutes les données sont non étiquetées

D) Lorsque les données sont en temps réel

Réponse : B

Explication : L'apprentissage semi-supervisé est utile lorsque seule une partie des données est étiquetée.

### 20. Quelle technique est utilisée pour normaliser les données avant l'entraînement ?

A) Division en sous-ensembles

- B) Normalisation min-max
- C) Tri rapide
- D) Augmentation de données

Réponse : B

Explication : La normalisation min-max est une méthode de mise à l'échelle

des données pour qu'elles se situent dans une plage donnée, souvent entre 0 et 1.

### 21. Quel est le rôle de l'algorithme K-nearest neighbors (KNN) ?

- A) Classer les données en fonction des k voisins les plus proches
- B) Trier des valeurs numériques
- C) Réduire la taille des données
- D) Prédire des valeurs continues

Réponse : A

Explication : KNN est un algorithme de classification basé sur la proximité des données par rapport à leurs k voisins les plus proches.

### 22. Quelle est une technique courante pour réduire la dimensionnalité d'un jeu de données ?

- A) Descente de gradient
- B) Tri rapide
- C) PCA (Analyse en Composantes Principales)
- D) Algorithme de Bellman-Ford

Réponse : C

Explication : PCA est une méthode utilisée pour réduire la dimensionnalité en extrayant les caractéristiques les plus importantes d'un jeu de données.

### 23. Quel est un exemple d'algorithme de classification binaire ?

- A) K-means
- B) Régression logistique
- C) Algorithme de clustering
- D) Tri rapide

Réponse : B

Explication : La régression logistique est couramment utilisée pour la classification binaire, c'est-à-dire prédire deux classes distinctes.

### 24. Qu'est-ce que la validation de modèle dans l'apprentissage automatique ?

- A) Vérifier que le modèle est sans erreur
- B) Évaluer la capacité d'un modèle à bien fonctionner sur des données inédites
- C) Réduire la taille des données d'entraînement
- D) Maximiser la complexité du modèle

Réponse : B

Explication : La validation de modèle consiste à évaluer si un modèle peut bien généraliser sur des données qu'il n'a pas vues pendant l'entraînement.

### 25. Quel est un inconvénient des réseaux de neurones profonds ?

- A) Leur lenteur
- B) Leur besoin en grandes quantités de données
- C) Leur faible précision
- D) Leur manque de capacité de mémorisation

Réponse : B

Explication : Les réseaux de neurones profonds nécessitent de grandes quantités de données pour être efficaces et bien entraînés.

### 26. Quelle technique permet de lutter contre le surapprentissage dans les réseaux neuronaux ?

- A) Augmentation de la taille des couches
- B) Utilisation de la régularisation L2 (ridge)
- C) Réduire la quantité de données
- D) Diminuer le nombre d'itérations

Réponse : B

Explication : La régularisation L2 est une méthode utilisée pour réduire la complexité d'un modèle et prévenir le surapprentissage.

### 27. Qu'est-ce que la fonction de coût dans un modèle d'apprentissage automatique ?

- A) Un algorithme de tri
- B) Une fonction qui mesure l'erreur du modèle
- C) Une méthode d'entraînement des réseaux de neurones
- D) Une méthode de sélection des caractéristiques

Réponse : B

Explication : La fonction de coût mesure la différence entre les prédictions du modèle et les résultats attendus, et guide l'optimisation.

### 28. Quel est l'objectif de l'apprentissage supervisé ?

- A) Trouver des relations cachées dans des données non étiquetées
- B) Apprendre à partir de données étiquetées pour faire des prédictions ou des classifications
- C) Minimiser la complexité des algorithmes
- D) Classer des données en clusters

Réponse : B

Explication : L'apprentissage supervisé utilise des données étiquetées pour entraîner un modèle qui pourra prédire ou classer des données similaires.

### 29. Dans l'apprentissage par renforcement, comment l'agent apprend-il ?

- A) En suivant des règles préétablies
- B) En recevant des récompenses ou punitions basées sur ses actions
- C) En analysant des données étiquetées
- D) En utilisant des algorithmes non supervisés

Réponse : B

Explication : L'agent apprend à prendre des décisions en fonction des récompenses ou punitions qu'il reçoit pour ses actions.

### 30. Quel est le principal avantage de l'utilisation de réseaux neuronaux profonds ?

- A) Ils sont faciles à interpréter
- B) Ils peuvent modéliser des relations complexes dans les données
- C) Ils nécessitent peu de données
- D) Ils consomment moins de mémoire

Réponse : B

Explication : Les réseaux neuronaux profonds sont capables de capturer et modéliser des relations complexes dans de grandes quantités de données.

---

Ces questions couvrent les concepts clés de l'apprentissage automatique et offrent des explications pour mieux comprendre chaque réponse.

Here are 30 multiple-choice questions in French, based on the key concepts from Note 4: Modèles linéaires, entraînement :

---

1. Qu'est-ce qu'un modèle linéaire ?

- a) Un modèle qui s'adapte à toutes les formes de données
- b) Un modèle qui prédit des valeurs continues avec une relation linéaire
- c) Un modèle de classification non linéaire

Réponse : b)

Explication : Un modèle linéaire prédit une variable dépendante en fonction d'une ou plusieurs variables indépendantes avec une relation linéaire.

---

2. Quelle est la fonction de coût couramment utilisée pour la régression linéaire ?

- a) Fonction logarithmique
- b) Entropie croisée
- c) Erreur quadratique moyenne (MSE)

Réponse : c)

Explication : L'erreur quadratique moyenne mesure la différence moyenne au carré entre les valeurs prédites et les valeurs réelles.

---

3. Qu'est-ce que le gradient dans le contexte de la descente de gradient ?

- a) La dérivée partielle de la fonction de coût par rapport aux paramètres

- b) Une constante qui ajuste la vitesse de convergence
- c) Une erreur aléatoire introduite dans le modèle

Réponse : a)

Explication : Le gradient est utilisé pour ajuster les poids du modèle afin de minimiser la fonction de coût.

---

4. Quel est l'objectif principal de la descente de gradient ?

- a) Maximiser la fonction de coût
- b) Trouver un minimum global de la fonction de coût
- c) Augmenter les prédictions incorrectes

Réponse : b)

Explication : La descente de gradient cherche à minimiser la fonction de coût en ajustant les poids itérativement.

