

## **Machine Learning : technologies et bonnes pratiques** CB031

**Durée:** 2 jours

1 180 €

13 au 14 mars

19 au 20 juin

9 au 10 octobre

11 au 12 décembre

### **Public:**

Chefs de projet, développeurs, data scientists, architectes souhaitant comprendre comment organiser le traitement des données et structurer les processus de Machine Learning.

### **Objectifs:**

Savoir définir les étapes de préparation des données, et les algorithmes de Machine Learning.

### **Connaissances préalables nécessaires:**

Connaissances des principes du BigData, et des architectures techniques mises en oeuvre.

### **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