

Maîtriser le MLOps moderne avec Databricks

Concevoir, déployer et superviser des modèles Machine Learning à l'échelle

DESCRIPTION

Accélérez l'industrialisation de vos modèles avec le MLOps sur Databricks.

Dans un contexte où l'intelligence artificielle s'impose comme un levier stratégique de performance, cette formation vous permet d'ancrer les bonnes pratiques du MLOps au cœur de vos projets Data et ML sur la plateforme Databricks.

Pensée pour les équipes Data Science, Data Engineering et Machine Learning, elle couvre l'ensemble du cycle de vie des modèles : de l'expérimentation à grande échelle jusqu'au déploiement opérationnel, en intégrant les standards de traçabilité, reproductibilité et gouvernance.

Les participants apprennent à :

- Piloter efficacement leurs modèles avec MLflow et Unity Catalog, véritables tours de contrôle du Machine Learning.
- Industrialiser les pipelines d'entraînement et de déploiement avec Databricks Asset Bundle et le CLI Databricks.
- Gérer les features à grande échelle avec Feature Store et orchestrer des déploiements batch ou temps réel via Model Serving.
- Mettre en œuvre des stratégies de rollout avancées (A/B testing, canary) et assurer le suivi continu des modèles grâce à Lakehouse Monitoring.

Grâce à une approche résolument pratique et orientée résultats, cette formation vous aide à accélérer vos cycles d'expérimentation, fiabiliser vos processus de déploiement et renforcer la gouvernance de vos modèles à l'échelle de l'entreprise.

OBJECTIFS PEDAGOGIQUES

- Identifier et analyser les principes clés du MLOps sur Databricks afin de comparer les différentes approches, en mettant en évidence leurs avantages opérationnels et leurs limites techniques.
- Concevoir, implémenter et superviser des modèles de machine learning sur Databricks en mobilisant les outils natifs pour mesurer les performances et ajuster les hyperparamètres en vue d'une optimisation continue.
- Concevoir et mettre en production des modèles de machine learning robustes sur Databricks, en argumentant les choix

Stage pratique

MLOps

Code :

MLODB

Durée :

2 jour(s) (14,00 heures)

Exposés : **20 %**

Cas pratiques : **50 %**

Echanges d'expérience : **30 %**

Inter-entreprises :

Prochaines sessions disponibles [sur notre site web](#).

Tarif : 1 800,00 € HT / participant

Intra-entreprise :

Tarifs et dates sur demande.

méthodologiques et en évaluant la performance et la résilience du déploiement en environnement réel.

- Automatiser et industrialiser les workflows de machine learning sur Databricks en intégrant les solutions de gestion de pipelines pour accroître l'efficacité, la traçabilité et la reproductibilité des processus.

PUBLIC CIBLE

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

PRE-REQUIS

- Connaissances de base en programmation et en scripting (Python notamment)
- Avoir suivi la formation "Fondamentaux de la Data Science" (DSFDX) ou savoir modéliser et développer un problème de Data Science simple

METHODE PEDAGOGIQUE

Formation avec apports théoriques, échanges sur les contextes des participants et retours d'expérience pratique des formateurs, complétés de travaux pratiques et de mises en situation.

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. 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

OUVERTURE DE SESSION

- Accueil des participants et tour de table des attentes
- Présentation du programme et déroulé de la formation

PRÉPARATION DES DONNÉES POUR LE MACHINE LEARNING

- Gestion et exploration de la data
- Pré-traitement des données et feature engineering pour le Machine Learning
- Utilisation d'un feature store
- Mises en pratique :
 - "Préparer les données"
 - "Instancier la feature table"
 - "Manipuler via l'API"
 - "Enrichir les features (merge)"
 - "Nettoyer son environnement"

DÉVELOPPEMENT DE MODÈLES POUR LE MACHINE LEARNING

- Workflow de développement de modèles avec MLflow
- Tuning et optimisation de modèles avec Optuna
- Mises en pratique :
 - "Connecter le feature store"
 - "Traçabilité rigoureuse (MLflow Tracking)"
 - "Enregistrement d'artefacts visuels"
 - "Exploration de l'UI"
 - "Définir un espace de recherche"
 - "Intégrer Optuna et MLflow (Nested Runs)"
 - "Identifier le modèle gagnant"
 - "Enregistrer le modèle en production"
- Automatisation du Machine Learning avec AutoML

MACHINE LEARNING ET DÉPLOIEMENT

- Conception et gestion des pipelines et traitements batch pour le Machine Learning
- Mises en pratique
 - "Préparation et création de la Feature table"
 - "Entraînement lié et inférence Batch"
 - "Optimisation des performances (Liquid Clustering)"
- Déploiement en temps réel avec les model serving
- Mises en pratique
 - "Approche UI (simple et visuelle)"
 - "Approche API"

Jour 2

MLOPS : GESTION OPÉRATIONNELLE DES MODÈLES DE MACHINE LEARNING

- Pipeline MLOps : contrôle de la qualité des données et analyse des features
- Monitoring et visualisation des données Lakehouse via un tableau de bord
- Mises en pratique
 - "Définir la référence (Baseline)"
 - "Ingérer et parser les requêtes de production (Inference tables)"
 - "Consolider les données"
 - "Activer le Lakehouse Monitor (via SDK)"

MLOPS AVANCÉ : DÉPLOIEMENT, MONITORING ET OPTIMISATION DES MODÈLES DE ML

- Méthodes et bonnes pratiques de déploiement de Machine Learning sur Databricks
- Création d'assets de Machine Learning en tant que code avec Databricks Asset Bundle (DAB)
- Mises en pratique sur mlops-stacks (standard industriel de Databricks)
 - "Initialiser le projet"
 - "Explorer l'architecture"
 - "Déployer et exécuter"
- Retours d'expérience

CLÔTURE DE SESSION

- Revue des concepts clés et bilan
- Echange autour des questions et 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