

DP-3014: Implementing a Machine Learning solution with Azure Databricks

Course Outline

1) Explore Azure Databricks

- Provision an Azure Databricks workspace
- Identify core workloads and personas for Azure Databricks
- Describe key concepts of an Azure Databricks solution
- Lab: Explore Azure Databricks

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
- Lab: Use Spark in Azure Databricks

3) Train a machine learning model in Azure Databricks

- Prepare data for machine learning
- Train a machine learning model
- Evaluate a machine learning model
- Lab: Train a machine learning model in Azure Databricks

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
- Lab: Use MLflow in Azure Databricks

5) Tune hyperparameters in Azure Databricks

- Use the Hyperopt library to optimize hyperparameters
- Distribute hyperparameter tuning across multiple worker nodes
- Lab: Optimize hyperparameters for machine learning in Azure Databricks

6) Use AutoML in Azure Databricks

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

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
- Lab: Train deep learning models on Azure Databricks