MLOps : mise en production d'un projet de Data Science

Favoriser le déploiement des changements de code, de données et de modèles en production

DESCRIPTION

L'industrialisation du code du projet de Data Science n'est qu'une étape parmi d'autres pour réaliser la mise en production du système et le faire évoluer. En effet, une fois ce premier pas franchi, il s'agira alors de mettre en place un **mécanisme d'orchestration**, mais aussi d'être **capable de déployer régulièrement**, et de **suivre le comportement du système sur la durée**.

Aujourd'hui encore, trop de modèles de Machine Learning ne sont pas déployés en production et demeurent dans les data labs. Une solution à ce problème consiste à **réunir les équipes de data science et d'informatique autour du concept de MLOps**.

Tout au long de cette formation, vous profiterez de la riche expérience de nos consultants-formateurs pour mettre en œuvre les pratiques de la Data Science en production.

Vous vous approprierez les outils et le **processus transversal**, **collaboratif** et **itératif** MLOps pour **réduire le délai de votre mise sur le marché** tout en **améliorant continuellement la valeur ajoutée** pour vos utilisateurs. Le MLOps prend notamment en charge l'intégration continue, ainsi que le déploiement rapide et reproductible des modèles.

OBJECTIFS PEDAGOGIQUES

- Identifier et utiliser les pratiques clés MLOps pour industrialiser son projet de Machine Learning
- Savoir découper un projet de Machine Learning en différentes briques et les organiser
- Construire l'architecture applicative d'un projet Data Science
- Déployer son application de Machine Learning en production et interagir avec
- Maintenir un flux continu de données pour alimenter son application de Machine Learning
- Mesurer et évaluer les performances de son application de Machine Learning

PUBLIC CIBLE

- Analyste
- Statisticien
- Data Scientist

Stage pratique
MLOps

Code :
DSIN2

Durée :
3 jour(s) (21,00 heures)

Exposés : **40 %**
Cas pratiques : **50 %**
Echanges d'expérience : **10 %**

Inter-entreprises :
Prochaines sessions disponibles sur notre site web.
Tarif : 2 510,00 € HT / participant

Intra-entreprise :
Tarifs et dates sur demande.

- Data Engineer
- Machine Learning Engineer
- Développeur

PRE-REQUIS

- Connaissances de base en programmation et en scripting
- Avoir suivi la formation "MLOps : industrialisation d'un projet de Data Science" (DSIND) ou savoir tester une application python
- Avoir suivi la formation "Fondamentaux de la Data Science" (DSFDX) ou savoir modéliser et développer un problème de Data Science simple

METHODE PEDAGOGIQUE

Une formation alternant théorie et pratique durant laquelle les participants partent d'un modèle de machine learning construit par le formateur qu'ils vont amener en production, étape par étape, en découvrant l'ensemble des briques techniques à mettre en place.

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

CLARIFICATION DE L'APPROCHE MLOPS

INTRODUCTION À LA CI / CD

- Les différents environnements nécessaires
- L'intérêt de l'intégration continue et du déploiement continu
- Les outils de CI / CD existants

PYRAMIDE DE TESTS

- Description des différents types de tests
- Focus sur les tests d'intégrations et fonctionnels
- Présentation de framework de tests

MISE EN PRATIQUE : "METTRE EN PLACE UN OUTIL D'INTÉGRATION CONTINUE"

CLOUD & INFRASTRUCTURE AS CODE

- Introduction au cloud et aux différents providers
- Choix d'un cloud provider pour les travaux pratiques
- Infrastructure as code, principe et découverte de Docker, Ansible et Terraform

ORCHESTRATION

- Le besoin d'orchestration dans un système de Data Science
- Présentation de différents orchestrateurs
- Mise en pratique : "Passer d'un script Python à une série de tâches orchestrées avec Airflow"

Jour 2

ARTEFACTS

- Patterns d'architecture technique pour favoriser la mise en

production et la maintenabilité du projet
- Identification des artefacts (modèles, données, code…)
- Cycles de vie des artefacts
- Mise en pratique : mise en place de MLFlow

DÉPLOIEMENT

- La démarche
- Le couplage / découplage
- Les stratégies de déploiement

EXPOSITION

- Méthodes pour exposer un modèle de Data Science
- Mise en pratique : “Interagir avec un modèle existant”

Jour 3

MISE EN PRATIQUE : “PASSER D'UN MODÈLE ONE-SHOT À UN MODÈLE QUI PRÉDIT TOUTES LES MINUTES”

- A partir d'un flux de données disponibles, mettre en place les outils nécessaires pour faire une prédiction par minute

FEEDBACK LOOPS

- L'intérêt des boucles de feedback
- Identifier les différentes boucles de feedbacks
- Mettre en place les différentes boucles de feedbacks

MONITORING

- L'enjeu du monitoring
- Les indicateurs à suivre
- Les pratiques pour mettre en place un monitoring utile tout en respectant les contraintes de SLA

AMÉLIORATION CONTINUE

- Choisir le bon moment pour réentraîner un modèle
- Les stratégies à mettre en place pour s'assurer que le modèle en production est toujours pertinent
- MLops

- Qualité de données pour un projet de *Machine Learning*

BILAN ET CLÔTURE

- Partage et retour d'expérience
- Questions / réponses
- Évaluation de la session

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