---

5. Quel hyperparamètre contrôle la taille des étapes dans la descente de gradient ?

- a) La régularisation
- b) Le taux d'apprentissage (learning rate)
- c) Le nombre d'époques

Réponse : b)

Explication : Le taux d'apprentissage contrôle la taille des ajustements des poids à chaque itération.

---

6. Quelle méthode est souvent utilisée pour éviter le surapprentissage (overfitting) dans un modèle linéaire ?

- a) Ajouter plus de couches
- b) Utiliser une régularisation L2
- c) Réduire le nombre de données

Réponse : b)

Explication : La régularisation L2 pénalise les poids excessivement grands et aide à généraliser le modèle.

---

7. Quel est l'effet d'une valeur de taux d'apprentissage trop élevée ?

- a) Convergence rapide vers un optimum global
- b) Non-convergence ou oscillations autour du minimum
- c) Surapprentissage immédiat

Réponse : b)

Explication : Un taux d'apprentissage trop élevé peut provoquer des oscillations ou empêcher la convergence.

---

8. Dans la régression linéaire simple, quelle est la forme générale de la fonction prédictive ?

- a)  $y = b_0 + b_1 x$
- b)  $y = b_0 + b_1 x^2$
- c)  $y = b_0 + e^x$

Réponse : a)

Explication : La régression linéaire simple est représentée par une droite où  $y$  dépend de  $x$  avec une pente  $b_1$ .

---

9. Qu'est-ce qu'une variable indépendante dans un modèle linéaire ?

- a) La variable cible
- b) La variable prédite
- c) La variable explicative

Réponse : c)

Explication : Les variables indépendantes sont les variables explicatives utilisées pour prédire la variable dépendante.

---

10. Quel est le rôle du biais (intercept) dans un modèle linéaire ?

- a) Ajuster la pente
- b) Introduire un décalage dans les prédictions
- c) Normaliser les données

Réponse : b)

Explication : Le biais permet au modèle de faire des prédictions qui ne passent pas forcément par l'origine.

---

11. Quel est l'effet d'un taux d'apprentissage trop bas dans la descente de gradient ?

- a) Convergence lente
- b) Surapprentissage
- c) Oscillation autour du minimum

Réponse : a)

Explication : Un taux d'apprentissage trop bas ralentit la convergence vers le minimum de la fonction de coût.

---

12. Que signifie "multicolinéarité" dans un modèle linéaire ?

- a) Quand les variables indépendantes sont faiblement corrélées
- b) Quand les variables indépendantes sont fortement corrélées entre elles
- c) Quand les variables dépendantes sont fortement corrélées

Réponse : b)

Explication : La multicolinéarité se produit lorsque les variables explicatives sont fortement corrélées, ce qui peut fausser les estimations du modèle.

---

13. Qu'est-ce que la validation croisée (cross-validation) ?

- a) Une technique pour diviser les données en plusieurs sous-ensembles
- b) Un moyen de tester la stabilité d'un modèle
- c) Les deux

Réponse : c)

Explication : La validation croisée aide à évaluer les performances d'un modèle en testant sur différents sous-ensembles de données.

---

14. Que fait la méthode "batch gradient descent" ?

- a) Utilise une partie aléatoire des données à chaque itération
- b) Utilise l'ensemble des données pour chaque mise à jour des poids
- c) Ne met à jour les poids qu'à la fin de l'entraînement

Réponse : b)

Explication : La descente de gradient par lot (batch) utilise toutes les données pour chaque mise à jour.

---

15. Pourquoi l'erreur quadratique moyenne est-elle utilisée comme fonction de coût ?

- a) Elle minimise les petites erreurs
- b) Elle punit les grandes erreurs de manière plus sévère
- c) Elle ne prend pas en compte les erreurs

Réponse : b)

Explication : L'erreur quadratique moyenne amplifie les grandes erreurs, ce qui pousse le modèle à les minimiser.

---

16. Quel est l'avantage principal de la descente de gradient stochastique (SGD) ?

- a) Elle est plus rapide pour les grands ensembles de données
- b) Elle garantit une meilleure solution

c) Elle évite les oscillations

Réponse : a)

Explication : La SGD met à jour les poids après chaque échantillon, ce qui la rend plus rapide pour les grands ensembles de données.

---

17. Comment un modèle linéaire se comporte-t-il en présence d'un surapprentissage ?

a) Il généralise bien sur les nouvelles données

b) Il s'adapte trop aux données d'entraînement, mais échoue sur les nouvelles données

c) Il échoue sur les données d'entraînement

Réponse : b)

Explication : Un modèle en surapprentissage s'ajuste trop bien aux données d'entraînement et manque de généralisation.

---

18. Qu'est-ce qu'un résidu dans un modèle de régression ?

a) Une erreur systématique

b) La différence entre la valeur observée et la valeur prédite

c) Une erreur dans les poids

Réponse : b)

Explication : Les résidus représentent les erreurs individuelles entre les valeurs réelles et prédites.

---

19. Que signifie une pente de régression linéaire égale à zéro ?

a) Aucune relation entre les variables

b) Une forte relation positive

c) Une forte relation négative

Réponse : a)

Explication : Une pente de zéro indique que les variables explicatives n'ont pas d'impact sur la variable dépendante.

---

20. Quelle technique est utilisée pour normaliser les données avant d'entraîner un modèle linéaire ?

a) Le scaling min-max

b) La dérivation

c) L'algorithme de Gauss-Newton

Réponse : a)



Explication : Le scaling min-max ajuste les données pour qu'elles soient dans une plage comprise entre 0 et 1.

---

21. Qu'est-ce qu'un modèle sous-entraîné (underfitted) ?

- a) Un modèle trop complexe
- b) Un modèle qui ne capte pas suffisamment la structure des données
- c) Un modèle qui s'adapte parfaitement aux données

Réponse : b)

Explication : Un modèle sous-entraîné est trop simple et échoue à capturer les tendances des données.

---

22. Quel est l'effet de la régularisation L1 sur les poids du modèle ?

- a) Elle force certains poids à devenir exactement nuls
- b) Elle réduit les poids de manière uniforme
- c) Elle augmente la complexité du modèle

Réponse : a)

Explication : La régularisation L1 favorise la parcimonie en éliminant certains poids, ce qui peut conduire à un modèle plus simple.

---

23. Quel est l'objectif de la régularisation dans un modèle linéaire ?

- a) Minimiser l'erreur quadratique
- b) Réduire la complexité du modèle pour éviter le surapprentissage
- c) Augmenter les poids pour de meilleures prédictions

Réponse : b)

Explication : La régularisation pénalise les grands poids pour éviter que le modèle ne s'ajuste trop aux données d'entraînement.

