

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 are solving a classification task.

You must evaluate your model on a limited data sample by using k-fold cross-validation. You start by configuring a k parameter as the number of splits.

You need to configure the k parameter for the cross-validation.

Which value should you use?

- A. k=1
- B. k=10
- C. k=0.5
- D. k=0.9

Correct Answer: B

Leave One Out (LOO) cross-validation

Setting $K = n$ (the number of observations) yields n-fold and is called leave-one out cross-validation (LOO), a special case of the K-fold approach.

LOO CV is sometimes useful but typically doesn't shake up the data enough. The estimates from each fold are highly correlated and hence their average can have high variance. This is why the usual choice is $K=5$ or 10 . It provides a good

compromise for the bias-variance tradeoff.

QUESTION 2

DRAG DROP

You create an Azure Machine Learning workspace and a new Azure DevOps organization. You register a model in the workspace and deploy the model to the target environment.

All new versions of the model registered in the workspace must automatically be deployed to the target environment.

You need to configure Azure Pipelines to deploy the model.

Which four actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

Select and Place:

Actions

Create a service connection

Create a release pipeline

Create a build pipeline

Create an Azure DevOps project

Install the Machine Learning extension for Azure Pipelines

Answer

Correct Answer:

Actions

Create a build pipeline

Answer

Create an Azure DevOps project
Create a release pipeline
Install the Machine Learning extension for Azure Pipelines
Create a service connection

Step 1: Create an Azure DevOps project

Step 2: Create a release pipeline

1.

Sign in to your Azure DevOps organization and navigate to your project.

2.

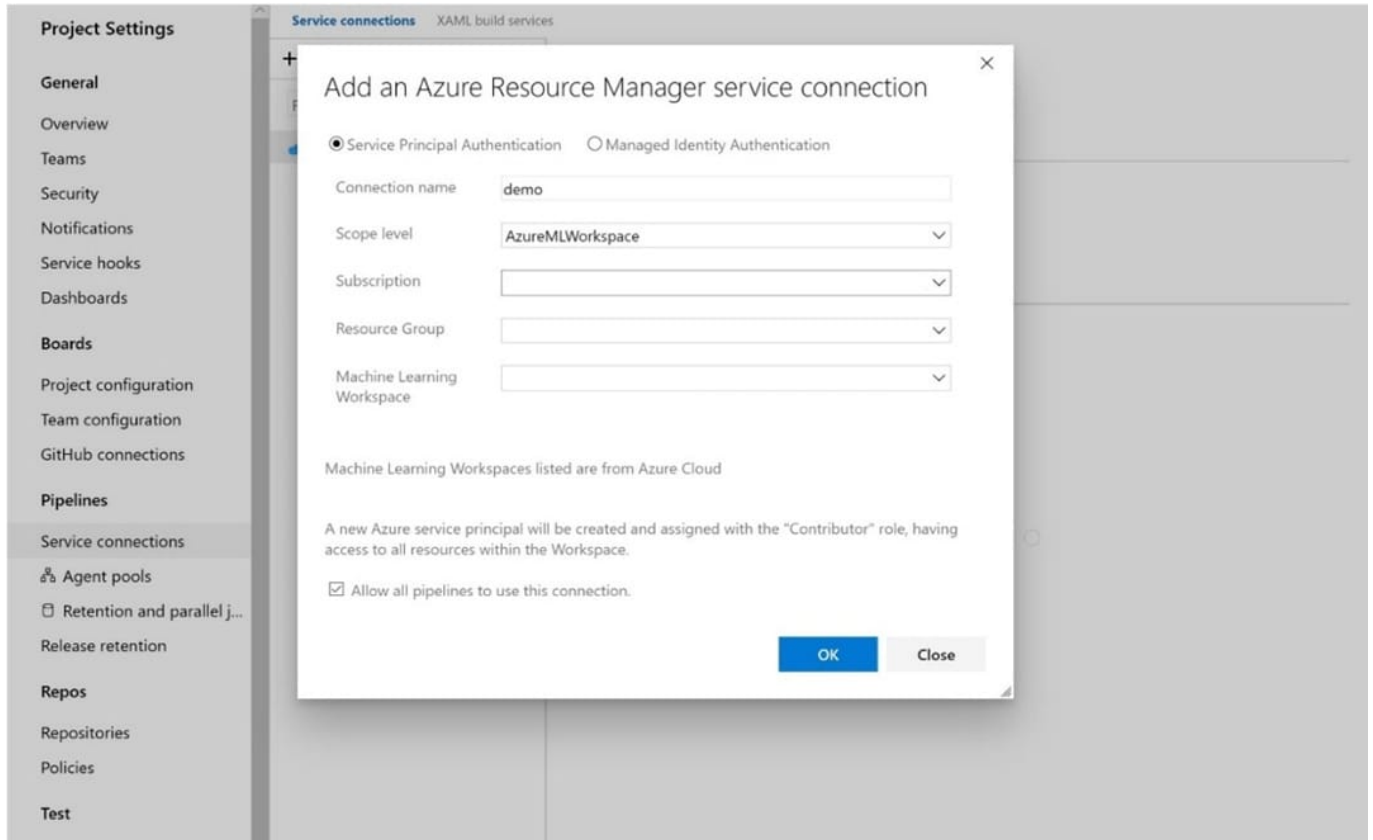
Go to Pipelines, and then select New pipeline.

