Kód kurzu: MOC DP-100

Kurz je určený pre dátových analytikov (data scientist), ktorí sa chcú zoznámiť s procesom analýzy dát (data science) a pochopiť, ako im vybrané služby v Microsoft Azure môžu pomôcť tento proces implementovať, zefektívniť a obohatiť. Zoznámite sa s použitím služieb Azure Notebooks a Azure Machine Learning a naučíte v Azure Machine Learning vytvárať vlastné modely strojového učenia. Kurz je vhodný ako príprava k získaniu titulu Microsoft Certified: Azure Data Scientist Associate.

Pre koho je kurz určený

Azure Data Scientist

Čo vás naučíme

- Set up an Azure Machine Learning workspace (30-35%)
- Run experiments and train models (25-30%)
- Optimize and manage models (20-25%)
- Deploy and consume models (20-25%)

Osnova kurzu

Module 1: Getting Started with 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.

Module 2: No-Code Machine Learning

This module introduces the Automated Machine Learning and Designer visual tools, which you can use to train, evaluate, and deploy machine learning models without writing any code.

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.

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.

Module 5: Working with Compute

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.

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.

Module 7: Deploying and Consuming Models

GOPAS

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.

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.

Module 9: Responsible Machine Learning

Data scientists have a duty to ensure they analyze data and train machine learning models responsibly; respecting individual privacy, mitigating bias, and ensuring transparency. This module explores some considerations and techniques for applying responsible machine learning principles.

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.