---

24. Quelle est la principale différence entre la régularisation L1 et L2 ?

- a) L1 encourage des poids petits, mais non nuls, tandis que L2 force des poids nuls
- b) L1 force certains poids à zéro, tandis que L2 réduit les poids sans les annuler
- c) L1 et L2 sont identiques

Réponse : b)

Explication : La régularisation L1 peut annuler certains poids, tandis que L2 les réduit proportionnellement sans les forcer à zéro.

---

25. Dans quel cas est-il approprié d'utiliser la régression linéaire ?

- a) Lorsque les données présentent une relation non linéaire
- b) Lorsque la relation entre les variables est linéaire
- c) Lorsque les données sont catégorielles

Réponse : b)

Explication : La régression linéaire est adaptée lorsque la relation entre les variables dépendantes et indépendantes est linéaire.

---

26. Pourquoi utilise-t-on un terme de biais (intercept) dans un modèle linéaire ?

- a) Pour améliorer la précision du modèle
- b) Pour ajuster la prédiction lorsque toutes les variables explicatives sont nulles
- c) Pour réduire l'impact des erreurs

Réponse : b)

Explication : Le terme de biais ajuste la prédiction du modèle lorsque toutes les variables indépendantes valent zéro.

---

27. Quel est l'effet d'une grande variance des paramètres dans un modèle linéaire ?

- a) Le modèle généralise bien aux nouvelles données
- b) Le modèle surapprend les données d'entraînement
- c) Le modèle sous-apprend et manque de précision

Réponse : b)

Explication : Une grande variance peut entraîner un surapprentissage, où le modèle s'ajuste trop aux données d'entraînement, mais échoue sur de nouvelles données.

---

28. Quelle est la principale hypothèse de la régression linéaire ?

- a) La relation entre les variables est non linéaire
- b) Les variables explicatives sont indépendantes
- c) La relation entre la variable dépendante et les variables explicatives est linéaire

Réponse : c)

Explication : La régression linéaire suppose une relation linéaire entre la variable dépendante et les variables explicatives.

---

29. Comment peut-on interpréter les coefficients d'un modèle de régression linéaire ?

- a) Ils indiquent l'importance relative de chaque variable explicative
- b) Ils sont uniquement utilisés pour normaliser les données
- c) Ils ne sont pas interprétables

Réponse : a)

Explication : Les coefficients indiquent comment une unité de changement dans une variable explicative affecte la variable dépendante.

---

30. Quel est le rôle du test d'hypothèse dans la régression linéaire ?

- a) Déterminer si les coefficients sont significativement différents de zéro
- b) Réduire les erreurs dans les données
- c) Augmenter la précision des prédictions

Réponse : a)

Explication : Un test d'hypothèse permet de vérifier si les coefficients du modèle sont significativement différents de zéro, ce qui indique l'importance de chaque variable explicative dans le modèle.

---

Voici 30 questions à choix multiples en français, basées sur les concepts clés de Note 5 :  
Modèles linéaires, entraînement .

---

1. Quel est l'objectif principal de la régularisation dans les modèles linéaires ?

- a) Réduire le temps d'entraînement
- b) Réduire le surapprentissage (overfitting)
- c) Augmenter l'erreur d'entraînement

Réponse : b)

Explication : La régularisation aide à contrôler la complexité du modèle en pénalisant les grands poids pour éviter un surajustement aux données d'entraînement.

---

2. Quelle méthode de régularisation applique une pénalité proportionnelle à la somme des carrés des coefficients ?

- a) Régularisation L1
- b) Régularisation L2
- c) Normalisation

Réponse : b)

Explication : La régularisation L2 applique une pénalité égale à la somme des carrés des coefficients, encourageant ainsi des poids plus petits.

---

3. Quelle technique de régularisation force certains coefficients à devenir exactement égaux à zéro ?

- a) Lasso (L1)
- b) Ridge (L2)
- c) Normalisation des données

Réponse : a)

Explication : La méthode Lasso (L1) encourage la parcimonie en éliminant certains coefficients, ce qui simplifie le modèle.

---

4. Dans le contexte de la validation croisée, quel est le but d'un "k-fold cross-validation" ?

- a) Tester le modèle sur l'ensemble des données d'entraînement
- b) Diviser les données en k sous-ensembles pour valider le modèle
- c) Minimiser la fonction de coût sur les données d'entraînement

Réponse : b)

Explication : La validation croisée en k-fold divise les données en k sous-ensembles et teste le modèle sur chacun d'eux pour obtenir une meilleure estimation de sa performance.

---

5. Comment appelle-t-on un modèle qui apprend trop bien les données d'entraînement et échoue sur les nouvelles données ?

- a) Un modèle sous-entraîné
- b) Un modèle en surapprentissage (overfitting)
- c) Un modèle bien ajusté

Réponse : b)

Explication : Le surapprentissage se produit lorsque le modèle est trop ajusté aux particularités des données d'entraînement et échoue à généraliser sur de nouvelles données.

---

6. Quelle technique peut être utilisée pour éviter le surapprentissage dans un modèle linéaire ?

- a) Utiliser un modèle plus complexe
- b) Réduire la taille du jeu d'entraînement
- c) Utiliser la régularisation L2

Réponse : c)

Explication : La régularisation L2 limite les coefficients du modèle pour éviter qu'il ne s'ajuste trop aux données d'entraînement.

---

7. Quel est l'avantage principal de la descente de gradient stochastique (SGD) par rapport à la descente de gradient par lot (batch) ?

- a) Elle est plus rapide sur les grands ensembles de données
- b) Elle nécessite plus de calculs à chaque étape
- c) Elle garantit une meilleure convergence

Réponse : a)

Explication : La descente de gradient stochastique met à jour les poids après chaque exemple, ce qui la rend plus rapide pour les grands ensembles de données.

---

8. Quelle est la principale différence entre la descente de gradient par mini-lots (mini-batch) et la descente de gradient par lot (batch) ?

- a) La mini-batch utilise toutes les données à chaque itération
- b) La mini-batch utilise un sous-ensemble des données à chaque itération
- c) La mini-batch est plus lente

Réponse : b)

Explication : La descente de gradient par mini-lots divise les données en petits sous-ensembles, combinant les avantages de la descente de gradient par lot et de la descente de gradient stochastique.