Step 3: Install the Machine Learning extension for Azure Pipelines You must install and configure the Azure CLI and ML extension.

Step 4: Create a service connection How to set up your service connection Select AzureMLWorkspace for the scope level, then fill in the following subsequent parameters.

The screenshot displays the Azure DevOps 'Project Settings' page. On the left, the 'Pipelines' section is expanded, and 'Service connections' is selected. The 'Service connections' dropdown menu is open, showing a list of providers. 'Azure Resource Manager' is highlighted with a red rectangle. The right pane shows the 'Service connection: demo' details, including tabs for 'Details', 'Roles', 'Request history', and 'Policies'. The 'INFORMATION' section is currently visible, and the 'ACTIONS' section lists several management options.

Project Settings	Service connections	Service connection: demo
General	+ New service connection	Details Roles Request history Policies
Overview	Azure Classic	INFORMATION
Teams	Azure Repos/Team Foundation Se...	
Security	Azure Resource Manager	
Notifications	Azure Service Bus	
Service hooks	Bitbucket Cloud	
Dashboards	Chef	ACTIONS
Boards	DLIS (Alpha)	List of actions that can be performed on this service connection:
Project configuration	Docker Host	Update service connection
Team configuration	Docker Registry	Manage service connection roles
GitHub connections	Generic	Manage Service Principal
Pipelines		Disconnect
Service connections		
Agent pools		
Retention and parallel jobs		
Release retention		
Repos		
Repositories		
Policies		
Test		



Note: How to enable model triggering in a release pipeline

1.

Go to your release pipeline and add a new artifact. Click on AzureML Model artifact then select the appropriate AzureML service connection and select from the available models in your workspace.

2.

Enable the deployment trigger on your model artifact as shown here. Every time a new version of that model is registered, a release pipeline will be triggered.

Reference: <https://marketplace.visualstudio.com/items?itemName=ms-air-aiagility.vss-services-azureml>
<https://docs.microsoft.com/en-us/azure/devops/pipelines/targets/azure-machine-learning>

QUESTION 3

HOTSPOT

You are performing sentiment analysis using a CSV file that includes 12,000 customer reviews written in a short sentence format.

You add the CSV file to Azure Machine Learning Studio and Configure it as the starting point dataset of an experiment.

You add the Extract N-Gram Features from Text module to the experiment to extract key phrases from the customer review column in the dataset.

You must create a new n-gram text dictionary from the customer review text and set the maximum n-gram size to trigrams.

You need to configure the Extract N Gram features from Text module.

What should you select? To answer, select the appropriate options in the answer area;

NOTE: Each correct selection is worth one point.

