Big Data, méthodes et solutions pratiques pour l’analyse des données

Lien : https://innov-systems.com/formation/big-data-methodes-et-solutions-pratiques-pour-lanalyse-des-donnees

DURÉE
5 jours (35h)

RÉFÉRENCE
BSI95

CATÉGORIE
Big Data - Techniques D'analyse

Innov Systems

OBJECTIFS DE LA FORMATION

À l'issue de cette formation, vous serez capable de :

- Connaître l'ensemble des concepts et enjeux du Big Data
- Identifier et s’approprier les outils de l’écosystème BIG DATA
- Gérer des flux de données complexes, non structurés et massifs
- Implémenter des modèles d'analyses statistiques pour répondre aux besoins métiers
- Appréhender un outil de data visualisation pour restituer des analyses dynamiques

POUR QUI ?

- Dataminers
- Chargés d'études statistiques
- Développeurs
- Chefs de projet

Programme détaillé

1 / Présentation du Big Data

- Les origines du Big Data
- Un déluge de données
- Données internes et externes
- Les dimensions en V du Big Data : les 3V (Vélocité, Variété et Volume)
- Les grands acteurs (éditeurs, intégrateurs) et l'Opensource

2 / Enjeux du Big Data

- Les données au cœur des enjeux
- Les nouveaux Métiers
- BIG DATA versus Business Intelligence
- Liens entre Big Data et internet des objets (IoT)
- CNIL et règlementations (GDPR)
- Sécurité IT (tokenisation, anonymisation)
- Concept de Datalake
- Datamanagement, traçabilité des données
- Qualité des données

3 / Les technologies

- Architecture et Composants Hadoop
- Le stockage avec Hadoop (NoSQL, HDFS)
- Fonctionnement de MapReduce et Storm...
- L'écosystème Hadoop (Hortonworks, Cloudera, MapR, Elastic Map Reduce, Biginsights)
- Installation d'Hadoop
- Technologies du datascientist
- Les solutions de visualisation (Tableau, Talend, Qlikview ...)

Innov Systems

4 / Gérer les données structurées et non structurées

- Système de fichiers distribué HDFS : principes
- Importer des données externes vers HDFS
- Requêter avec Hive
- Analyser les données (Pig)
- Les outils permettant d'extraire les données (ETL)
- Streaming Data (NIFI, Kafka, Spark, Storm...)

5 / Technique et méthodes Big data analytics

- Machine Learning
- Régression, Classification et Clustering
- Préparation des données (data preparation, feature engineering)
- Générer des modèles en R ou Python
- Outils du marché : Jupyter Notebook, Dataïku, Amazon Machine Learning...

6 / Visualisation des données

- Besoin de la data visualisation
- Analyse et visualisation des données
- Principes fondamentaux et objectifs de la dataviz
- Limites des outils et méthodes traditionnels (Excel, BusinessObjects, etc)
- Panorama des solutions de datavisualisation : Tableau, Qlik, Microstrategy, Microsoft Power BI...
- Quels types de visualisation pour quels usages ? comparaison QlikView, Tibco Spotfire, Tableau

Innov Systems

Approche pédagogique

- Support Ecrit et Projection
- Exposés Interactifs, Podcasts et Vidéos
- Brainstorming et Jeux de Rôle
- Cas Pratiques et Labs inclus pour leur impact opérationnel
- Test de Validation des Acquis des Connaissances

Prochaines dates programmées

01 au 05 Fév. 2027	Présentiel - Casablanca
29 Mars au 02 Avr. 2027	Présentiel - Casablanca
24 au 28 Mai 2027	Présentiel - Casablanca
19 au 23 Juil. 2027	Présentiel - Casablanca
13 au 17 Sep. 2027	Présentiel - Casablanca

Autres dates possibles sur demande. Contactez-nous pour organiser une session intra-entreprise.

Réservation & Renseignements

Téléphone : +212 522 247 210
Email : contact@innov-systems.com
Web : https//www.innov-systems.com

Innov Systems — Tous droits réservés