

## Build machine learning solutions using Azure Databricks (DP-3014)

Durée: 1 Jour    Réf de cours: M-DP3014    Méthodes d'apprentissage: Classe à distance

### Résumé:

Built as a joint effort by Microsoft and the team that started Apache Spark, Azure Databricks provides data science, engineering, and analytical teams with a single platform for big data processing and machine learning. In this course, you'll learn how to use Azure Databricks to train and deploy machine learning models.

Classe à Distance - site Client

Cette formation peut être suivie à distance en synchrone depuis n'importe quel site pourvu d'une connexion internet (2 Mb/s en symétrique recommandés). Le programme (théorie et pratique) suit le même déroulé pédagogique qu'en présentiel. La solution technologique adoptée permet aux apprenants à distance de suivre les présentations faites au tableau, de voir et d'entendre l'instructeur et les participants en temps réel, mais également d'échanger avec eux.

### Public visé:

Data scientists and machine learning engineers.

### Objectifs pédagogiques:

- Students will learn to,
- Explore Azure Databricks
- Use Apache Spark in Azure Databricks
- Train a machine learning model in Azure Databricks
- Use MLflow in Azure Databricks
- Tune hyperparameters in Azure Databricks
- Use AutoML in Azure Databricks
- Train deep learning models in Azure Databricks
- Manage machine learning in production with Azure Databricks

### Pré-requis:

- This learning path assumes that you have experience of using Python to explore data and train machine learning models with common open source frameworks, like Scikit-Learn, PyTorch, and TensorFlow. Consider completing the Create machine learning models learning path before starting this one.

## Contenu:

### Module 1 : Explore Azure Databricks

- Provision an Azure Databricks workspace.
- Identify core workloads and personas for Azure Databricks.
- Use Data Governance tools Unity Catalog and Microsoft Purview
- Describe key concepts of an Azure Databricks solution.

### Module 2 : Use Apache Spark in Azure Databricks

- Describe key elements of the Apache Spark architecture.
- Create and configure a Spark cluster.
- Describe use cases for Spark.
- Use Spark to process and analyze data stored in files.
- Use Spark to visualize data.

### Module 3 : Train a machine learning model in Azure Databricks

- Prepare data for machine learning
- Train a machine learning model
- Evaluate a machine learning model

### Module 4 : Use MLflow in Azure Databricks

- Use MLflow to log parameters, metrics, and other details from experiment runs.
- Use MLflow to manage and deploy trained models.

### Module 5 : Tune hyperparameters in Azure Databricks

- Use the Hyperopt library to optimize hyperparameters.
- Distribute hyperparameter tuning across multiple worker nodes.

### Module 6 : Use AutoML in Azure Databricks

- Use the AutoML user interface in Azure Databricks
- Use the AutoML API in Azure Databricks

### Module 7 : Train deep learning models in Azure Databricks

- Train a deep learning model in Azure Databricks
- Distribute deep learning training by using the Horovod library

### Module 8 : Manage machine learning in production with Azure Databricks

- Automate feature engineering and data pipelines
- Model development and training
- Model deployment strategies
- Model versioning and lifecycle management

## Autres moyens pédagogiques et de suivi:

- Compétence du formateur : Les experts qui animent la formation sont des spécialistes des matières abordées et ont au minimum cinq ans d'expérience d'animation. Nos équipes ont validé à la fois leurs connaissances techniques (certifications le cas échéant) ainsi que leur compétence pédagogique.
- Suivi d'exécution : Une feuille d'émargement par demi-journée de présence est signée par tous les participants et le formateur.
- En fin de formation, le participant est invité à s'auto-évaluer sur l'atteinte des objectifs énoncés, et à répondre à un questionnaire de satisfaction qui sera ensuite étudié par nos équipes pédagogiques en vue de maintenir et d'améliorer la qualité de nos prestations.

### Délais d'inscription :

- Vous pouvez vous inscrire sur l'une de nos sessions planifiées en inter-entreprises jusqu'à 5 jours ouvrés avant le début de la formation sous réserve de disponibilité de places et de labs le cas échéant.
- Votre place sera confirmée à la réception d'un devis ou "\*\*\*\*\*booking form\*\*\*\*\*" signé. Vous recevrez ensuite la convocation et les modalités d'accès en présentiel ou distanciel.
- Attention, si cette formation est éligible au Compte Personnel de Formation, vous devrez respecter un délai minimum et non négociable fixé à 11 jours ouvrés avant le début de la session pour vous inscrire via [moncompteformation.gouv.fr](http://moncompteformation.gouv.fr).

### Accueil des bénéficiaires :

- En cas de handicap : plus d'info sur [globalknowledge.fr/handicap](http://globalknowledge.fr/handicap)
- Le Règlement intérieur est disponible sur [globalknowledge.fr/reglement](http://globalknowledge.fr/reglement)