

EXAMEN - Machine learning

Nombre de questions: 14

Score de passage: 70.0%

Temps limite: 30 minutes

Question 1 (1.0 point)

Qu'est-ce que le machine learning ?

- ☒ Un sous-domaine de l'intelligence artificielle
- ☐ Un type de logiciel
- ☐ Un langage de programmation
- ☐ Un système d'exploitation

Explication: Le machine learning est défini comme un sous-domaine de l'intelligence artificielle qui permet aux systèmes d'apprendre à partir de données sans être explicitement programmés, comme discuté dans le chapitre 2 du module.

Question 2 (1.0 point)

Quel est l'objectif principal du machine learning ?

- ☐ Automatiser les tâches manuelles
- ☒ Faire des prédictions basées sur des données
- ☐ Développer des logiciels
- ☐ Créer des sites web

Explication: L'objectif principal du machine learning est de faire des prédictions ou des classifications basées sur des données, ce qui est essentiel pour le développement d'applications intelligentes.

Question 3 (1.0 point)

Quelle est la première étape dans un projet de machine learning ?

- Évaluation du modèle
- ✓ Préparation des données
- Déploiement du modèle
- Choix de l'algorithme

Explication: La préparation des données est la première étape dans un projet de machine learning, car des données propres et bien structurées sont essentielles pour un modèle performant.

Question 4 (1.0 point)

Quel type de machine learning utilise des données étiquetées pour l'apprentissage ?

- ☒ Apprentissage supervisé
- ☐ Apprentissage non supervisé
- ☐ Apprentissage par renforcement
- ☐ Apprentissage semi-supervisé

Explication: L'apprentissage supervisé utilise des données étiquetées pour entraîner les modèles, permettant ainsi de prédire des résultats sur de nouvelles données.

Question 5 (1.0 point)

Quel algorithme est souvent utilisé pour la classification dans le machine learning ?

- K-means
- Régression linéaire
- ✓ Arbre de décision
- Analyse en composantes principales

Explication: L'arbre de décision est un algorithme couramment utilisé pour des tâches de classification en machine learning, car il divise les données en fonction de critères spécifiques.

Question 6 (1.0 point)

Dans quel scénario utiliseriez-vous l'apprentissage non supervisé ?

- ☐ Pour prédire le prix d'une maison
- ☒ Pour segmenter des clients en groupes
- ☐ Pour classer des emails comme spam ou non spam
- ☐ Pour recommander des films

Explication: L'apprentissage non supervisé est utilisé pour segmenter des données en groupes, comme dans le cas de la segmentation de clients, où les étiquettes ne sont pas disponibles.

Question 7 (1.0 point)

Qu'est-ce qu'un surapprentissage (overfitting) dans le contexte du machine learning ?

- ☐ Un modèle qui généralise bien
- ☒ Un modèle qui est trop complexe et s'adapte trop aux données d'entraînement
- ☐ Un modèle qui ne peut pas apprendre
- ☐ Un modèle qui est trop simple

Explication: Le surapprentissage se produit lorsque le modèle est trop complexe et s'adapte trop aux données d'entraînement, ce qui nuit à sa capacité de généralisation sur de nouvelles données.

Question 8 (1.0 point)

Quelle méthode peut être utilisée pour éviter le surapprentissage ?

- ☐ Augmenter le nombre d'étiquettes
- ☐ Utiliser un modèle plus complexe
- ☒ Utiliser la validation croisée
- ☐ Réduire le nombre de données

Explication: La validation croisée est une méthode efficace pour éviter le surapprentissage, car elle permet de tester le modèle sur différentes sous-ensembles de données.

Question 9 (1.0 point)

Quel est le rôle de la fonction de coût dans l'entraînement d'un modèle de machine learning ?

- Mesurer la qualité des données
- ✓ Évaluer la performance du modèle
- Déterminer la complexité du modèle
- Ajuster le taux d'apprentissage

Explication: La fonction de coût évalue la performance du modèle en mesurant l'erreur entre les prédictions du modèle et les valeurs réelles, permettant ainsi d'ajuster les paramètres du modèle.

Question 10 (1.0 point)

Quelle technique est souvent utilisée pour réduire la dimensionnalité des données ?

- Régression logistique
- K-means
- ✓ Analyse en composantes principales (PCA)
- Arbre de décision

Explication: L'analyse en composantes principales (PCA) est une technique de réduction de dimensionnalité qui aide à simplifier les données tout en préservant autant d'information que possible.

Question 11 (1.0 point)

Comment évalueriez-vous la précision d'un modèle de classification ?

- ☒ En utilisant la matrice de confusion
- ☐ En comptant les erreurs de prédiction
- ☐ En mesurant le temps d'entraînement
- ☐ En testant avec un ensemble de données non étiquetées

Explication: La matrice de confusion est un outil essentiel pour évaluer la précision d'un modèle de classification en montrant les vrais positifs, faux positifs, vrais négatifs et faux négatifs.

Question 12 (1.0 point)

Quel est l'impact de la taille de l'ensemble de données sur l'apprentissage d'un modèle ?

- Un ensemble de données plus petit améliore la précision
- La taille de l'ensemble de données n'a pas d'impact
- ✓ Un ensemble de données plus grand aide à généraliser mieux
- Un ensemble de données plus grand augmente le surapprentissage

Explication: *Un ensemble de données plus grand aide souvent à améliorer la généralisation du modèle, car il a plus d'exemples pour apprendre les variations des données.*

Question 13 (1.0 point)

Quel type de problème résout un algorithme de régression dans le machine learning ?

- ☐ Classification
- ☐ Clustering
- ☒ Prédiction continue
- ☐ Segmentation

Explication: Un algorithme de régression résout des problèmes de prédiction continue, où il prédit une valeur numérique basée sur des données d'entrée.

Question 14 (1.0 point)

Quel est le but de la normalisation des données dans le machine learning ?

- Augmenter le nombre de données
- Équilibrer les classes
- Réduire la variance des données
- ✓ Mettre toutes les données à la même échelle

Explication: La normalisation des données met toutes les données à la même échelle, ce qui est crucial pour certains algorithmes de machine learning qui sont sensibles à l'échelle des données.