MLOps : industrialisation d'un projet de Data Science

Adopter les bonnes pratiques de développement pour la Data Science

Stage pratique
MLOps

Code :
DSIND

Durée :
2 jour(s) (14,00 heures)

Exposés : **10 %**
Cas pratiques : **80 %**
Echanges d'expérience : **10 %**

Inter-entreprises :
Prochaines sessions disponibles sur notre site web.
Tarif : 1 920,00 € HT / participant

Intra-entreprise :
Tarifs et dates sur demande.

DESCRIPTION

La maturité du marché sur les sujets de la Data Science ne cesse de grandir chaque jour.

C'est pourquoi, il est important d'être capable de déployer ses projets en production, rapidement, pour **apporter de la valeur à ses utilisateurs**.

Durant cette formation, vous profiterez de la riche expérience de nos consultants pour découvrir la **Data Science industrialisée**. Vous vous approprierez les bonnes pratiques issues du **software craftsmanship** essentielles pour rendre votre projet maintenable et qui vous permettront par la suite d'initier une approche MLOps.

OBJECTIFS PEDAGOGIQUES

- Découvrir les concepts de Software Craftsmanship appliqués à la Data Science
- Passer d'un notebook à un package testé, documenté et versionné
- Constituer le socle technique requis pour accompagner le déploiement en production et aborder la pratique MLOps

PUBLIC CIBLE

- Analyste
- Statisticien
- Data Scientist
- Data Engineer
- Machine Learning Engineer
- Développeur

PRE-REQUIS

- Connaissance des bases de la Data Science (modèles, biais, variance, etc.)
- Connaissance des librairies de manipulation de donnée en python (pandas, numpy, etc.)
- Connaissance de la ligne de commande linux (bash par exemple)

METHODE PEDAGOGIQUE

Une formation pratique durant laquelle les participants partent d'un

code d'exploration préparé par le formateur et qu'ils vont amener en production, étape par étape, en démontrant que la transition doit se faire de manière la plus fluide possible.

PROFIL DES INTERVENANTS

Cette formation est dispensée par un·e ou plusieurs consultant·es d'OCTO Technology ou de son réseau de partenaires, expert·es reconnus des sujets traités.

Le processus de sélection de nos formateurs et formatrices est exigeant et repose sur une évaluation rigoureuse leurs capacités techniques, de leur expérience professionnelle et de leurs compétences pédagogiques.

MODALITÉS D'ÉVALUATION ET FORMALISATION À L'ISSUE DE LA FORMATION

L'évaluation des acquis se fait tout au long de la session au travers des ateliers et des mises en pratique.

Afin de valider les compétences acquises lors de la formation, un formulaire d'auto-positionnement est envoyé en amont et en aval de celle-ci.

En l'absence de réponse d'un ou plusieurs participants, un temps sera consacré en ouverture de session pour prendre connaissance du positionnement de chaque stagiaire sur les objectifs pédagogiques évalués.

Une évaluation à chaud est également effectuée en fin de session pour mesurer la satisfaction des stagiaires et un certificat de réalisation leur est adressé individuellement.

PROGRAMME PEDAGOGIQUE DETAILLE

Jour 1

INTRODUCTION

Rôle du data scientist

Convictions

INSTALLATION DES ENVIRONNEMENTS

Jupyter Notebook

PyCharm

Environnement anaconda

CLEAN CODE ET NOTEBOOK

Nettoyer son notebook

Noms des variables

Utilisation de fichiers de configuration

Programmation fonctionnelle

Immutabilité

Créer des fonctions

Mise en pratique : "Appliquer les techniques de Clean code à un notebook"

TESTER SON CODE

Découverte du TDD

Initiation à unittest

- Les classes de test
- Écrire son premier test
- Les méthode setup et teardown

Intégration à setuptools

Mise en pratique : "Écrire ses premiers tests"

Jour 2

RAPPELS ET RÉVISION DU JOUR 1

DOCUMENTER SON PROJET

Découverte de Sphinx

Découverte de sphinx-quickstart

Mise en pratique : "Documenter son projet"

GÉRER LES VERSIONS

Introduction à Git

Gérer les versions de code

Utilisation des tags

Stratégies de versionning

Gérer des datasets et des modèles

RENDRE SON CODE DÉPLOYABLE

Définition du packaging

Introduction à setuptools

Installer son code en local

Gestion des dépendances

Règles simples de démarrage

Mise en pratique : "Créer et installer son package"

Bien séparer entraînement et inférence

Sauvegarder les objets de préparations de données

Mise en pratique : "Séparer train & predict"

CONCLUSION

Synthèse des points abordés

Partage sur la formation

Questions/Réponses additionnelles

Accessibilité

L'inclusion est sujet important pour OCTO Academy.
Nos référent·es sont à votre disposition pour faciliter l'adaptation de votre formation à vos besoins spécifiques.
Pour les contacter : academy.accessibilite@octo.com