

## FORMATION DATA CLASSIFICATION ET MACHINE LEARNING

Cette formation vous permettra d'appréhender le machine learning du point de vue du traitement de la donnée (collecte, nettoyage, exploration des données...) ainsi qu'au processus de modélisation statistique. Grâce à un ensemble de techniques efficaces, vous serez en mesure, tel un data scientist, de créer des modèles prédictifs à partir de données sans avoir été explicitement programmées

### PROGRAMME

#### 1/ Introduction

- Zoom sur les données : format, volumes, structures, ...et les requêtes, attentes des utilisateurs.
- Étapes de la préparation des données.
- Définitions, présentation du data munging
- Le rôle du data scientist.

#### 2/ Gouvernance des données

- Qualité des données.
- Transformation de l'information en donnée. Qualification et enrichissement.
- Sécurisation et étanchéité des lacs de données.
- Flux de données et organisation dans l'entreprise. De la donnée maître à la donnée de travail. MDM.
- Mise en œuvre pratique des différentes phases :
- Nettoyage, enrichissement, organisation des données.

#### 3/ Traitements statistiques de base

- Introduction aux calculs statistiques. Paramétrisation des fonctions.
- Applications aux fermes de calculs distribués. Problématiques induites.
- Approximations. Précision des estimations.

#### 4/ Data Mining

- Besoin, apports et enjeux.
- Extraction et organisation des classes de données.
- Analyse factorielle.

#### 5/ Machine Learning

- Apprentissage automatique
- Définition, les attentes par rapport au Machine Learning
- Les valeurs d'observation, et les variables cibles.
- Ingénierie des variables.
- Les méthodes : apprentissage supervisé et non-supervisé
- Classification des données,
- Algorithmes : régression linéaire, k-moyennes, k-voisins, classification naïve bayésienne, arbres de décision, forêts aléatoires, etc ...
- Création de jeux d'essai, entraînement et construction de modèles.
- Prévisions à partir de données réelles. Mesure de l'efficacité des algorithmes.
- Courbes ROC.
- Parallélisation des algorithmes. Choix automatique.

PRIX (INTER-ENTREPRISE) : 1350 euros

14h jours

ML02

Vous souhaitez organiser cette formation dans vos locaux ?

**Demandez Houilly au  
01 84 25 05 10**

### OBJECTIFS

- Savoir mettre en œuvre les étapes de préparation des données
- Comprendre et mettre en place l'apprentissage automatique
- Les techniques de classification de données
- Les apports des réseaux de neurones et du Deep Learning

### PRE-REQUIS

- Connaissances des principes du BigData
- et des architectures techniques mises en œuvre

### PUBLIC CONCERNE

- Chefs de projet
- Développeurs
- Data Scientists
- Architectes

### DATES INTER-ENTREPRISES

Delais d'entree : sans  
A distance

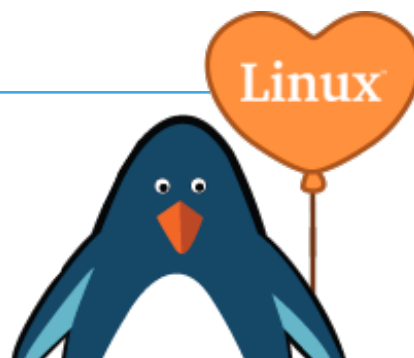
Paris

Si vous souhaitez organiser cette formation a une autre date contactez-nous.

## NOUS CONTACTER

**Openska**  
21 rue Louise Weiss  
75013 Paris

Tel : 01 84 25 05 10  
Tel : 01 84 17 44 76  
[www.openska.com](http://www.openska.com)



## **6/ IA**

- Un mot sur les réseaux de neurones.
- Réseaux de neurones à convolution. Modèles de CNN.
- Les types de couches : convolution, pooling et pertes.
- L'approche du Deep Learning. Deeplearning4j sur Spark.

## **7/ Les risques et écueils**

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

## **8/ Visualisation des données**

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

## ELEMENTS COMPLÉMENTAIRES À TRANSMETTRE IMPÉRATIVEMENT À VOTRE OPCO POUR VOTRE DEMANDE DE PRISE EN CHARGE

### HORAIRES

- 9h30-13h
- 14h-17h30

### ACCÈS HANDICAP

Oui, Consultez notre référent interne.

### MODALITÉS D'ÉVALUATION

L'évaluation des acquis se fait tout au long de la formation au travers de multiples exercices et mise en situation

### MODALITÉS DE SUIVI DE L'EXÉCUTION

- Contrôle systématique des présences par demi-journée
- Attestation de présence remis à chaque participant

### COMPÉTENCES DU FORMATEUR

Les experts qui animent la formation sont des spécialistes des matières abordées. Ils ont été validés par nos équipes pédagogiques tant sur le plan des connaissances métiers que sur celui de la

pédagogie, et ce pour chaque formation qu'ils animent. Ils ont majoritairement cinq à dix années d'expérience dans leur domaine de compétences et ont une grande pratique en entreprise

### MOYENS PÉDAGOGIQUES

Les supports pédagogiques sont imprimés par nos soins et transmis à chacun des stagiaires. Les supports sont aussi remis au format électronique aux participants. Ce support est projeté via un vidéo projecteur afin d'animer la formation. Des exercices d'application ou études de cas sont prévus afin de valider les acquis des stagiaires.

### MOYENS D'ACCOMPAGNEMENT

8 personnes maximum, formation animée par un formateur expert sur le sujet.  
Mise en situation : Pédagogie active et participative.  
Apport théorique et méthodologique.  
Etude de cas. Atelier pratique.

### NATURE DES TRAVAUX DEMANDÉS AU STAGIAIRE

Un exercice est réalisé par le participant à la fin de chaque chapitre. Le participant dispose de 20 min avant de passer à la correction avec le formateur.

### MODALITÉS TECHNIQUES EN CAS DE PROBLÈMES

Les connexions et installations d'outil peuvent être testées en amont de la formation avec le formateur et les stagiaires ou le donneur d'ordre. Ce test permet d'éviter tout accident technique lié aux outils de connexion à distance.

Le formateur prend la main sur le poste du participant en cas de difficulté durant la formation.



ORGANISME DE FORMATION RÉFÉRENCÉ SOUS LE NUMÉRO : 117 555 432 75