

Industrialisation d'un Projet Data en Python

Description

Découvrez comment passer en production un projet data avec cette formation industrialisation de projets data en python.

Durée

3.00 jours (21.00 heures)

Profils des stagiaires et prérequis

Profils :

- Data analyst et data scientist désirant acquérir des compétences en industrialisation
- Développeur Python souhaitant accompagner la mise en production de modèles

Prérequis :

- Connaissances de base en traitement de données (statistique et tables de données)
- Connaissances de base du langage Python

Objectifs pédagogiques

- Découvrir les techniques de Software Engineering appliquées à la Data Science
- Créer une version packagée et testée d'un modèle simple
- Réaliser les différentes étapes d'une mise en production d'un modèle de Data Science

Programme détaillé

- Qu'est-ce que l'industrialisation ?
 - Quelles sont les grandes thématiques de l'industrialisation d'un projet ?
 - Pourquoi et pour qui industrialiser ?
- Le version control, et les outils qui l'entourent (Gitlab, Github)
- Découverte des principes de clean code
 - Application des différentes règles dans le contexte d'une MR
 - Écriture d'un code en suivant des standards de qualité élevés
 - Application des principes de clean code aux notebooks
- Découverte des outils de vérification statique
 - Utilisation d'un linter
 - Utilisation du typage statique
- Les différents types de tests, et leur application au contexte de la data science
 - Écriture d'un test unitaire, intégration et e2e
 - Application des principes du testing à la thématique ML
 - Écriture d'un pipeline CI / CD lançant plusieurs types de test

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

- Les outils de documentation et leur principe
 - Documentation automatique du code avec Sphinx
 - La documentation comme substitut à la revue scientifique à partir d'outils collaboratifs tels que Slite
- Découverte du processus de recette d'un algorithme
- Mise en production d'un algorithme
 - Rédaction et application d'un processus de recette d'un algorithme
 - Description du principe d'itération
- Suivi de la qualité et monitoring des performances d'un algorithme
- Mise en production et dockerisation d'un projet

Organisation de la formation

Moyens pédagogiques et techniques

- Supports sous forme de Jupyter Notebooks
- Nombreux cas pratiques

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

- Évaluation par des exercices par le formateur
- Quiz en classe

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