Architecture des données : stockage et accès

Appréhender et prendre en main les nouvelles architectures de données : Hadoop, NoSQL, Spark

Stage pratique
Data Science

Code :
DSARC

Durée :
3 jour(s) (21,00 heures)

Exposés : **50 %**
Cas pratiques : **40 %**
Echanges d'expérience : **10 %**

Inter-entreprises :
Prochaines sessions disponibles sur notre site web.
Tarif : 2 500,00 € HT / participant

Intra-entreprise :
Tarifs et dates sur demande.

DESCRIPTION

Si les algorithmes de Machine Learning ont connu des avancées majeures ces dernières années, c'est avant tout grâce à la quantité d'information disponible pour les entraîner. Accumuler toute cette donnée, la traiter, et la rendre disponible sont les enjeux principaux du mouvement Big Data.

Au cours de cette formation, nos consultants mettent à disposition les connaissances issues de leurs retours d'expériences auprès de nos clients, et vous font découvrir les bases des architectures permettant de répondre à ces enjeux de stockage et d'accès.

OBJECTIFS PEDAGOGIQUES

- Découvrir les notions centrales de stockage de données
- Appréhender les enjeux des nouvelles architectures de données (Hadoop, NoSQL, Spark), et positionner leurs usages au sein de l'univers Big Data
- Savoir manipuler ces technologies et les bases de données de façon conjointe, pour mener à bien des analyses efficaces

PUBLIC CIBLE

- Analyste
- Statisticien
- Développeur

PRE-REQUIS

- Notions de programmation sur la base d'un langage quelconque
- Manipulation basique de la ligne de commande Linux

METHODE PEDAGOGIQUE

Formation avec apports théoriques, échanges sur les contextes des participants et retours d'expérience pratique du formateur, complétés de travaux pratiques et de mises en situation.

PROFIL DES INTERVENANTS

Cette formation est dispensée par un·e ou plusieurs consultant·es d'OCTO Technology ou de son réseau de partenaires, expert·es reconnu·es des sujets traités.

Le processus de sélection de nos formateurs et formatrices est exigeant et repose sur une évaluation rigoureuse leurs capacités techniques, de leur expérience professionnelle et de leurs compétences pédagogiques.

MODALITÉS D'ÉVALUATION ET FORMALISATION À L'ISSUE DE LA FORMATION

L'évaluation des acquis se fait tout au long de la session au travers des ateliers et des mises en pratique.

Afin de valider les compétences acquises lors de la formation, un formulaire d'auto-positionnement est envoyé en amont et en aval de celle-ci.

Une évaluation à chaud est également effectuée en fin de session pour mesurer la satisfaction des stagiaires et un certificat de réalisation leur est adressé individuellement.

PROGRAMME PEDAGOGIQUE DETAILLE

Jour 1

INTRODUCTION

ACCÈS AUX DONNÉES

Les fichiers

- Arborescence
- Formats

Les bases de données

- Notion de Catalogue
- Le langage SQL
- Data Warehouses et Data Lake
- Bases d'analyse

API

- Définition
- Web Scraping

Traitements en mémoire

ARCHITECTE DE DONNÉES

Limites des systèmes traditionnels

- Limites des fichiers
- Limites des SGBD

Les architectures distribuées

- Patterns d'accès
- 1OLTP
- 2OLAP
- Distribution vs Réplication

Concepts essentiels

- Disponibilités
- Cohérence
- Tolérance à la partition

Le théorème CAP

Quorums

BASES NOSQL

Avantages et inconvénients

Modèles de données

- Key-Value
- Documents
- Column-Family
- Graph

Exemple : MongoDB

Les moteurs de recherche

JOUR 2

HADOOP

Introduction à Hadoop

- Histoire
- Ecosystème

HDFS

Map-Reduce

- Les phases de Map-Reduce
- Notion de job
- Exemple

YARN

Les distributions

La ligne de commande

Administration d'un cluster

- Resource Manager
- History Server
- Hue

ÉTUDES DE CAS

Traitements de courbes de charge

- Contexte et hypothèses
- Raisonnements

Analyse de logs

- Contexte et Hypothèses
- Raisonnements

CONCLUSION

Rappels des points abordés

Questions et réponses

Synthèse

Jour 3

DÉCOUVERTE DE SPARK

Spark Core

- RDD
- Transformations
- Pair RDD

Spark SQL

Spark Streaming

Structured Streaming

Accessibilité

L'inclusion est sujet important pour OCTO Academy.
Nos référent·es sont à votre disposition pour faciliter l'adaptation de votre formation à vos besoins spécifiques.
Pour les contacter : academy.accessibilite@octo.com