---

9. Quel est le rôle du taux d'apprentissage dans la descente de gradient ?

- a) Contrôler la direction du gradient
- b) Déterminer la taille des étapes pour mettre à jour les poids
- c) Réduire l'erreur quadratique moyenne

Réponse : b)

Explication : Le taux d'apprentissage détermine la taille des ajustements des poids à chaque itération dans la descente de gradient.

---

10. Que se passe-t-il si le taux d'apprentissage est trop élevé ?

- a) Le modèle convergera rapidement
- b) Le modèle oscillera et ne convergera pas
- c) Le modèle ne changera pas ses poids

Réponse : b)

Explication : Un taux d'apprentissage trop élevé peut provoquer des oscillations autour du minimum, empêchant la convergence.

---

11. Que représente l'erreur quadratique moyenne (MSE) dans un modèle de régression linéaire ?

- a) L'erreur absolue entre les valeurs réelles et prédites

- b) La moyenne des erreurs au carré entre les valeurs réelles et prédites
- c) Le biais dans les prédictions

Réponse : b)

Explication : L'erreur quadratique moyenne mesure la différence moyenne au carré entre les valeurs prédites par le modèle et les valeurs réelles.

---

12. Pourquoi la normalisation des données est-elle importante avant l'entraînement d'un modèle linéaire ?

- a) Pour augmenter la complexité du modèle
- b) Pour garantir que les variables aient la même échelle
- c) Pour éviter le surapprentissage

Réponse : b)

Explication : La normalisation met les variables à la même échelle, ce qui permet d'améliorer la convergence lors de la descente de gradient.

---

13. Qu'est-ce qu'un modèle sous-entraîné (underfitting) ?

- a) Un modèle qui est trop complexe et s'adapte parfaitement aux données
- b) Un modèle qui ne capture pas suffisamment la structure des données
- c) Un modèle avec un faible biais

Réponse : b)

Explication : Un modèle sous-entraîné est trop simple et ne parvient pas à capturer les tendances dans les données.

---

14. Qu'est-ce que la "dérivée partielle" dans la descente de gradient ?

- a) La somme de toutes les erreurs
- b) La pente de la fonction de coût par rapport à un poids spécifique
- c) Le taux d'apprentissage

Réponse : b)

Explication : La dérivée partielle indique comment la fonction de coût change par rapport à un paramètre particulier, aidant à ajuster les poids.

---

15. Quelle est la forme générale de la fonction prédictive dans un modèle de régression linéaire multiple ?

- a)  $y = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n$
- b)  $y = b_0 + e^x$
- c)  $y = b_1 + x_1^2$

Réponse : a)

Explication : Dans la régression linéaire multiple, la prédiction est une combinaison linéaire des variables explicatives pondérées par leurs coefficients respectifs.

---

16. Quel est l'impact de la multicollinéarité dans un modèle de régression linéaire ?

- a) Elle améliore la précision des prédictions
- b) Elle fausse les estimations des coefficients
- c) Elle réduit l'erreur quadratique moyenne

Réponse : b)

Explication : La multicollinéarité rend difficile l'estimation précise des coefficients car les variables explicatives sont fortement corrélées entre elles.

---

17. Que se passe-t-il lorsque le modèle n'a pas suffisamment de données pour l'entraînement ?

- a) Le modèle est plus susceptible de surapprendre
- b) Le modèle sera toujours performant
- c) Le modèle aura une meilleure généralisation

Réponse : a)

Explication : Un manque de données augmente le risque de surapprentissage, car le modèle peut s'ajuster trop aux particularités des données d'entraînement.

---

18. Quel est l'effet d'un taux d'apprentissage trop faible dans la descente de gradient ?

- a) Convergence très lente
- b) Convergence rapide
- c) Surapprentissage immédiat

Réponse : a)

Explication : Un taux d'apprentissage trop faible entraîne une convergence lente, car les ajustements des poids sont trop petits.

---

19. Pourquoi la descente de gradient stochastique (SGD) est-elle souvent utilisée pour les grands ensembles de données ?

- a) Elle converge plus rapidement en utilisant moins de données à chaque étape
- b) Elle est plus précise que la descente de gradient par lot

descente de gradient par lot

- c) Elle utilise toutes les données à chaque mise à jour

Réponse : a)

Explication : La SGD utilise un échantillon de données à chaque itération, ce qui accélère la convergence pour les grands ensembles de données.

---

20. Que mesure le coefficient de détermination  $R^2$  dans un modèle de régression ?

- a) La proportion de la variance des données expliquée par le modèle
- b) La somme des erreurs au carré
- c) Le taux d'apprentissage optimal

Réponse : a)

Explication :  $R^2$  mesure la proportion de la variance de la variable dépendante qui est expliquée par les variables explicatives du modèle.

---

21. Qu'est-ce qu'un hyperparamètre dans un modèle de machine learning ?

- a) Un paramètre appris pendant l'entraînement
- b) Un paramètre fixé avant l'entraînement
- c) Un paramètre optimisé automatiquement

Réponse : b)

Explication : Les hyperparamètres sont fixés avant l'entraînement et influencent la performance du modèle, comme le taux d'apprentissage.

---

22. Quelle est la fonction de coût la plus couramment utilisée dans les problèmes de régression linéaire ?

- a) Entropie croisée
- b) Erreur quadratique moyenne (MSE)
- c) Logarithme de vraisemblance

Réponse : b)

Explication : L'erreur quadratique moyenne (MSE) est utilisée pour mesurer la différence entre les prédictions et les valeurs réelles dans la régression linéaire.

---

23. Qu'est-ce qu'un modèle bien ajusté (good fit) ?

- a) Un modèle qui prédit parfaitement toutes les valeurs
- b) Un modèle qui généralise bien sur des données nouvelles
- c) Un modèle avec une très faible erreur d'entraînement

Réponse : b)

Explication : Un modèle bien ajusté est capable de faire de bonnes prédictions sur de nouvelles données sans être surajusté aux données d'entraînement.



---

24. Que fait la régularisation L2 ?

- a) Pénalise les grands coefficients pour éviter le surapprentissage
- b) Augmente la variance des poids pour améliorer l'ajustement
- c) Supprime les données redondantes

Réponse : a)

Explication : La régularisation L2 pénalise les grands coefficients, ce qui aide à éviter que le modèle ne devienne trop complexe et qu'il ne surapprenne.

---

25. Que signifie "erreur de biais" dans un modèle linéaire ?

- a) Une erreur due à un modèle trop complexe
- b) Une erreur due à un modèle trop simple
- c) Une erreur introduite par une mauvaise optimisation

Réponse : b)

Explication : L'erreur de biais provient d'un modèle trop simple qui ne capture pas les relations complexes dans les données.

