

DP-100^{Q&As}

Designing and Implementing a Data Science Solution on Azure

Pass Microsoft DP-100 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

<https://www.leads4pass.com/dp-100.html>

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Microsoft
Official Exam Center

- ⚙️ **Instant Download** After Purchase
- ⚙️ **100% Money Back** Guarantee
- ⚙️ **365 Days** Free Update
- ⚙️ **800,000+** Satisfied Customers



QUESTION 1

You plan to build a team data science environment. Data for training models in machine learning pipelines will be over 20 GB in size. You have the following requirements:

1.

Models must be built using Caffe2 or Chainer frameworks.

2.

Data scientists must be able to use a data science environment to build the machine learning pipelines and train models on their personal devices in both connected and disconnected network environments.

Personal devices must support updating machine learning pipelines when connected to a network.

You need to select a data science environment.

Which environment should you use?

- A. Azure Machine Learning Service
- B. Azure Machine Learning Studio
- C. Azure Databricks
- D. Azure Kubernetes Service (AKS)

Correct Answer: A

The Data Science Virtual Machine (DSVM) is a customized VM image on Microsoft's Azure cloud built specifically for doing data science. Caffe2 and Chainer are supported by DSVM. DSVM integrates with Azure Machine Learning.

Incorrect Answers:

B: Use Machine Learning Studio when you want to experiment with machine learning models quickly and easily, and the built-in machine learning algorithms are sufficient for your solutions.

References: <https://docs.microsoft.com/en-us/azure/machine-learning/data-science-virtual-machine/overview>

QUESTION 2

You create an Azure Machine Learning workspace.

You must use the Python SDK v2 to implement an experiment from a Jupyter notebook in the workspace. The experiment must log string metrics.

You need to implement the method to log the string metrics.

Which method should you use?

- A. `mlflow.log_artifact()`

B. `mlflow.log.dict()`

C. `mlflow.log_metric()`

D. `mlflow.log_text()`

Correct Answer: D

Logging files

In general, files in MLflow are called artifacts. You can log artifacts in multiple ways in MLflow:

*

Log text in a text file

```
mlflow.log_text("text string", "notes.txt")
```

Text is persisted inside of the run in a text file with name notes.txt.

Incorrect:

*

`mlflow.log.dict`

Log dictionaries as JSON and YAML files

```
mlflow.log_dict(dictionary, "file.yaml")
```

dictionary is a dictionary object containing all the structure that you want to persist as JSON or YAML file.

*

`mlflow.log_artifact`

Log a trivial file already existing

```
mlflow.log_artifact("path/to/file.pkl")
```

Files are always logged in the root of the run. If `artifact_path` is provided, then the file is logged in a folder as indicated in that parameter.

*

Log all the artifacts in an existing folder

```
mlflow.log_artifacts("path/to/folder")
```

Folder structure is copied to the run, but the root folder indicated is not included.

Reference:

<https://learn.microsoft.com/en-us/azure/machine-learning/how-to-log-view-metrics>

QUESTION 3

You are retrieving data from a large datastore by using Azure Machine Learning Studio.

You must create a subset of the data for testing purposes using a random sampling seed based on the system clock.

You add the Partition and Sample module to your experiment.

You need to select the properties for the module.

Which values should you select? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area

▲ Partition and Sample

Partition or sample mode

	▼
Assign to Folds	
Pick Fold	
Sampling	
Head	

Rate of sampling

.2

Random seed for sampling

	▼
0	
1	
time.clock()	
utcNow()	

Stratified split for sampling

False	▼
-------	---

Correct Answer:

Answer Area

Partition and Sample

Partition or sample mode

	▼
Assign to Folds	
Pick Fold	
Sampling	
Head	

Rate of sampling

.2

Random seed for sampling

	▼
0	
1	
time.clock()	
utcNow()	

Stratified split for sampling

False	▼
-------	---

Box 1: Sampling Create a sample of data This option supports simple random sampling or stratified random sampling. This is useful if you want to create a smaller representative sample dataset for testing.

1. Add the Partition and Sample module to your experiment in Studio, and connect the dataset.
2. Partition or sample mode: Set this to Sampling.
3. Rate of sampling. See box 2 below.

Box 2: 0

3. Rate of sampling. Random seed for sampling: Optionally, type an integer to use as a seed value.

This option is important if you want the rows to be divided the same way every time. The default value is 0, meaning that a starting seed is generated based on the system clock. This can lead to slightly different results each time you run the experiment.

References: <https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/partition-and-sample>

QUESTION 4

You use the Azure Machine Learning SDK for Python v1 and notebooks to train a model. You create a compute target, an environment, and a training script by using Python code.

You need to prepare information to submit a training run.

Which class should you use?

- A. ScriptRun
- B. ScriptRunConfig
- C. RunConfiguration
- D. Run

Correct Answer: B

A ScriptRunConfig is used to configure the information necessary for submitting a training job as part of an experiment.

Reference: <https://learn.microsoft.com/en-us/azure/machine-learning/v1/how-to-set-up-training-targets>

QUESTION 5

You run a script as an experiment in Azure Machine Learning.

You have a Run object named run that references the experiment run. You must review the log files that were generated during the experiment run.

You need to download the log files to a local folder for review.

Which two code segments can you run to achieve this goal? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

- A. `run.get_details()`
- B. `run.get_file_names()`
- C. `run.get_metrics()`
- D. `run.download_files(output_directory='\\./runfiles\\')`
- E. `run.get_all_logs(destination='\\./runlogs\\')`

Correct Answer: AE

The run Class `get_all_logs` method downloads all logs for the run to a directory.

The run Class `get_details` gets the definition, status information, current log files, and other details of the run.

Incorrect Answers:

B: The run `get_file_names` list the files that are stored in association with the run.

Reference: [https://docs.microsoft.com/en-us/python/api/azureml-core/azureml.core.run\(class\)](https://docs.microsoft.com/en-us/python/api/azureml-core/azureml.core.run(class))

[DP-100 PDF Dumps](#)

[DP-100 VCE Dumps](#)

[DP-100 Exam Questions](#)