

---

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

Varighed: 1 Day    Kursus Kode: M-DP3014

---

### Beskrivelse:

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.

---

### Målgruppe:

Data scientists and machine learning engineers.

---

### Agenda:

- 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
- 

### Forudsætninger:

- 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.
-

## Indhold:

### 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

---

## Flere Informationer:

For yderligere informationer eller booking af kursus, kontakt os på tlf.nr.: 44 88 18 00

[training@globalknowledge.dk](mailto:training@globalknowledge.dk)

[www.globalknowledge.com/da-dk/](http://www.globalknowledge.com/da-dk/)

Global Knowledge, Stamholmen 110, 2650 Hvidovre