Machine Learning - Technologies et bonnes pratiques

2 j (14 heures) Ref : MALE

Public

Chefs de projets, développeurs, data scientists, architectes ou toute personne souhaitant comprendre comment organiser le traitement des données et structurer les processus de Machine Learning

Pré-requis

Avoir connaissance des principes du Big Data et des architectures techniques mises en oeuvre

Moyens pédagogiques

Formation réalisée en présentiel ou à distance selon la formule retenue
Exposés, cas pratiques, synthèse, assistance post-formation pendant trois mois
Un poste par stagiaire, vidéoprojecteur, support de cours fourni à chaque stagiaire

Modalités de suivi et d'évaluation

Feuille de présence émargée par demi-journée par les stagiaires et le formateur
Exercices de mise en pratique ou quiz de connaissances tout au long de la formation permettant de mesurer la progression des stagiaires
Questionnaire d'évaluation de la satisfaction en fin de stage
Auto-évaluation des acquis de la formation par les stagiaires
Attestation de fin de formation

Objectifs

Définir les étapes de préparation des données et les algorithmes de Machine Learning

Découvrir les différentes méthodes d'apprentissage automatique

Préparer vos données avant de les exploiter

Programme détaillé

INTRODUCTION

Zoom sur les données

- Format

- Volumes

- Structures

Requêtes

Attentes

Utilisateurs

Etapes de la préparation des données

Définitions

Présentation du Data Munging

Le rôle du data scientist

ETUDE DE CAS

Mise en oeuvre pratique des différentes phases

- Nettoyage
- Enrichissement
- Organisation des données

MACHINE LEARNING

Définition

Les attentes

Les valeurs d'observation et les variables cibles

Ingénierie des variables

APPRENTISSAGE AUTOMATIQUE

Les méthodes

- Apprentissage supervisé et non supervisé

Classification des données

Algorithmes

- Régression linéaire
- k-voisins
- Classification naïve bayésienne
- Arbres de décision...

LES RISQUES ET ECUEILS

Importance de la préparation des données

L'écueil du "surapprentissage"

VISUALISATION DES DONNEES

L'intérêt de la visualisation

Outils disponibles

Exemples de visualisation avec R et Python