
# Microsoft

## Exam Questions DP-100

Designing and Implementing a Data Science Solution on Azure

**NEW QUESTION 1**
- (Exam Topic 3)
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 create a model to forecast weather conditions based on historical data.
You need to create a pipeline that runs a processing script to load data from a datastore and pass the processed data to a machine learning model training script.
Solution: Run the following code:

```
datastore = ws.get_default_datastore()
data_output = PipelineData("processed_data", datastore=datastore)
process_step = PythonScriptStep(script_name="process.py",
 arguments=["--data_for_train", data_output],
 outputs=[data_output], compute_target=aml_compute,
 source_directory=process_directory)
pipeline = Pipeline(workspace=ws, steps=[process_step])
```

Does the solution meet the goal?

A. Yes
B. No

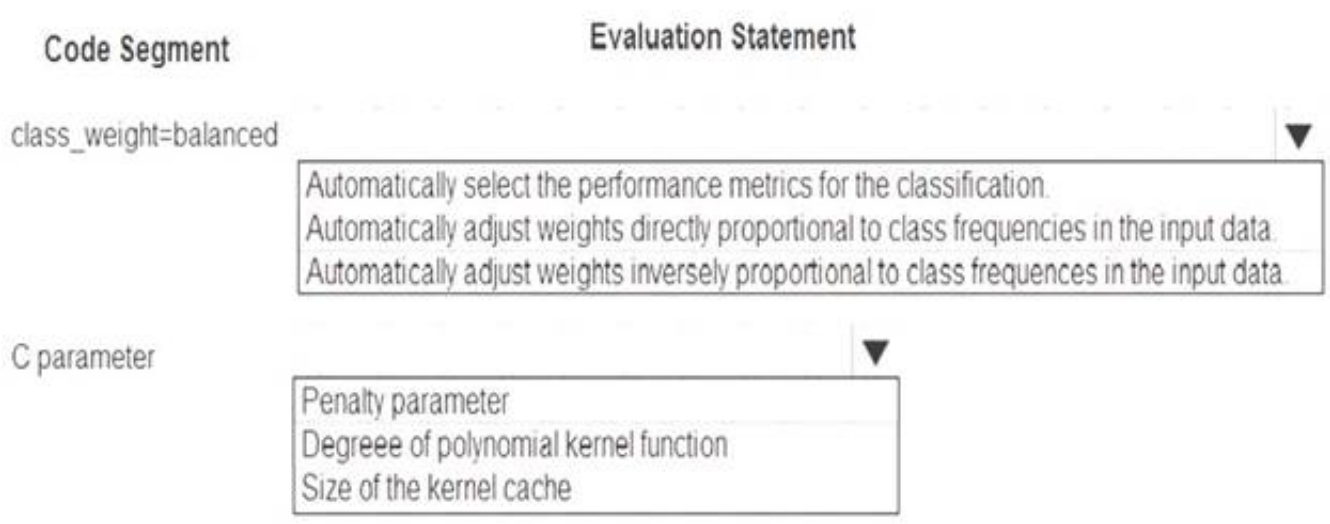
**Answer:** B

**Explanation:**
train_step is missing. Reference:
https://docs.microsoft.com/en-us/python/api/azureml-pipeline-core/azureml.pipeline.core.pipelinedata?view=azu

**NEW QUESTION 2**
- (Exam Topic 3)
You are using C-Support Vector classification to do a multi-class classification with an unbalanced training dataset. The C-Support Vector classification using Python code shown below:

```
from sklearn.svm import svc
import numpy as np
svc = SVC(kernel= 'linear', class_weight= 'balanced', C-1.0, random_state-0)
model1 = svc.fit(X_train, y)
```

You need to evaluate the C-Support Vector classification code.
Which evaluation statement should you use? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.




A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Box 1: Automatically adjust weights inversely proportional to class frequencies in the input data
The “balanced” mode uses the values of y to automatically adjust weights inversely proportional to class frequencies in the input data as n_samples / (n_classes * np.bincount(y)).
Box 2: Penalty parameter
Parameter: C : float, optional (default=1.0) Penalty parameter C of the error term. References:
https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html

**NEW QUESTION 3**
- (Exam Topic 3)
HOTSPOT

You register the following versions of a model.

| Model name | Model version | Tags | Properties |
|---|---|---|---|
| healthcare_model | 3 | 'Training context':'CPU Compute' | value:87.43 |
| healthcare_model | 2 | 'Training context':'CPU Compute' | value:54.98 |
| healthcare_model | 1 | 'Training context':'CPU Compute' | value:23.56 |

