

Data Classification et Machine Learning

Machine Learning Technologies et bonnes pratiques

Ces 2 jours de formation vous permettront d'appréhender le machine learning (aussi appelé apprentissage automatique en français). Une fois que le ou la Data Scientist a effectué son travail de collecte, de nettoyage et d'exploration des données, il s'attaque à la partie de modélisation statistique.

C'est ce processus que nous allons explorer ensemble dans cette formation au machine learning, un ensemble de techniques puissantes permettant de créer des modèles prédictifs à partir de données, sans avoir été explicitement programmées.

Détails

- **Code** : DB-MCH
- **Durée** : 2 jours (14 heures)

Public

- Architectes
- Chefs de projets
- Data Scientist
- Développeurs

Pré-requis

Objectifs

- Définir les étapes de préparation des données, et les algorithmes de Machine Learning.
- Organiser le traitement des données et structurer les processus de Machine Learning.

Programme

Introduction

- Zoom sur les données : format, volumes, structures, ...
- et les requêtes, attentes, des 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 par rapport au Machine Learning
- 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, etc ..

Les risques et écueils

- Importance de la préparation des données.
- L'écueil du « surapprentissage ».

Visualisation des données

- L'intérêt de la visualisation.
- Outils disponibles,
- Exemples de visualisation avec R et Python

Modalités

- **Type d'action** : Acquisition des connaissances
- **Moyens de la formation** : Formation présentielle – 1 poste par stagiaire – 1 vidéo projecteur – Support de cours fourni à chaque stagiaire
- **Modalités pédagogiques** : Exposés – Cas pratiques – Synthèse
- **Validation** : Exercices de validation – Attestation de stages