MACHINE LEARNING, MÉTHODES ET SOLUTIONS

Formation proposée par BIZOÏ RAZVAN

Machine Learning

Le Machine Learning couvre l'ensemble des méthodes et concepts qui permettent d'extraire automatiquement à partir de données, des modèles de prédiction et de prise de décision. Vous mettrez en œuvre toute la chaîne de conception appliquée au Machine Learning dans un contexte Big Data batch et streaming.

◷DURÉE
5 Jour(s)
▣MODALITÉS
Présentiel / Distanciel
♙PARTICIPANTS
Jusqu'à 10 participants
★LANGUE
Français
LIEU
France entière
♙CERTIFICATION
Attestation de formation
▦MISE À JOUR
11/04/2023

VOTRE FORMATEUR

BR **BIZOÏ RAZVAN**
Formateur professionnel

TARIFS INDICATIFS

Inter-entreprise	**Pas de dates planifiées**
Intra-entreprise	**Sur devis**
Individuelle	**Sur devis**
Durée	**5 Jour(s)**

Tarifs indicatifs, sous réserve d'analyse du besoin et de devis.

◎Objectifs pédagogiques

- ✓Comprendre les différents modèles d'apprentissage.
- ✓Modéliser un problème pratique sous forme abstraite.
- ✓Identifier les méthodes d'apprentissage pertinentes pour résoudre un problème.
- ✓Appliquer et évaluer les méthodes identifiées sur un problème.
- ✓Faire le lien entre les différentes techniques d'apprentissage.

⌘Compétences visées

Machine Learning

♙Public cible et prérequis

Public cible

- Ingénieurs/chefs de projet souhaitant considérer les techniques d'apprentissage automatique dans la résolution de problèmes industriels.

Prérequis

- Connaissances équivalentes à celles apportées par le stage Data Science avec Python/R

Programme détaillé

Introduction au Machine Learning
- Le Big Data et le Machine Learning.
- Les algorithmes d'apprentissage supervisés, non supervisés et par renforcement.
- Les étapes de construction d'un modèle prédictif.
- Détecter les valeurs aberrantes et traiter les données manquantes.
- Comment choisir l'algorithme et les variables de l'algorithme -
Travaux pratiques
Prise en main de l'environnement R/Python à l'aide de Jupyter Notebook. Visualiser plusieurs exemples de modèles fournis.

Procédures d'évaluation de modèles
- Les techniques de ré-échantillonnage en jeu d'apprentissage, de validation et de test.
- Test de représentativité des données d'apprentissage.
- Les techniques de bootstrap
- La validation croisée
- Définition d'une métrique de performance.
- Mesures de performance des modèles prédictifs.
- Matrice de confusion, de coût et la courbe ROC et AUC.
Travaux pratiques
Evaluation et comparaison des différents algorithmes sur les modèles fournis.

Les algorithmes non supervisés
- Le clustering hiérarchique.
- Le clustering non hiérarchique.
- Les approches mixtes.
- Le clustering par mesure de densité DBSCAN
- Autres approches du Clustering
Travaux pratiques
Traitements de clustering non supervisés sur plusieurs jeux de données.

Les algorithmes supervisés régressions
- Le principe de régression linéaire univariée.
- La régression multivariée.
- La régression polynomiale.
- La régression régularisée.
Travaux pratiques
Mise en œuvre des régressions sur plusieurs types de données.

Les algorithmes supervisés classifications
- Le Naive Bayes.
- La régression logistique.
- Les arbres de décision
- Machines à vecteurs supports et méthodes à noyaux.
- Quantification vectorielle.
Travaux pratiques
Mise en œuvre des classifications sur plusieurs types de données.

Les algorithmes ensemblistes
- L'ensemble de méthodes bagging
- Ramdom Forest.
- Les méthodes de Boosting.
- L'assemblage de plusieurs modèles.
Travaux pratiques
Mise en œuvre des analyses avec un assemblage des algorithmes.

Concepts fondamentaux d'un réseau de neurones
- Rappel de bases mathématiques.
- Le réseau de neurones : architecture, fonctions d'activation et de pondération des activations précédentes.
- L'apprentissage d'un réseau de neurones : fonctions de coût, back-propagation, stochastic gradient descent.
- Modélisation d'un réseau de neurones.
- Data Augmentation : comment équilibrer un dataset

- Généralisation des résultats d'un réseau de neurones.
- Initialisations et régularisations d'un réseau de neurones.
- Optimisations et algorithmes de convergence.
Travaux pratiques
Approximation d'une fonction et d'une distribution par un réseau de neurones.

Machine Learning en production
- Le cycle de vie d'un projet de machine Learning.
- Les spécificités liées au développement d'un modèle en environnement distribué.
- Les outils du marché pour le traitement de la donnée.
- Le déploiement Big Data avec Spark et la MLlib.
- Le Cloud : DataBrics, Dataiku, DataDog, Amazon, Microsoft Azure ML, IBM Watson...
- La maintenance du modèle.
Travaux pratiques
Mise en production d'un modèle prédictif avec l'intégration dans des processus de batch et dans des flux de traitements.

▣Méthodes pédagogiques & évaluation

Méthodes pédagogiques

- ✓ Apports théoriques et retours d'expérience
- ✓ Exercices pratiques et mises en situation
- ✓ Travaux de groupe et échanges entre pairs
- ✓ Études de cas concrets

Modalités d'évaluation

- ✓ Quiz de validation des acquis
- ✓ Mises en situation évaluées
- ✓ Projet fil rouge
- ✓ Feuille de route post-formation

♿Accessibilité

Une formation accessible à tous

Ce programme peut être adapté pour les personnes en situation de handicap. Contactez-nous pour étudier vos besoins spécifiques.