You use the Azure ML Python SDK to run a training experiment. You use a variable named run to reference the experiment run. After the run has been submitted and completed, you run the following code:

```
run.register_model(model_path='outputs/model.pkl',
 model_name='healthcare_model',
 tags={'Training context':'CPU Compute'} )
```

For each of the following statements, select Yes if the statement is true. Otherwise, select No.
NOTE: Each correct selection is worth one point.

| | Yes | No |
|---|---|---|
| The code will cause a previous version of the saved model to be overwritten. | ○ | ○ |
| The version number will now be 4. | ○ | ○ |
| The latest version of the stored model will have a property of value: 87.43. | ○ | ○ |

A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-deploy-and-where

**NEW QUESTION 4**
- (Exam Topic 3)
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.
An IT department creates the following Azure resource groups and resources:

| Resource group | Resources |
|---|---|
| ml_resources | • an Azure Machine Learning workspace named amlworkspace<br>• an Azure Storage account named amlworkspace12345<br>• an Application Insights instance named amlworkspace54321<br>• an Azure Key Vault named amlworkspace67890<br>• an Azure Container Registry named amlworkspace09876 |
| general_compute | A virtual machine named mlvm with the following configuration:<br>• Operating system: Ubuntu Linux<br>• Software installed: Python 3.6 and Jupyter Notebooks<br>• Size: NC6 (6 vCPUs, 1 vGPU, 56 Gb RAM) |

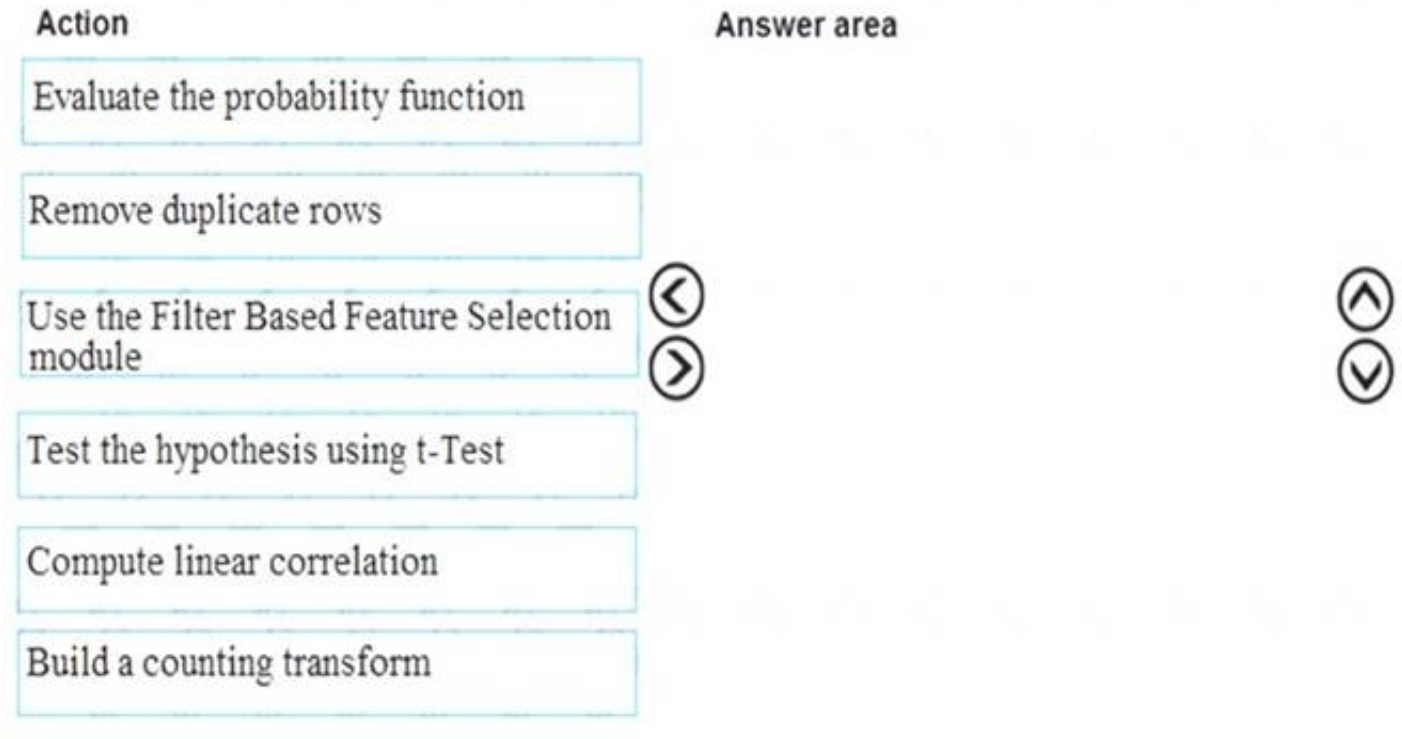
The IT department creates an Azure Kubernetes Service (AKS)-based inference compute target named
aks-cluster in the Azure Machine Learning workspace. You have a Microsoft Surface Book computer with a GPU. Python 3.6 and Visual Studio Code are installed.
You need to run a script that trains a deep neural network (DNN) model and logs the loss and accuracy metrics.
Solution: Install the Azure ML SDK on the Surface Book. Run Python code to connect to the workspace. Run the training script as an experiment on the aks-cluster compute target.
Does the solution meet the goal?