---

26. Que fait un modèle avec une faible erreur de variance ?

- a) Il surapprend les données d'entraînement
- b) Il généralise bien sur de nouvelles données
- c) Il ajuste parfaitement toutes les données d'entraînement

Réponse : b)

Explication : Un modèle avec une faible variance est capable de généraliser efficacement sur des données non vues sans être trop ajusté aux données d'entraînement.

---

27. Quel est le principal problème lié à la multicolinéarité dans les modèles de régression ?

- a) Elle augmente l'erreur quadratique moyenne
- b) Elle empêche la convergence du modèle
- c) Elle rend difficile l'interprétation des coefficients

Réponse : c)

Explication : La multicolinéarité fausse l'estimation des coefficients et rend difficile l'interprétation de l'importance de chaque variable explicative.

---

28. Que fait le "dropout" dans l'entraînement des modèles linéaires ?

- a) Il élimine des caractéristiques pendant l'entraînement

- b) Il ajuste le taux d'apprentissage automatiquement
- c) Il force l'interruption de l'entraînement après une période définie

Réponse : a)

Explication : Le "dropout" est une technique utilisée pour éviter le surapprentissage en éliminant aléatoirement des neurones pendant l'entraînement.

---

29. Pourquoi utiliser un ensemble de test distinct pour évaluer un modèle ?

- a) Pour ajuster les hyperparamètres du modèle
- b) Pour évaluer la capacité du modèle à généraliser sur des données non vues
- c) Pour optimiser les coefficients du modèle

Réponse : b)

Explication : Un ensemble de test distinct est utilisé pour évaluer la performance du modèle sur des données nouvelles, non utilisées pendant l'entraînement.

---

30. Quelle technique est utilisée pour sélectionner les meilleures caractéristiques dans un modèle linéaire ?

- a) Régression Ridge
- b) Sélection par validation croisée
- c) Lasso (L1)

Réponse : c)

Explication : La méthode Lasso (régularisation L1) peut être utilisée pour sélectionner les meilleures caractéristiques en forçant certains coefficients à devenir nuls.

---

Ces questions permettent de réviser les concepts importants comme la régularisation, la descente de gradient, la validation croisée, et l'évaluation des performances d'un modèle.

Voici 30 questions à choix multiples en français, basées sur les concepts clés de Note 6 : Régression logistique .

---

1. Quel est le but principal de la régression logistique ?

- a) Prédire des valeurs continues
- b) Classifier des données en catégories
- c) Optimiser des séries temporelles

Réponse : b)

Explication : La régression logistique est utilisée pour la classification, c'est-à-dire pour prédire l'appartenance à des classes binaires ou multiples.

---

2. Quelle fonction est utilisée dans la régression logistique pour transformer la sortie en probabilité ?

- a) Fonction exponentielle
- b) Fonction logarithmique
- c) Fonction sigmoïde

Réponse : c)

Explication : La fonction sigmoïde transforme une valeur réelle en une probabilité comprise entre 0 et 1, ce qui est utile pour la classification.

---

3. Quelle est la formule de la fonction sigmoïde ?

- a)  $\frac{1}{1+e^{-z}}$
- b)  $z = mx + b$
- c)  $z = e^x$

Réponse : a)

Explication : La fonction sigmoïde est définie par  $\frac{1}{1+e^{-z}}$ , où  $z$  est une combinaison linéaire des caractéristiques d'entrée.

---

4. Quel type de problème est mieux adapté à la régression logistique ?

- a) Régression linéaire
- b) Classification binaire
- c) Clustering

Réponse : b)

Explication : La régression logistique est utilisée pour les problèmes de classification binaire où la sortie est soit 0 soit 1.

---

5. Que représente la sortie de la fonction sigmoïde dans la régression logistique ?

- a) Une valeur réelle
- b) Une probabilité
- c) Une somme des carrés

Réponse : b)

Explication : La sortie de la fonction sigmoïde est une probabilité qui indique la probabilité d'appartenance à une classe donnée.

---

6. Quelle fonction de coût est couramment utilisée dans la régression logistique ?

- a) Erreur quadratique moyenne
- b) Entropie croisée
- c) Perte huber

Réponse : b)

Explication : L'entropie croisée est utilisée pour mesurer l'erreur dans les problèmes de classification en comparant les probabilités prédites aux vraies étiquettes.

---

7. Que signifie une probabilité de 0,8 dans la régression logistique ?

- a) La classe prédit 0 avec 80% de confiance
- b) La classe prédit 1 avec 80% de confiance
- c) Aucune conclusion ne peut être tirée

Réponse : b)

Explication : Une probabilité de 0,8 indique une confiance de 80% que l'élément appartient à la classe 1.

---

8. Quel critère est utilisé pour prendre une décision de classification en régression logistique ?

- a) Seuil de probabilité
- b) Minimisation de la variance
- c) Maximisation de la dérivée

Réponse : a)

Explication : Un seuil de probabilité (généralement fixé à 0,5) est utilisé pour décider si la sortie est classée dans la classe 0 ou 1.

---

9. Qu'est-ce que le log-odds (logarithme des cotes) dans la régression logistique ?

- a) Le logarithme de la sortie de la fonction sigmoïde
- b) Le logarithme du ratio entre les probabilités d'appartenance aux deux classes
- c) Le logarithme de la somme des erreurs

Réponse : b)

Explication : Le log-odds est le logarithme du rapport entre la probabilité d'appartenance à la classe 1 et la probabilité d'appartenance à la classe 0.

---

10. Comment s'interprète un coefficient positif dans un modèle de régression logistique ?

- a) Il diminue les probabilités d'appartenance à la classe 1
- b) Il augmente les probabilités d'appartenance à la classe 1
- c) Il n'a aucun effet sur la classification

Réponse : b)

Explication : Un coefficient positif indique que, lorsque la variable correspondante augmente, la probabilité d'appartenir à la classe 1 augmente.

---

11. Que signifie une perte d'entropie croisée élevée ?

- a) Le modèle fait de bonnes prédictions
- b) Le modèle a de mauvaises performances
- c) La fonction de coût est optimisée

Réponse : b)

Explication : Une perte d'entropie croisée élevée signifie que le modèle ne prédit pas correctement les classes des données d'entraînement.