Hot Area:

vocabulary mode

	▼
Create	
ReadOnly	
Update	
Merge	

N-Grams size

	▼
3	
4	
4,000	
12,000	

Correct Answer:

vocabulary mode

	▼
Create	
ReadOnly	
Update	
Merge	

N-Grams size

	▼
3	
4	
4,000	
12,000	

QUESTION 4

You define a datastore named ml-data for an Azure Storage blob container. In the container, you have a folder named train that contains a file named data.csv. You plan to use the file to train a model by using the Azure Machine Learning SDK.

You plan to train the model by using the Azure Machine Learning SDK to run an experiment on local compute.

You define a DataReference object by running the following code:


```
from azureml.core import Workspace, Datastore, Environment
from azureml.train.estimator import Estimator
ws = Workspace.from_config()
ml_data = Datastore.get(ws, datastore_name='ml-data')
data_ref = ml_data.path('train').as_download(path_on_compute='train_data')
estimator = Estimator(source_directory='experiment_folder',
    script_params={'--data-folder': data_ref},
    compute_target = 'local',
    entry_script='training.py')
run = experiment.submit(config=estimator)
run.wait_for_completion(show_output=True)
```

You need to load the training data. Which code segment should you use?

- A.

```
import os
import argparse
import pandas as pd

parser = argparse.ArgumentParser()
parser.add_argument('--data-folder', type=str, dest='data_folder')
data_folder = args.data_folder
data = pd.read_csv(os.path.join(data_folder, 'ml-data', 'train_data', 'data.csv'))
```
- B.

```
import os
import argparse
import pandas as pd

parser = argparse.ArgumentParser()
parser.add_argument('--data-folder', type=str, dest='data_folder')
data_folder = args.data_folder
data = pd.read_csv(os.path.join(data_folder, 'train', 'data.csv'))
```
- C.

```
import pandas as pd

data = pd.read_csv('./data.csv')
```
- D.

```
import os
import argparse
import pandas as pd

parser = argparse.ArgumentParser()
parser.add_argument('--data-folder', type=str, dest='data_folder')
data_folder = args.data_folder
data = pd.read_csv(os.path.join('ml_data', data_folder, 'data.csv'))
```
- E.

```
import os
import argparse
import pandas as pd

parser = argparse.ArgumentParser()
parser.add_argument('--data-folder', type=str, dest='data_folder')
data_folder = args.data_folder
data = pd.read_csv(os.path.join(data_folder, 'data.csv'))
```

A. Option A

B. Option B

C. Option C

D. Option D

E. Option E

Correct Answer: E

Example:

```
data_folder = args.data_folder
```

```
# Load Train and Test data
```

```
train_data = pd.read_csv(os.path.join(data_folder, '\\data.csv\\'))
```

Reference:

<https://www.element61.be/en/resource/azure-machine-learning-services-complete-toolbox-ai>

QUESTION 5

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while

others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You are analyzing a numerical dataset which contains missing values in several columns.

You must clean the missing values using an appropriate operation without affecting the dimensionality of the feature set.

You need to analyze a full dataset to include all values.

Solution: Replace each missing value using the Multiple Imputation by Chained Equations (MICE) method.

Does the solution meet the goal?

A. Yes

B. No

Correct Answer: A

Replace using MICE: For each missing value, this option assigns a new value, which is calculated by using a method described in the statistical literature as "Multivariate Imputation using Chained Equations" or "Multiple Imputation by Chained Equations". With a multiple imputation method, each variable with missing data is modeled conditionally using the other variables in the data before filling in the missing values.

Note: Multivariate imputation by chained equations (MICE), sometimes called "fully conditional specification" or "sequential regression multiple imputation" has emerged in the statistical literature as one principled method of addressing missing data. Creating multiple imputations, as opposed to single imputations, accounts for the statistical uncertainty in the imputations. In addition, the chained equations approach is very flexible and can handle variables of

varying types (e.g., continuous or binary) as well as complexities such as bounds or survey skip patterns.

References: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3074241/>

<https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/clean-missing-data>

[Latest DP-100 Dumps](#)

[DP-100 Exam Questions](#)

[DP-100 Braindumps](#)