



Big Data - MLOps et DataOps

Microsoft Azure - Concevoir, déployer et surveiller un pipeline MLOps

5 jours (35h00) | 9 4,6/5 | MLOPS-AZU | Évaluation qualitative de fin de stage | Formation délivrée en présentiel ou distanciel

Formations Informatique > Data > Big Data - MLOps et DataOps

Contenu mis à jour le 13/10/2023. Document téléchargé le 27/07/2024.

Objectifs de formation

A l'issue de cette formation, vous serez capable de :

- Concevoir et tester un modèle de ML (Machine Learning) sur Microsoft Azure
- Concevoir et tester un pipeline MLOps robuste
- Déployer et automatiser votre pipeline MLOps de production et le passer à l'échelle
- Surveiller votre modèle de production et corriger les écarts.

Modalités, méthodes et moyens pédagogiques

Formation délivrée en présentiel ou distanciel* (blended-learning, e-learning, classe virtuelle, présentiel à distance).

Le formateur alterne entre méthode** démonstrative, interrogative et active (via des travaux pratiques et/ou des mises en situation).

Variables suivant les formations, les moyens pédagogiques mis en oeuvre sont :

- Ordinateurs Mac ou PC (sauf pour certains cours de l'offre Management), connexion internet fibre, tableau blanc ou paperboard, vidéoprojecteur ou écran tactile interactif (pour le distanciel)
- Environnements de formation installés sur les postes de travail ou en ligne
- Supports de cours et exercices

En cas de formation intra sur site externe à M2i, le client s'assure et s'engage également à avoir toutes les ressources matérielles pédagogiques nécessaires (équipements informatiques...) au bon déroulement de l'action de formation visée conformément aux prérequis indiqués dans le programme de formation communiqué.

* nous consulter pour la faisabilité en distanciel

** ratio variable selon le cours suivi

Prérequis

Avoir des connaissances en bases de données et/ou analytiques.

Public concerné

Data Scientists, Data Analysts, ingénieurs Data, ingénieurs DevOps, développeurs, architectes Data, chefs de projets, managers.

Cette formation :

- Est animée par un consultant-formateur dont les compétences techniques, professionnelles et pédagogiques ont été validées par des diplômes et/ou testées et approuvées par l'éditeur et/ou par M2i Formation
- Bénéficie d'un suivi de son exécution par une feuille de présence émarginée par demi-journée par les stagiaires et le formateur.

Programme

Jour 1

Fondamentaux

- Qu'est-ce qu'une donnée ?
- Cycle de vie d'une donnée
- Cycle de vie de la Data Science
- Evolution des analytiques
- Apports des analytiques et du Cloud
- Qu'est-ce qu'une organisation Data Driven ?
- Gouvernance et qualité des données
- Lean, Agilité et DevOps, MLOps appliqués aux données
- Conteneurisation et architectures micro-services
- Apports de MLOps
- Workflow de MLOps
- DevOps vs DataOps vs MLOps
- Machine Learning et opérations : briser le mur de la confusion
- Constituer et former une équipe MLOps

Exemple de travaux pratiques (à titre indicatif)

- Mise en pratique : création d'un pipeline MLOps simple

Jour 2

Mise en place de l'infrastructure et des outils

- Prise en main d'Azure Machine Learning
- Prise en main d'Azure DevOps
- Installation d'Apache MLflow
- Installation de JupyterHub
- Les 10 principes de gestion du code pour le ML
- Prétraitement des données (qualité, attributs, corrélations, filtrage)
- Enregistrement et versioning des données dans le Feature Store

Jour 3

Pipelines de ML

- Fondamentaux des pipelines de ML
- Ingestion des données et réduction de dimension
- Apprentissage (SVM, forêts aléatoires) et optimisation des hyperparamètres

- Métriques d'évaluation de modèles
- Evaluation des modèles
- Tests de production :
 - Tests batch
 - A/B testing
 - Stage test vs shadow test
 - Tests en CI-CD
- Packaging du modèle : enregistrement et production des artefacts
- Paquetage vs conteneurisation
- Génération du micro-service et déploiement
- Modèles prêts pour l'inférence : importation et chargement des artefacts

Déployer des modèles de ML à l'échelle

- Principes clés du déploiement de systèmes de ML
- Comprendre les types d'inférence de ML en production
- Mapper l'infrastructure à la solution de ML
- Déploiement du modèle avec ACI (Azure Container Instance)
- Déploiement du modèle avec AKS (Azure Kubernetes Service)
- Déploiement du service avec MLflow
- Comprendre le besoin de CI-CD

Jour 4

Construire des pipelines de ML robustes

- CI et CD dans MLOps
- Créer le pipeline de CI-CD et l'environnement de test avec Azure DevOps
- Exécution et test du pipeline
- "Triggers" de pipeline
- Comprendre les API micro-services
- Implémentation et serving d'un modèle comme API
- Créer un micro-service avec Docker
- Tester l'API
- Sécuriser son déploiement face à diverses attaques
- Mettre en place l'infrastructure de production avec AML workspace et SDK (Software Development Kit)
- Mettre en place l'environnement de production dans le pipeline CI-CD
- Tester le pipeline de production
- Configurer les "triggers" pour l'automatisation (Git trigger, Artifactory trigger, Schedule trigger)
- Gestion de la release du pipeline

Jour 5

Monitoring de modèle de ML en production

- Principes clés de la surveillance du modèle : Model Drift, Model Bias, Explainable AI...
- Monitoring dans le workflow MLOps
- Activer le monitoring continu du service
- Activer le réapprentissage manuel ou automatique du modèle
- Gouvernance : alertes, actions, assurance qualité, contrôle

Le contenu de ce programme peut faire l'objet d'adaptation selon les niveaux, prérequis et besoins des apprenants.

Modalités d'évaluation des acquis

- En cours de formation, par des études de cas ou des travaux pratiques
- Et, en fin de formation, par un questionnaire d'auto-évaluation

Accessibilité de la formation

Le groupe M2i s'engage pour faciliter l'accessibilité de ses formations. Les détails de l'accueil des personnes en situation de handicap sont consultables sur la page Accueil et Handicap.

Modalités et délais d'accès à la formation

Les formations M2i sont disponibles selon les modalités proposées sur la page programme. Les inscriptions sont possibles jusqu'à 48 heures ouvrées avant le début de la formation. Dans le cas d'une formation financée par le CPF, ce délai est porté à 11 jours ouvrés.