---

12. Pourquoi la régression logistique ne convient-elle pas aux problèmes de régression ?

- a) Elle ne peut pas prédire des valeurs continues
- b) Elle est trop lente
- c) Elle a besoin de grands ensembles de données

Réponse : a)

Explication : La régression logistique est conçue pour les problèmes de classification, pas pour prédire des valeurs continues comme dans la régression linéaire.

---

13. Quel est l'effet d'une augmentation du coefficient régularisation ( $\lambda$ ) dans la régression logistique régularisée ?

- a) Augmentation du surapprentissage
- b) Réduction des coefficients des variables explicatives
- c) Augmentation des prédictions incorrectes

Réponse : b)

Explication : La régularisation pénalise les grands coefficients, ce qui aide à réduire le surapprentissage en simplifiant le modèle.

---

14. Qu'est-ce que la régularisation L2 applique dans la régression logistique ?

- a) Une pénalité basée sur la somme des carrés des coefficients
- b) Une pénalité sur la somme des valeurs absolues des coefficients
- c) Une modification du taux d'apprentissage

Réponse : a)

Explication : La régularisation L2 pénalise la somme des carrés des coefficients, encourageant des coefficients plus petits pour éviter le surapprentissage.

---

15. Qu'est-ce qu'un modèle sous-entraîné (underfitting) dans la régression logistique ?

- a) Un modèle qui ne capture pas suffisamment de la structure des données
- b) Un modèle qui s'adapte parfaitement aux données d'entraînement
- c) Un modèle avec une faible régularisation

Réponse : a)

Explication : Un modèle sous-entraîné ne capture pas suffisamment de la complexité des données, ce qui conduit à de mauvaises performances aussi bien sur l'entraînement que sur les tests.

---

16. Que fait la descente de gradient dans la régression logistique ?

- a) Elle maximise la fonction de coût
- b) Elle minimise la fonction de coût en ajustant les coefficients
- c) Elle ajuste uniquement le taux d'apprentissage

Réponse : b)

Explication : La descente de gradient est utilisée pour minimiser la fonction de coût en ajustant les coefficients du modèle.

---

17. Quelle est la différence entre la régression logistique et la régression linéaire ?

- a) La régression logistique prédit des valeurs continues, tandis que la régression linéaire prédit des classes
- b) La régression logistique prédit des probabilités pour des classes, tandis que la régression linéaire prédit des valeurs continues
- c) Elles sont identiques, sauf dans leur fonction de coût

Réponse : b)

Explication : La régression logistique est utilisée pour la classification, alors que la régression linéaire est utilisée pour prédire des valeurs continues.

---

18. Qu'est-ce que le terme "multicolinéarité" dans la régression logistique ?

- a) L'absence de corrélation entre les variables explicatives
- b) Une forte corrélation entre deux ou plusieurs variables explicatives
- c) Une relation non linéaire entre la variable dépendante et les variables explicatives

Réponse : b)

Explication : La multicolinéarité survient lorsque deux ou plusieurs variables explicatives sont fortement corrélées, ce qui peut rendre difficile l'interprétation des coefficients.

---

19. Quel est le rôle de la fonction softmax dans la régression logistique multinomiale ?

- a) Maximiser les cotes d'une classe particulière
- b) Transformer des scores en probabilités pour plusieurs classes
- c) Normaliser les caractéristiques

Réponse : b)

Explication : La fonction softmax est utilisée dans la régression logistique multinomiale pour transformer les scores de chaque classe en probabilités, garantissant que la somme est égale à 1

.

---

20. Quelle métrique est souvent utilisée pour évaluer la performance d'un modèle de régression logistique binaire ?

- a) L'erreur quadratique moyenne
- b) L'aire sous la courbe ROC (AUC)
- c) Le coefficient de détermination ( $R^2$ )

Réponse : b)

Explication : L'aire sous la courbe ROC (AUC) est une métrique couramment utilisée pour évaluer la performance des modèles de classification binaire.

---

21. Qu'est-ce que la régularisation L1 dans la régression logistique ?

- a) Une méthode pour réduire les erreurs quadratiques
- b) Une technique pour supprimer certains coefficients en les forçant à zéro
- c) Une technique pour augmenter le surapprentissage

Réponse : b)

Explication : La régularisation L1 encourage la parcimonie en supprimant certains coefficients, ce qui peut améliorer la simplicité du modèle.

---

22. Quel est l'avantage d'utiliser la régularisation dans un modèle de régression logistique ?

- a) Augmenter la précision sur les données d'entraînement
- b) Réduire le surapprentissage en pénalisant les grands coefficients
- c) Maximiser la probabilité des résultats incorrects

Réponse : b)

Explication : La régularisation limite la complexité du modèle en pénalisant les grands coefficients, ce qui aide à réduire le surapprentissage.

---

23. Quelle est l'hypothèse principale de la régression logistique ?

- a) Il existe une relation linéaire entre les variables explicatives et la sortie
- b) La sortie est une variable continue
- c) La relation entre les variables explicatives et la sortie est modélisée par une fonction sigmoïde

Réponse : c)

Explication : La régression logistique suppose que la relation entre les variables explicatives et la sortie peut être modélisée par une fonction sigmoïde.

---

24. Comment interpréter un coefficient nul dans un modèle de régression logistique ?

- a) La variable correspondante n'a aucun effet sur la probabilité de classification
- b) La variable augmente la probabilité de la classe 1
- c) La variable augmente la probabilité de la classe 0

Réponse : a)

Explication : Un coefficient nul signifie que la variable correspondante n'a aucun effet sur la probabilité que l'échantillon soit classé dans une certaine catégorie.

---

25. Que fait un modèle en surapprentissage dans la régression logistique ?

- a) Il généralise bien sur de nouvelles données
- b) Il s'ajuste parfaitement aux données d'entraînement mais échoue sur les nouvelles données

c) Il sous-estime les classes

Réponse : b)

Explication : Un modèle en surapprentissage s'adapte trop aux données d'entraînement, ce qui entraîne de mauvaises performances sur de nouvelles données.

---

26. Quelle est l'utilité d'une courbe ROC dans l'évaluation de la régression logistique ?

- a) Evaluer les résidus
- b) Mesurer la performance de classification pour différents seuils
- c) Calculer la perte de log-vraisemblance

Réponse : b)

Explication : Une courbe ROC montre la performance du modèle pour différents seuils de classification, en traçant la sensibilité contre la spécificité.

---



27. Que signifie un AUC de 0,5 dans une courbe ROC ?

- a) Le modèle prédit parfaitement
- b) Le modèle fait des prédictions aléatoires
- c) Le modèle ne parvient pas à classer correctement

Réponse : b)

Explication : Un AUC de 0,5 indique que le modèle ne fait pas mieux qu'un tirage au sort aléatoire dans ses prédictions.

