

DP-100T01-A: Designing and Implementing a Data Science Solution on Azure

OBJECTIVE

Learn how to operate machine learning solutions at cloud scale using Azure Machine Learning. This course teaches you to leverage your existing knowledge of Python and machine learning to manage data ingestion and preparation, model training and deployment, and machine learning solution monitoring in Microsoft Azure.

COURSE TOPICS

Module 1: Introduction to Azure Machine Learning

In this module, you will learn how to provision an Azure Machine Learning workspace and use it to manage machine learning assets such as data, compute, model training code, logged metrics, and trained models. You will learn how to use the web-based Azure Machine Learning studio interface as well as the Azure Machine Learning SDK and developer tools like Visual Studio Code and Jupyter Notebooks to work with the assets in your workspace.

Lessons

- Getting Started with Azure Machine Learning
- Azure Machine Learning Tools

After completing this module, you will be able to

- Provision an Azure Machine Learning workspace
- Use tools and code to work with Azure Machine Learning

Module 2: No-Code Machine Learning with Designer

This module introduces the Designer tool, a drag and drop interface for creating machine learning models without writing any code. You will learn how to create a training pipeline that encapsulates data preparation and model training, and then convert that training pipeline to an inference pipeline that can be used to predict values from new data, before finally deploying the inference pipeline as a service for client applications to consume.

Lessons

- Training Models with Designer
- Publishing Models with Designer

After completing this module, you will be able to

- Use designer to train a machine learning model
- Deploy a Designer pipeline as a service

Module 3: Running Experiments and Training Models

In this module, you will get started with experiments that encapsulate data processing and model training code, and use them to train machine learning models.

Lessons

- Introduction to Experiments
- Training and Registering Models

After completing this module, you will be able to

- Run code-based experiments in an Azure Machine Learning workspace
- Train and register machine learning models

Module 4: Working with Data

Data is a fundamental element in any machine learning workload, so in this module, you will learn how to create and manage datastores and datasets in an Azure Machine Learning workspace, and how to use them in model training experiments.

Lessons

- Working with Datastores
- Working with Datasets

After completing this module, you will be able to

- Create and consume datastores
- Create and consume datasets

Module 5: Compute Contexts

One of the key benefits of the cloud is the ability to leverage compute resources on demand, and use them to scale machine learning processes to an extent that would be infeasible on your own hardware. In this module, you'll learn how to manage experiment environments that ensure consistent runtime consistency for experiments, and how to create and use compute targets for experiment runs.

Lessons

- Working with Environments
- Working with Compute Targets

After completing this module, you will be able to

- Create and use environments
- Create and use compute targets

Module 6: Orchestrating Operations with Pipelines

Now that you understand the basics of running workloads as experiments that leverage data assets and compute resources, it's time to learn how to orchestrate these workloads as pipelines of connected steps. Pipelines are key to implementing an effective Machine Learning Operationalization (ML Ops) solution in Azure, so you'll explore how to define and run them in this module.

Lessons

- Introduction to Pipelines
- Publishing and Running Pipelines

After completing this module, you will be able to

- Create pipelines to automate machine learning workflows
- Publish and run pipeline services

Module 7: Deploying and Consuming Models

Models are designed to help decision making through predictions, so they're only useful when deployed and available for an application to consume. In this module learn how to deploy models for real-time inferencing, and for batch inferencing.

Lessons

- Real-time Inferencing
- Batch Inferencing

After completing this module, you will be able to

- Publish a model as a real-time inference service
- Publish a model as a batch inference service

Module 8: Training Optimal Models

By this stage of the course, you've learned the end-to-end process for training, deploying, and consuming machine learning models; but how do you ensure your model produces the best predictive outputs for your data? In this module, you'll explore how you can use hyperparameter tuning and automated machine learning to take advantage of cloud-scale compute and find the best model for your data.

Lessons

- Hyperparameter Tuning
- Automated Machine Learning

After completing this module, you will be able to

- Optimize hyperparameters for model training
- Use automated machine learning to find the optimal model for your data

Module 9: Interpreting Models

Many of the decisions made by organizations and automated systems today are based on predictions made by machine learning models. It's increasingly important to be able to understand the factors that influence the predictions made by a model, and to be able to determine any unintended biases in the model's behavior. This module describes how you can interpret models to explain how feature importance determines their predictions.

Lessons

- Introduction to Model Interpretation
- using Model Explainers

After completing this module, you will be able to

- Generate model explanations with automated machine learning
- Use explainers to interpret machine learning models

Module 10: Monitoring Models

After a model has been deployed, it's important to understand how the model is being used in production, and to detect any degradation in its effectiveness due to data drift. This module describes techniques for monitoring models and their data.

Lessons

- Monitoring Models with Application Insights
- Monitoring Data Drift

After completing this module, you will be able to

- Use Application Insights to monitor a published model
- Monitor data drift

PREREQUISITES

Before attending this course, students must have:

- A fundamental knowledge of Microsoft Azure
- Experience of writing Python code to work with data, using libraries such as Numpy, Pandas, and Matplotlib.
- Understanding of data science; including how to prepare data, and train machine learning models using common machine learning libraries such as Scikit-Learn, PyTorch, or Tensorflow.

TRAINING APPROACH

This course includes lectures, course notes, exercises and hands-on practice.

COURSE DURATION

16 Hours (in 2 days)

Time: 9:00am to 6:00pm

Lunch Time: 1:00pm to 2:00pm

CERTIFICATION COMPLETION

A certificate of completion is provided for all trainees attending the course.