



# Startel

*Dé IT-opleider*

## Implementing a Machine Learning solution with Azure Databricks (DP-3014) Zelfstudiepakket

In dit zelfstudiepakket zit het officiële cursusmateriaal, een labomgeving en een achievement code (badge). Je hebt 180 dagen toegang tot de labomgeving.

### Algemene omschrijving

In dit zelfstudiepakket zit het officiële cursusmateriaal, een labomgeving en een achievement code (badge). Je hebt 180 dagen toegang tot de labomgeving.

Duik in de dynamische wereld van machine learning en ontdek hoe jij met het zelfstudiepakket Microsoft: Implementing a Machine Learning Solution with Azure Databricks (DP-3014) waardevolle data om kunt zetten in actiegerichte inzichten. Dit zelfstudiepakket is speciaal ontworpen voor datawetenschappers, data engineers en iedereen die geïnteresseerd is in het verkennen van de mogelijkheden van Azure Databricks met betrekking tot machine learning.

Leer de basisbeginselen van machine learning en hoe jij deze krachtige technologie toe kunt passen binnen Azure Databricks. Van het voorbereiden van datasets tot het trainen en beoordelen van modellen, dit zelfstudiepakket biedt een grondige basis voor iedereen die zijn/haar kennis en vaardigheden in data-analyse en machine learning uit wil breiden.

Verbeter jouw kennis en vaardigheden in het efficiënt beheren van machine learning-projecten door gebruik te maken van Azure Databricks. Leer hoe jij werkprocessen kunt automatiseren, de prestaties kunt verbeteren en hoe jij een goede samenwerking binnen jouw team kunt faciliteren, allemaal binnen één geïntegreerde omgeving.

### Doelgroep

Dit zelfstudiepakket is voornamelijk geschikt voor data engineers en datawetenschappers.

### Leerdoelen

Door dit zelfstudiepakket door te nemen zul je de volgende kennis en vaardigheden verkrijgen:

- Inzicht in de mogelijkheden van Azure Databricks voor machine learning.

- Het opzetten van Azure Databricks.
- Het verwerken en verkennen van data.
- Machine learning met Azure Databricks.
- Machine learning-modellen creëren en afstemmen.
- Machine learning-modellen inzetten.
- Inzicht in gezamenlijke machine learning-ontwikkeling.
- Inzicht in de beste werkwijzen en het optimaliseren van de prestaties.

## Voorkennis

Voordat je dit zelfstudiepakket doorneemt is het van belang dat jij in staat bent om de programmeertaal Python te gebruiken om data te onderzoeken en machine learning-modellen te trainen met algemene Open Source-frameworks, zoals Scikit-Learn, PyTorch en TensorFlow.

## Onderwerpen

### Module 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

### 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.

#### Lab: Use Spark in Azure Databricks

### 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.

#### Lab: Train a machine learning model in Azure Databricks

### 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.

#### Lab: Use MLflow in Azure Databricks

### Module 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**

### **Module 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**

### **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.

### **Lab: Train deep learning models on Azure Databricks**