

Introduction à la data science

Durée: 2 Jours Réf de cours: GKDSAIN

Résumé:

Cette formation a pour objectif de donner une introduction précise et claire des problématiques et outils de la data science. On définit la data science et on explique le métier du data scientist, On présentera les grandes familles de modèles qui sont utilisés en data science, leurs vocations et leurs problématiques. Ces modèles seront illustrés grâce à des business cases et des exemples concrets avec des programmes Python-Sckitlearn et IBM SPSS. Les jeux de données concerneront des domaines divers comme le marketing, la santé et la maintenance prédictive. Les modèles et algorithmes présentés dans ce cours seront développés dans des cours spécifiques du cursus Data Scientist de Global Knowledge.

Public visé:

Cette formation s'adresse à toute personne qui souhaite avoir des connaissances sur la data science, ses enjeux, ses problématiques et outils.

Objectifs pédagogiques:

- A l'issue de la formation, les participants seront capables de :
 - Identifier les grandes familles de modèles utilisés en data science
 - Identifier à la fois les problématiques et les outils de la data science
-

Pré-requis:

Aucun

Après cette formation, nous vous conseillons le(s) module(s) suivant(s):

- GKBDPYT - Python pour la data science
 - GKDAML - Introduction à l'analyse prédictive et cognitive et au Machine Learning
 - GKDPYT - Le langage Python pour Machine Learning
 - GKIADS - Intelligence Artificielle pour la Data Science
 - M-DP100 - Concevoir et mettre en oeuvre une solution de Data Science dans Azure
 - M-DP200 - Mettre en oeuvre une solution de données Azure
 - M-DP201 - Concevoir une solution de données Azure
-

Contenu:

INTRODUCTION A LA DATA SCIENCE

- Définition, historique et enjeux
- Domaines d'application et problématiques de la Data Science
- Environnement Big Data pour la Data Science
- Métier de la data science : journée type du data scientist
- Les acteurs et les outils du marché pour la data science

COMPETENCES TECHNIQUES POUR LA DATA SCIENCE

Extraction manipulation des données

- Objectif et démarche
- Exemple pratique

Exploration et visualisation des données

- Objectif et démarche
- Exemple pratique

La modélisation

- Objectif et démarche
- Les grandes familles de modèles et leurs domaines d'application :
- Classifications supervisée et non supervisée et système de recommandations
- Modèles mathématique statistiques
- Machine learning et modèles cognitifs
- Exemples pratiques

Modèles de classification supervisée :

- Domaine d'application
- Programmes et langages pour les modèles
- Introduction à quelques modèles :
- Principe du modèle de la régression
- Principe du modèle des Réseaux bayésiens
- Principe du modèle d'Arbre de décision et Random Forest
- Exemples pratiques avec Python Scikitlearn et IBM SPSS

Modèle de classification non supervisée :

- Domaine d'application
- Programmes et langages pour les modèles
- Introduction à quelques modèles : Principe du modèle K-means, Principe du modèle de Réseaux de neurones
- Exemples pratiques avec Python Scikitlearn et IBM SPSS

Modèles de recommandation et analyse d'association :

- Domaine d'application
- Programmes et langages pour les modèles
- Introduction à quelques modèles : Principe du filtrage par le contenu, Principe du filtrage Collaboratif
- Exemples pratiques avec Python Scikitlearn et IBM SPSS

Méthodes pédagogiques :

Support de cours en français remis aux participants

Autres moyens pédagogiques et de suivi:

- Compétence du formateur : Les experts qui animent la formation sont des spécialistes des matières abordées et ont au minimum cinq ans d'expérience d'animation. Nos équipes ont validé à la fois leurs connaissances techniques (certifications le cas échéant) ainsi que leur compétence pédagogique.
- Suivi d'exécution : Une feuille d'émargement par demi-journée de présence est signée par tous les participants et le formateur.
- Modalités d'évaluation : le participant est invité à s'auto-évaluer par rapport aux objectifs énoncés.
- Chaque participant, à l'issue de la formation, répond à un questionnaire de satisfaction qui est ensuite étudié par nos équipes pédagogiques en vue de maintenir et d'améliorer la qualité de nos prestations.

Délais d'inscription :

- Vous pouvez vous inscrire sur l'une de nos sessions planifiées en inter-entreprises jusqu'à 5 jours ouvrés avant le début de la formation sous réserve de disponibilité de places et de labs le cas échéant.
- Votre place sera confirmée à la réception d'un devis ou ""booking form"" signé. Vous recevrez ensuite la convocation et les modalités d'accès en présentiel ou distanciel.
- Attention, si vous utilisez votre Compte Personnel de Formation pour financer votre inscription, vous devrez respecter un délai minimum et non négociable fixé à 11 jours ouvrés.