---

28. Comment une régression logistique multinomiale diffère-t-elle d'une régression logistique binaire ?

- a) Elle prédit des classes multiples au lieu de deux
- b) Elle utilise une fonction de coût différente
- c) Elle est basée sur une régression linéaire

Réponse : a)

Explication : La régression logistique multinomiale permet de prédire plus de deux classes, alors que la régression logistique binaire est limitée à deux classes.

---

29. Comment choisir le seuil optimal pour un modèle de régression logistique ?

- a) En maximisant la précision
- b) En utilisant la courbe ROC pour évaluer différents seuils
- c) En augmentant les coefficients

Réponse : b)

Explication : Le seuil optimal peut être choisi en utilisant la courbe ROC pour trouver le meilleur compromis entre sensibilité et spécificité.

---

30. Qu'est-ce que l'overfitting dans un modèle de régression logistique ?

- a) Lorsque le modèle généralise bien
- b) Lorsque le modèle s'ajuste trop aux données d'entraînement et échoue à généraliser
- c) Lorsque le modèle réduit l'erreur sur les données de test

Réponse : b)

Explication : L'overfitting se produit lorsque le modèle apprend trop bien les particularités des données d'entraînement, au détriment de sa capacité à généraliser sur de nouvelles données.

---

Ces questions couvrent des concepts clés comme la fonction sigmoïde, la fonction de coût, la régularisation, et les métriques de performance dans la régression logistique.

The note you provided is a comprehensive presentation on model evaluation techniques in machine learning, particularly focusing on issues of underfitting and overfitting. It uses various learning models such as linear regression and polynomial regression to demonstrate these concepts. The key points covered include:

- Underfitting : Occurs when a model is too simple to capture the underlying data patterns, leading to poor performance on both training and test sets.
- Overfitting : Happens when a model is too complex, fitting the training data too well but performing poorly on new, unseen data.
- Model Evaluation : Emphasizes methods to evaluate the performance of machine learning models, with a focus on metrics like Root Mean Squared Error (RMSE).
- Learning Curves : These curves are used to illustrate the performance of models on both training and test data as the training set size increases, providing insight into whether a model is underfitting or overfitting.

---

#### ### 30 Multiple Choice Questions (in French)

1. Qu'est-ce que le sous-apprentissage ?

- A) Un modèle trop simple
- B) Un modèle trop complexe
- C) Un modèle parfaitement ajusté
- Réponse : A

Explication : Un modèle trop simple ne capture pas bien les données et a des performances médiocres.

2. Le surapprentissage se produit lorsque :

- A) Le modèle est bien ajusté
- B) Le modèle s'ajuste trop aux données d'entraînement
- C) Le modèle a de mauvaises performances sur les données d'entraînement
- Réponse : B

Explication : Le surapprentissage signifie que le modèle s'ajuste trop aux données d'entraînement et échoue sur les nouvelles données.

3. Quelle méthode peut être utilisée pour éviter le surapprentissage ?

- A) Réduction de la complexité du modèle
- B) Augmenter la complexité du modèle
- C) Ignorer les données de test
- Réponse : A

Explication : La réduction de la complexité du modèle permet d'éviter un ajustement excessif des données d'entraînement.

4. Quel est le but des courbes d'apprentissage ?

- A) Évaluer la performance du modèle
- B) Créer des modèles plus complexes
- C) Augmenter la taille des données
- Réponse : A

Explication : Les courbes d'apprentissage sont utilisées pour évaluer la performance du modèle sur les ensembles de données d'entraînement et de test.

5. Un modèle qui présente un sous-apprentissage aura :

- A) Une erreur faible sur l'ensemble de test
- B) Une erreur élevée sur l'ensemble d'entraînement
- C) Une erreur faible sur l'ensemble d'entraînement
- Réponse : B

Explication : Un modèle sous-appris a des erreurs élevées à la fois sur les ensembles d'entraînement et de test.

6. Un modèle surajusté :

- A) Fonctionne bien sur les nouvelles données
- B) Fonctionne mal sur les nouvelles données
- C) A une complexité réduite
- Réponse : B

Explication : Un modèle surajusté a de bonnes performances sur les données d'entraînement, mais échoue sur les nouvelles données.

7. Quel est l'objectif principal de l'évaluation d'un modèle ?

- A) Construire un modèle avec une faible complexité
- B) Mesurer la performance du modèle sur des données invisibles
- C) Optimiser le modèle uniquement sur les données d'entraînement
- Réponse : B

Explication : L'évaluation d'un modèle consiste à vérifier ses performances sur des données qui n'ont pas été vues pendant l'entraînement.

8. La régression polynomiale peut aider à :

- A) Réduire le nombre d'attributs
- B) Créer des modèles plus complexes pour capturer les relations non linéaires
- C) Simplifier les données d'entrée
- Réponse : B

Explication : La régression polynomiale est utilisée pour capturer des relations non linéaires.

9. Qu'est-ce que RMSE (Root Mean Squared Error) mesure ?

- A) La précision d'un modèle
- B) L'écart entre les valeurs prédites et les valeurs réelles
- C) La vitesse de convergence du modèle
- Réponse : B

Explication : RMSE mesure la différence entre les valeurs prédites par un modèle et les valeurs réelles.

10. Comment réduire le risque de sous-apprentissage ?

- A) Utiliser un modèle plus complexe
- B) Réduire le nombre d'exemples
- C) Diminuer la taille de l'ensemble d'entraînement
- Réponse : A

Explication : L'utilisation d'un modèle plus complexe permet de mieux capturer les données.

11. Un modèle de régression linéaire est approprié pour :

- A) Des relations linéaires entre les variables
- B) Des relations non linéaires
- C) Des données catégoriques
- Réponse : A

Explication : La régression linéaire est utilisée pour des relations linéaires entre les variables.

12. Qu'est-ce que l'ingénierie des attributs ?

- A) La création de nouveaux modèles
- B) La transformation des données brutes pour améliorer la performance du modèle
- C) L'évaluation des performances du modèle
- Réponse : B

Explication : L'ingénierie des attributs consiste à transformer les données pour améliorer la performance des modèles.

13. La régularisation est utilisée pour :

- A) Réduire le surapprentissage
- B) Augmenter le surapprentissage
- C) Réduire la taille de l'ensemble de test
- Réponse : A

Explication : La régularisation ajoute une pénalité pour réduire la complexité du modèle et éviter le surapprentissage.

