

Formation Valoriser vos données dormantes avec le Machine Learning

Mise à jour janvier 2024

Inter 1800€ HT/participant

Intra 6750€ HT* groupe de 6 participants

*hors frais de déplacement et de personnalisation de formation sur-mesure

Démarrez en toute sérénité vos projets de Machine Learning : cette formation vous donnera les connaissances nécessaires pour identifier les applications vous permettant une valorisation rapide des données dormantes que stocke votre entreprise.

Nous verrons comment analyser ces dernières, les nettoyer et vérifier qu'elles satisfont les pré-requis propres à chaque type de méthode. Nous vous donnerons les bases pour réaliser rapidement des visualisations capables de les mettre en valeur. Enfin, nous étudierons les solutions de Machine Learning ayant fait leurs preuves.

Durée: 21.00 heures (3.00 jours)

À QUI S'ADRESSE CETTE FORMATION ?

Profil du participant

- Ingénieurs, Développeurs, Chercheurs,
- Data scientist

Prérequis

- Connaissances de base en algorithmie
- Connaissances de base en mathématiques (probabilités et statistiques : notion de variable aléatoire, de probabilité conditionnelle, de moyenne et de variance / vecteurs et espaces de dimensions N et fonction à N variables)
- Connaissance du langage Python (structures algorithmiques de base (if, for, while) / chaînes, booléens, nombres flottants et entiers / listes, tuples, dictionnaires / les fonctions / visibilité des variables / utilisation des packages et modules)

OBJECTIFS PÉDAGOGIQUES

1. Être capable de déterminer si vos données sont adéquates pour une application de type Machine Learning
2. Savoir créer rapidement une visualisation pertinente pour vos données
3. Connaître les architectures les plus célèbres en Machine Learning et Deep Learning
4. Être capable d'identifier les problèmes pouvant bénéficier de ce type d'approche

CONTENU (PROGRESSION PÉDAGOGIQUE)

Introduction

- Introduction au Machine Learning : principes, concepts de base
- Quelques exemples d'applications
- Dans quel cas de figure peut-on utiliser du Machine Learning ?

Préparer ses données avec Numpy et Pandas

- Numpy : traitement de données numériques à N dimensions
- Pandas : traitement des données tabulaires et séries temporelles
- Rechercher les données erronées
 - Par la visualisation : Seaborn, Plotly, Folium, Superset
 - Algorithmes détection automatiques d'outliers : Isolation Forest, Local Outlier Factor, Elliptic Envelope
 - Clustering avec détection du bruit : HDBSCAN
- Écarter ou corriger les erreurs
- Constituer un jeu de tests et un ensemble de données d'apprentissage
- Pré-traitements usuels : PCA, normalisation, centrage

Data visualisation

- Règles de bases pour une Data Viz pertinente
- Quels types de graphiques pour quels usages
- Jupiter Notebook
- Seaborn : graphique statistique
- Folium : données cartographiques
- Superset et MapBox
- Plotly

Utilisation des méthodes standards de Machine Learning

- Pour quel type d'applications ?
- Skicit-learn
- Validation croisée
- Automatisation de la recherche d'une méthode adéquate : AutoSklearn et TPOT
- Sauvegarde d'un modèle entraîné

Utilisation du Deep Learning

- Les problèmes maîtrisés
- Architectures reconnues
- Tests rapides avec Keras

ORGANISATION

Formateur

Formation assurée par un expert Data Science et en Machine Learning

Moyens pédagogiques et techniques

- Accueil des stagiaires dans une salle dédiée à la formation
- Documents supports de formation partagés
- Exposés théoriques
- Étude de cas concrets
- Quiz en ligne
- Mise à disposition en ligne de documents supports à la suite de la formation

Dispositif de suivi de l'exécution de l'évaluation des résultats de la formation

- Feuilles de présence dématérialisées
- Questions orales ou écrites (QCM)
- Mises en situation
- Formulaire d'évaluation de la formation

Délais d'accès

La convocation est envoyée 7 jours avant le début de la formation après réception du bon de commande signé.

Coordonnées de l'équipe pédagogique

- Responsable formation, handicap et votre formateur : Cécile Chardonneau formation@makina-corpus.com
- Suivi facturation : Nathalie Carles Salmon administration@makina-corpus.com