

Data Science

Data Mining - Synthèse

2 jours (14h00) | ★★★★★ 4,6/5 | BIG-DM | Évaluation qualitative de fin de stage |
Formation délivrée en présentiel ou distanciel ⁽¹⁾

Formations Informatique > Big Data > Data Science



À l'issue de ce stage vous serez capable de :

- Décrire les principes et la finalité du Data Mining (DM)
- Identifier les principales techniques du DM et leur cas d'utilisation
- Mettre en oeuvre les méthodes de scoring et de géomarketing sur un cas simple
- Différencier les méthodes prédictives et les méthodes descriptives du DM
- Lister les principales étapes d'un projet Data Mining.

Niveau requis

Avoir des connaissances sur l'aide à la décision, l'analyse décisionnelle et des notions en statistiques.

Public concerné

Chefs de projets, DSI, statisticiens et/ou gestionnaires de bases de données.

Cette formation :

- Est animée par un consultant-formateur dont les compétences techniques, professionnelles et pédagogiques ont été validées par des diplômes et/ou testées et approuvées par l'éditeur et/ou par M2i Formation
- Bénéficie d'un suivi de son exécution par une feuille de présence émarginée par demi-journée par les stagiaires et le formateur.

(1) Modalité et moyens pédagogique :

Formation délivrée en présentiel ou distanciel * (e-learning, classe virtuelle, présentiel à distance). Le formateur alterne entre méthodes ** démonstrative, interrogative et active (via des travaux pratiques et/ou des mises en situation). La validation des acquis peut se faire via des études de cas, des quiz et/ou une certification.

Les moyens pédagogiques mis en oeuvre (variables suivant les formations) sont : ordinateurs Mac ou PC (sauf pour les cours de l'offre Management), connexion internet fibre, tableau blanc ou paperboard, vidéoprojecteur ou écran tactile interactif (pour le distanciel). Environnements de formation installés sur les postes de travail ou en ligne. Supports de cours et exercices.

* Nous consulter pour la faisabilité en distanciel. ** Ratio variable selon le cours suivi.

Programme

Introduction au système d'information décisionnel

- Comprendre les principes du DM
- Identifier les principales techniques du DM
- Cycle de projet DM
- Découvrir les méthodes prédictives et descriptives du DM
- Connaître les principales étapes d'un projet DM

Introduction au DM

- Définition du DM
- Avantages du DM
- Inconvénients du DM
- La donnée (structure et cas d'usage)
- Data Warehouse
- DM et OLAP (Online Analytical Processing)

Data Mining Process

- Que peut-on faire avec le DM ?
- Types de jeux de données
- Comment inventorier, décrire et classer les données ?
- Comment concevoir et alimenter la base DM ?
- Le processus de la fouille de données
 - Définition du problème
 - Collecte et préparation des données
 - Construction de modèles et évaluations
 - Déploiement des connaissances

Les outils du DM

- Introduction aux outils d'exploration de données
- Catégories d'outils d'exploration de données
- Outils traditionnels d'exploration de données
- Tableaux de bord
- Outils d'exploration de texte
- Autres applications et programmes

Les techniques du DM

- Introduction aux techniques d'exploration de données
- Réseaux de neurones artificiels
- Arbres de décision
- "Rule induction"
- Algorithmes génétiques
- Méthode du "plus proche voisin"
- Approches de K-Means, K-Medoids et hiérarchiques
- Avantages et inconvénients

Utilité et avenir du DM

- Atelier et échange avec les participants

Modalités d'évaluation des acquis

- En cours de formation, par des études de cas ou des travaux pratiques
- Et, en fin de formation, par un questionnaire d'auto-évaluation ou une certification (M2i ou éditeur)