
DP-3014 Implementing a Machine Learning solution with Azure Databricks

Overview

Course Duration: 1 Day

About This Course

Azure Databricks is a cloud-scale platform for data analytics and machine learning. Data scientists and machine learning engineers can use Azure Databricks to implement machine learning solutions at scale.

Course Outline

Module 1: Explore Azure Databricks.

- Get started with Azure Databricks.
- Identify Azure Databricks workloads.
- Understand key concepts.
- Exercise - Explore Azure Databricks.

Module 2: Use Apache Spark in Azure Databricks.

- Get to know Spark.
- Create a Spark cluster.
- Use Spark in notebooks.
- Use Spark to work with data files.
- Visualize data.
- Exercise - Use Spark in Azure Databricks.

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

- Understand principles of machine learning.
- Machine learning in Azure Databricks.
- Prepare data for machine learning.
- Train a machine learning model.
- Evaluate a machine learning model.
- Exercise - Train a machine learning model in Azure Databricks.

Module 4: Use MLflow in Azure Databricks.

- Capabilities of MLflow.
- Run experiments with MLflow.
- Register and serve models with MLflow.
- Exercise - Use MLflow in Azure Databricks.

Module 5: Tune hyperparameters in Azure Databricks.

- Optimize hyperparameters with Hyperopt.
- Review Hyperopt trials.
- Scale Hyperopt trials.
- Exercise - Optimize hyperparameters for machine learning in Azure Databricks.

Module 6: Use AutoML in Azure Databricks.

- What is AutoML.
- Use AutoML in the Azure Databricks user interface.
- Use code to run an AutoML experiment.
- Exercise - Use AutoML in Azure Databricks.

Module 7: Train deep learning models in Azure Databricks.

- Understand deep learning concepts.
- Train models with PyTorch.
- Distribute PyTorch training with Horovod.
- Exercise - Train deep learning models on Azure Databricks.

Prerequisites

Experience of using Python to explore data and train machine learning models with common open-source frameworks, like Scikit-Learn, PyTorch, and TensorFlow.