

Acquérir les fondamentaux mathématiques et statistiques pour l'IA

Date et durée
Code formation : IA039FR Durée : 1 jour Nombre d'heures : 7 heures
Description
<p>L'intelligence artificielle (IA) repose sur des fondations mathématiques solides qui sont souvent perçues comme une barrière à l'entrée. Cette formation d'une journée est conçue pour démystifier ces concepts essentiels et vous donner les clés de lecture nécessaires pour comprendre ce qui se passe réellement sous le capot des algorithmes.</p> <p>Le programme revisite les notions fondamentales d'algèbre linéaire (vecteurs, matrices) et d'analyse (dérivées, gradients) en les connectant directement à leur utilité dans les réseaux de neurones. Vous explorerez également les statistiques descriptives et inférentielles pour apprendre à analyser vos données avec rigueur et à interpréter la fiabilité de vos modèles prédictifs.</p> <p>L'approche se veut résolument pragmatique : pas de théorie abstraite sans application immédiate. À travers des ateliers pratiques utilisant Python et ses bibliothèques scientifiques (NumPy, Pandas, Matplotlib), vous manipulerez concrètement les concepts mathématiques pour ancrer votre compréhension et aborder sereinement vos futurs projets Data.</p>
Objectifs
<p>À l'issue de cette formation, vous atteindrez les objectifs de compétences suivants :</p> <ul style="list-style-type: none">• identifier les concepts mathématiques clés (algèbre, probabilités) utilisés dans les algorithmes d'IA ;• appliquer les statistiques descriptives pour analyser et visualiser la distribution des données ;• utiliser les statistiques inférentielles pour valider des hypothèses et fiabiliser les résultats ;• interpréter les métriques de performance et les biais (surapprentissage) d'un modèle de Machine Learning ;• manipuler les outils Python (NumPy, Scikit-learn) pour explorer les données et tester des modèles.
Points forts
<ul style="list-style-type: none">• Pédagogie appliquée : les concepts mathématiques abstraits sont immédiatement traduits en code Python exécutable pour une compréhension concrète.• Socle essentiel : vous acquérerez le bagage indispensable pour ne plus être bloqué par la lecture d'articles scientifiques ou la documentation technique.• Focus IA : le programme ne traite que des mathématiques utiles à l'IA, éliminant le superflu pour une efficacité maximale sur une journée.• Boîte à outils : vous repartirez avec des exemples de code réutilisables pour vos analyses statistiques et visualisations.
Modalités d'évaluation

Travaux Pratiques

Pré-requis

Suivre cette formation nécessite les prérequis suivants :

- **Connaissances techniques** : des bases en informatique et en programmation Python sont recommandées pour les ateliers.
- **Connaissances théoriques** : des notions élémentaires en mathématiques (niveau lycée).

Public

Cette formation s'adresse aux profils techniques souhaitant consolider leur socle théorique. Le public inclut notamment :

- les **développeurs et ingénieurs débutants en IA** qui souhaitent comprendre la logique mathématique derrière les bibliothèques qu'ils utilisent ;
- les **Data Analysts et Data Scientists juniors** qui veulent renforcer leur rigueur statistique pour mieux interpréter leurs résultats ;
- les **professionnels en reconversion** vers les métiers de la Data, ayant besoin d'une remise à niveau ciblée en mathématiques ;
- les **étudiants en informatique** qui désirent faire le lien entre la théorie mathématique et la pratique du Machine Learning.

Programme

Module 1 : maîtriser l'algèbre linéaire et les bases de l'optimisation

- Les vecteurs et matrices : opérations fondamentales et application aux réseaux de neurones.
- Les fonctions courantes en IA (Sigmoïde, ReLU) et leurs dérivées.
- Le principe de l'optimisation par descente de gradient pour l'entraînement des modèles.

Travaux pratiques

- Manipuler des matrices et visualiser des fonctions mathématiques avec les bibliothèques NumPy et Matplotlib.

Module 2 : analyser les données avec les statistiques et probabilités

- Les mesures de tendance (moyenne, médiane) et de dispersion (écart-type, variance).
- La visualisation des données : interprétation d'histogrammes et de boxplots.
- Les lois de probabilités classiques (Loi Normale, Binomiale) et leur rôle en modélisation.

Travaux pratiques

- Réaliser l'analyse exploratoire d'un jeu de données avec Pandas et simuler des distributions statistiques.

Module 3 : valider les modèles grâce aux statistiques inférentielles

- Les tests d'hypothèses (Student, Khi²) pour valider la pertinence des données.
- L'interprétation des métriques d'évaluation (Précision, Rappel, ROC) et la gestion du surapprentissage (overfitting).
- La validation croisée pour assurer la robustesse des prédictions.

Travaux pratiques

- Mener un test d'hypothèse sur un cas réel (ex. Titanic) et évaluer un modèle de classification avec Scikit-learn.

Les noms de marques et logos éventuellement cités dans cette fiche (ex. Python, NumPy, Pandas, Scikit-learn) sont la propriété de leurs détenteurs respectifs. Leur mention à des fins pédagogiques ne constitue ni un engagement ni un partenariat.