

Formation Spark, Les fondamentaux du Big Data

Présentation

Cette formation vous permettra de maîtriser les concepts fondamentaux de Spark.

Durée : 14,00 heures (2 jours)

Tarif INTRA : Nous consulter

Objectifs de la formation

- Comprendre les fondamentaux de Spark
- Découvrir les différents modules de Spark
- Manipuler des données avec Spark SQL

Prérequis

- Bonnes connaissances du langage Python

Public

- Développeurs
- Architectes

Programme de la formation

Présentation d'Apache Spark

- Introduction et historique
- Les langages de programmation supportés par Spark (Scala, Python et Java)
- Comparaison avec l'environnement Apache Hadoop
- Les modules de Spark (Spark SQL, Spark Streaming, MLlib, GraphX)

Resilient Distributed Dataset (RDD)

- C'est quoi un RDD ?
- Créer, manipuler et réutiliser des RDD
- Accumulateurs et variables broadcastées



- Utiliser des partitions

Manipuler des données structurées avec Spark SQL

- SQL, DataFrames et Datasets
- Les différents types de sources de données
- Interopérabilité avec les RDD
- Performance de Spark SQL
- JDBC/ODBC server et Spark SQL CLI

Spark sur un cluster

- Les différents types d'architectures : Standalone, Apache Mesos ou Hadoop YARN
- Configurer un cluster en mode Standalone
- Packager une application avec ses dépendances
- Déployer des applications avec Spark-submit
- Dimensionner un cluster

Organisation

Formateur

Les formateurs de Docaposte Institute sont des experts de leur domaine, disposant d'une expérience terrain qu'ils enrichissent continuellement. Leurs connaissances techniques et pédagogiques sont rigoureusement validées en amont par nos référents internes.

Moyens pédagogiques et techniques

- Alternance d'apports théoriques et de mises en pratique contextualisées.
- Progression pédagogique structurée, de la compréhension à la mise en œuvre opérationnelle.
- Pédagogie immersive favorisant une mise en application immédiate des compétences.
- Formation interactive et participative s'appuyant sur des ateliers pratiques, études de cas et échanges collectifs (+ 60 % du temps de formation.)
- Séquences pédagogiques rythmées et évaluations formatives régulières pour renforcer l'ancrage des apprentissages.

- Supports de cours, ressources documentaires et références mis à disposition par le formateur

Dispositif de suivi de l'exécution et de l'évaluation des résultats de la formation

En amont de la formation *(évaluation en ligne dans l'extranet apprenant)*

- Recueil des besoins et des attentes des participants.
- Auto-positionnement permettant d'évaluer le niveau initial des apprenants.

Tout au long de la formation

- Évaluations continues réalisées par le formateur à travers des questions orales, exercices, QCM, études de cas et mises en situation pratiques.

À l'issue de la formation *(évaluations en ligne dans l'extranet apprenant)*

- Auto-positionnement afin de mesurer l'acquisition des compétences visées.
- Validation de l'acquisition des compétences par le formateur
- Questionnaire de satisfaction « à chaud » à l'issue de la formation
- Questionnaire de satisfaction "à froid", 45 jours après la formation afin d'évaluer l'ancrage des acquis et leur mise en application en situation professionnelle.

Accessibilité

Nos formations peuvent être adaptées à certaines conditions de handicap. Nous contacter pour toute information et demande spécifique.