14. Une courbe d'apprentissage qui montre des erreurs similaires sur l'entraînement et le test indique :

- A) Surapprentissage
- B) Sous-apprentissage
- C) Ajustement parfait
- Réponse : B

Explication : Des erreurs similaires sur les ensembles d'entraînement et de test indiquent un sous-apprentissage.

15. Quel indicateur montre qu'un modèle surajusté performe mal sur des données non vues ?

- A) RMSE élevé sur les données d'entraînement
- B) RMSE faible sur les données d'entraînement mais élevé sur les données de test
- C) RMSE faible sur les deux ensembles de données
- Réponse : B

Explication : Un modèle surajusté a une faible erreur sur l'ensemble d'entraînement mais une erreur élevée sur les données de test.

16. Quelle technique est utilisée pour créer de nouvelles variables à partir des données brutes ?

- A) Ingénierie des attributs
- B) Régression linéaire
- C) Validation croisée
- Réponse : A

Explication : L'ingénierie des attributs consiste à créer de nouvelles variables à partir des données brutes pour améliorer les modèles.

17. La validation croisée permet :

- A) De réduire la taille de l'ensemble d'entraînement
- B) De mieux évaluer la performance du modèle en utilisant plusieurs sous-ensembles de données
- C) D'ajuster la complexité du modèle
- Réponse : B

Explication : La validation croisée permet d'évaluer la performance du modèle en utilisant différentes partitions des données.

18. Quelle est la conséquence d'un surapprentissage sur les performances du modèle ?

- A) Bonne performance sur les données de test
- B) Mauvaise performance sur les nouvelles données
- C) Bonne généralisation
- Réponse : B

Explication : Un modèle surajusté performe bien sur l'ensemble d'entraînement mais échoue à généraliser sur les nouvelles données.

19. Quel modèle est le plus susceptible de sous-apprendre les données ?

- A) Un modèle linéaire pour des données non linéaires
- B) Un modèle polynomial de degré élevé
- C) Un réseau neuronal profond
- Réponse : A

Explication : Un modèle linéaire est trop simple pour capturer des relations non linéaires, ce qui conduit au sous-apprentissage.

20. La régularisation L2 (ridge) aide à :

- A) Réduire la complexité du modèle en pénalisant les coefficients
- B) Augmenter la précision du modèle sur les données d'entraînement

- C) Supprimer les attributs non pertinents
- Réponse : A

Explication : La régularisation L2 ajoute une pénalité pour réduire les coefficients, ce qui réduit le surapprentissage.

21. Qu'est-ce qui distingue une mesure micro d'une mesure macro dans l'évaluation des modèles ?

- A) Micro agrège toutes les classes, macro calcule une moyenne par classe
- B) Macro agrège toutes les classes, micro calcule une moyenne par classe
- C) Macro est utilisée uniquement pour la régression
- Réponse : A

Explication : Les mesures micro agrégeront toutes les classes tandis que les mesures macro feront la moyenne des performances sur chaque classe.

22. Pourquoi fixer une graine dans un générateur de nombres aléatoires est-il important ?

- A) Pour obtenir des résultats différents à chaque exécution
- B) Pour assurer la reproductibilité des résultats
- C) Pour augmenter la complexité du modèle
- Réponse : B

Explication : Fixer une graine permet de reproduire exactement la même séquence de nombres aléatoires, assurant la reproductibilité des expériences.

23. Quel modèle est plus susceptible de surajuster les données d'entraînement ?

- A) Modèle simple avec peu de paramètres
- B) Modèle complexe avec beaucoup de paramètres
- C) Modèle de régression linéaire
- Réponse : B

Explication : Un modèle complexe avec beaucoup de paramètres est plus susceptible de surajuster les données d'entraînement.

24. Dans la régression polynomiale, un modèle avec un degré très élevé :

- A) A toujours une meilleure généralisation
- B) Peut surajuster les données
- C) Est toujours plus rapide à entraîner
- Réponse : B

Explication : Un degré élevé peut entraîner un surajustement, capturant des variations spécifiques à l'ensemble d'entraînement.

25. Quelle est la conséquence d'utiliser trop peu de données d'entraînement ?

- A) Augmente la généralisation du modèle
- B) Augmente le risque de surapprentissage ou de sous-apprentissage
- C) Améliore les performances du modèle
- Réponse : B

Explication : Utiliser trop peu de données d'entraînement augmente les risques d'erreurs de généralisation, de surapprentissage ou de sous-apprentissage.

26. Comment la régularisation L1 (lasso) agit-elle sur les coefficients d'un modèle ?

- A) Réduit les coefficients de manière non linéaire
- B) Peut amener certains coefficients à être exactement nuls
- C) Augmente tous les coefficients proportionnellement
- Réponse : B

Explication : La régularisation L1 peut amener certains coefficients à être exactement nuls, effectuant une forme de sélection d'attributs.

27. Lequel des éléments suivants est une méthode pour évaluer la qualité de prédiction d'un modèle ?

- A) RMSE
- B) Validation croisée
- C) Ingénierie des attributs
- Réponse : A

Explication : RMSE (Root Mean Squared Error) mesure la qualité des prédictions en évaluant l'écart entre les valeurs prédites et réelles.

28. Une méthode de validation croisée à k plis :

- A) Divise les données en k sous-ensembles égaux
- B) Évalue le modèle sur toutes les données d'entraînement
- C) N'est pas utilisée pour l'évaluation de modèle
- Réponse : A

Explication : La validation croisée à k plis divise les données en k sous-ensembles pour tester la performance sur différents segments de données.

29. Qu'est-ce que la régularisation tente d'éviter dans un modèle ?

- A) Surapprentissage
- B) Sous-apprentissage
- C) Augmentation de la taille des données
- Réponse : A

Explication : La régularisation réduit la complexité du modèle pour éviter le surapprentissage.

30. Quel est l'objectif principal des courbes d'apprentissage ?

- A) Déterminer la performance du modèle sur des ensembles de test et d'entraînement à différentes tailles
- B) Réduire la taille des données
- C) Augmenter la complexité des modèles
- Réponse : A

Explication : Les courbes d'apprentissage permettent d'évaluer la performance d'un modèle à différents niveaux de données d'entraînement et de test.

---

Ces questions couvrent différents aspects du modèle d'évaluation, y compris sous-apprentissage, surapprentissage, régularisation, et les concepts de validation croisée.