

Scikit-learn pour le machine learning

Description

Une formation scikit-learn pour le machine learning vous permettant de maîtriser l'utilisation du package scikit-learn, référence mondiale pour le machine learning en python. Par des cas pratiques développés dans des notebooks, vous découvrirez toutes les spécificités de cet outil complet pour la construction de modèles de machine learning. Cette formation de deux jours convient parfaitement à la formation d'utilisateur de python désireux de se perfectionner à l'utilisation de scikit-learn.

Cette formation scikit-learn pour le machine learning vous permet d'approfondir l'utilisation et de maîtriser les bonnes pratiques du package scikit-learn.

Formation en petits groupes (maximum 7 participants) pour assurer un échange optimal avec le formateur.

Durée

2.00 jours (14.00 heures)

Profils des stagiaires et prérequis

Profils :

- Data analyst ou data scientist désireux de monter en compétence sur scikit-learn

Prérequis :

- Bases en python et en traitement de données ou avoir suivi la session Python pour la data science

Objectifs pédagogiques

- Connaître les principaux concepts du machine learning
- Construire un pipeline de machine learning avec scikit-learn
- Maîtriser les étapes de création d'un modèle de machine learning (de la préparation à la validation)

Programme détaillé

- Introduction
 - Les concepts du machine learning
 - Le processus de construction d'un modèle de machine learning
 - L'écosystème de développement autour de scikit-learn (python, numpy, pandas...)
- Utilisation de Scikit-Learn
 - Les principes de scikit-learn pour les modèles de machine learning
 - La notion de pipeline en machine learning
 - Un premier modèle de machine learning avec scikit-learn
- Préparation des données et preprocessing
 - Comment préparer des données pour un modèle de machine learning ?
 - Les preprocessing de scikit-learn
 - Le traitement des données qualitatives
 - Le traitement des données manquantes

Stat4decision

37-39 avenue Ledru Rollin Paris 75012

Tel. 01.72.25.40.82 | E-mail : info@stat4decision.com | www.stat4decision.com

Numéro SIRET: 81048985600015 | Numéro de déclaration d'activité: 11755352275 (auprès du préfet de région de: 75)

Organisme de formation certifié Qualiopi pour ses actions de formation

- L'utilisation du ColumnTransformer de scikit-learn
- Sélection de meilleur modèle avec scikit-learn
 - Gérer le sur-apprentissage
 - Validation du modèle (train / test, validation croisée...)
 - Les métriques d'évaluation des modèles de machine learning
- L'ajustement des hyperparamètres avec scikit-learn
 - Gestion automatique
 - Combinaison dans un pipeline
- Focus sur certaines méthodes de machine learning
 - Modèles linéaires
 - Modèles basés sur des arbres (forets aléatoires, GBM...)
 - Autre modèles
- Le cas des modèles non supervisés dans scikit-learn
 - Mise en œuvre de modèle de machine learning non supervisés avec scikitlearn et comparaisons.
- Passage en production d'un modèle de machine learning avec scikit-learn - les différentes approches.

Organisation de la formation

Moyens pédagogiques et techniques

- Documents supports de formation projetés.
- Exposés théoriques
- Etude de cas concrets
- Mise à disposition en ligne de documents supports à la suite de la formation.
- Utilisation de Notebooks Jupyter et d'un environnement JupyterHub pour les exercices pratiques.

Dispositif de suivi de l'exécution et d'évaluation des objectifs de la formation

- Mises en situation.
- Questions orales ou écrites (QCM).
- Exercices effectués dans des notebooks

Stat4decision

37-39 avenue Ledru Rollin Paris 75012

Tel. 01.72.25.40.82 | E-mail : info@stat4decision.com | www.stat4decision.com

Numéro SIRET: 81048985600015 | Numéro de déclaration d'activité: 11755352275 (auprès du préfet de région de: 75)

Organisme de formation certifié Qualiopi pour ses actions de formation