



Le langage Python pour Machine Learning

Durée: 3 Jours **Réf de cours: GKDPYT**

Résumé:

La Data Science nécessite des compétences spécifiques afin de prévoir le comportement des clients ou d'anticiper des tendances. Cette formation vous permettra de découvrir l'analyse de données ainsi que le Machine Learning via Python, les différentes étapes qui interviennent lors de l'élaboration d'un modèle statistique, ainsi que l'analyse textuelle et la Data Visualisation.

Public visé:

Toute personne intéressée par une prise en main rapide du langage Python dans le domaine de la data science. Data Miner, CDO, DSI, développeurs en charge de la mise en production des modèles, Statisticiens.

Objectifs pédagogiques:

- Comprendre l'enjeu stratégique des données pour l'entreprise
 - Comprendre le principe de la création d'un modèle statistique
 - Préparer les données
 - Évaluer la performance d'un modèle en Data Science
 - Choisir entre régression, clustering et classification en fonction des données
 - Data Visualisation
-

Pré-requis:

- Connaissance en statistiques et Data Mining
 - Connaissance en programmation Python
-

Contenu:

Le langage Python

- C'est quoi ?
- Pourquoi Python
- Les principaux packages

Installation

Shell, IDE, NoteBook

- Interpréteur I Python

La base de langage Python

Premiers pas

- Structure de données
- Fonction
- Fichier

Python et la donnée

- NumPy, Pandas
- Chargement, stockage de données
- Interaction avec les données

Data Pré-Processing

- Manipulation des valeurs vides
- Transformation de données
- Expressions régulières

Data Wrangling

- Indexation de données
- Jointure ; combinaison
- Remodelage
- Agrégation ; Grouping

Data Exploring

- Explorations visuelles (Graphes)
- Explorations statistiques

Machine Learning

- Les types de Machine Learning :
L'apprentissage supervisé,
L'apprentissage non-supervisé,
L'apprentissage renforcé
- Les Méthodes de Machine Learning
- Les étapes de l'élaboration de modèle
- Travaux pratiques sur les méthodes de
Machine Learning : Introduction à
Scikit-Learn, Use cases

Text Mining

- Collection de données textuelles
- Préparation de données
- Étiquetage grammatical, analyse
syntaxique, analyse sémantique.

Data Visualisation

- Package Python
- Autres technologies

Méthodes pédagogiques :

Appréciation : évaluation qualitative de fin de stage

Modalités et moyens pédagogiques : démonstrations, cas pratiques, synthèse et évaluation des acquis

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.