A. Yes
B. No

**Answer:** B

**Explanation:**
Need to attach the mlvm virtual machine as a compute target in the Azure Machine Learning workspace.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/concept-compute-target

**NEW QUESTION 5**
- (Exam Topic 3)
You are producing a multiple linear regression model in Azure Machine Learning Studio. Several independent variables are highly correlated.
You need to select appropriate methods for conducting effective feature engineering on all the data.
Which three 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.




A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Step 1: Use the Filter Based Feature Selection module
Filter Based Feature Selection identifies the features in a dataset with the greatest predictive power.
The module outputs a dataset that contains the best feature columns, as ranked by predictive power. It also outputs the names of the features and their scores from the selected metric.
Step 2: Build a counting transform
A counting transform creates a transformation that turns count tables into features, so that you can apply the transformation to multiple datasets.
Step 3: Test the hypothesis using t-Test References:
https://docs.microsoft.com/bs-latn-ba/azure/machine-learning/studio-module-reference/filter-based-feature-selec
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/build-counting-transform

**NEW QUESTION 6**
- (Exam Topic 3)
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 creating a new experiment in Azure Machine Learning Studio.
One class has a much smaller number of observations than the other classes in the training set. You need to select an appropriate data sampling strategy to compensate for the class imbalance. Solution: You use the Stratified split for the sampling mode.
Does the solution meet the goal?

A. Yes
B. No

**Answer:** B

**Explanation:**
Instead use the Synthetic Minority Oversampling Technique (SMOTE) sampling mode.
Note: SMOTE is used to increase the number of underepresented cases in a dataset used for machine learning. SMOTE is a better way of increasing the number of rare cases than simply duplicating existing cases.
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/smote

**NEW QUESTION 7**
- (Exam Topic 3)
You plan to use automated machine learning to train a regression model. You have data that has features which have missing values, and categorical features with few distinct values.
You need to configure automated machine learning to automatically impute missing values and encode categorical features as part of the training task.
Which parameter and value pair should you use in the AutoMLConfig class?

A. featurization = 'auto'
B. enable_voting_ensemble = True
C. task = 'classification'
D. exclude_nan_labels = True
E. enable_tf = True

**Answer:** A

**Explanation:**
Featurization str or FeaturizationConfig Values: 'auto' / 'off' / FeaturizationConfig
Indicator for whether featurization step should be done automatically or not, or whether customized featurization should be used.
Column type is automatically detected. Based on the detected column type preprocessing/featurization is done as follows:
Categorical: Target encoding, one hot encoding, drop high cardinality categories, impute missing values. Numeric: Impute missing values, cluster distance, weight of evidence.
DateTime: Several features such as day, seconds, minutes, hours etc. Text: Bag of words, pre-trained Word embedding, text target encoding. Reference:
https://docs.microsoft.com/en-us/python/api/azureml-train-automl-client/azureml.train.automl.automlconfig.auto

**NEW QUESTION 8**
- (Exam Topic 3)
You are preparing to build a deep learning convolutional neural network model for image classification. You create a script to train the model using CUDA devices.
You must submit an experiment that runs this script in the Azure Machine Learning workspace. The following compute resources are available:

- a Microsoft Surface device on which Microsoft Office has been installed. Corporate IT policies prevent the installation of additional software
- a Compute Instance named ds-workstation in the workspace with 2 CPUs and 8 GB of memory
- an Azure Machine Learning compute target named cpu-cluster with eight CPU-based nodes
- an Azure Machine Learning compute target named gpu-cluster with four CPU and GPU-based nodes

You need to specify the compute resources to be used for running the code to submit the experiment, and for running the script in order to minimize model training time.
Which resources should the data scientist use? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.

| Resource type | Option |
|---|---|
| Run code to submit the experiment | the Microsoft Surface device<br>the ds-workstation notebook VM<br>the cpu-cluster compute target<br>the gpu-cluster compute target |
| Run the training script | the ds-workstation notebook VM<br>the cpu-compute target<br>the gpu-compute target<br>the Microsoft Surface device |

A. Mastered
B. Not Mastered

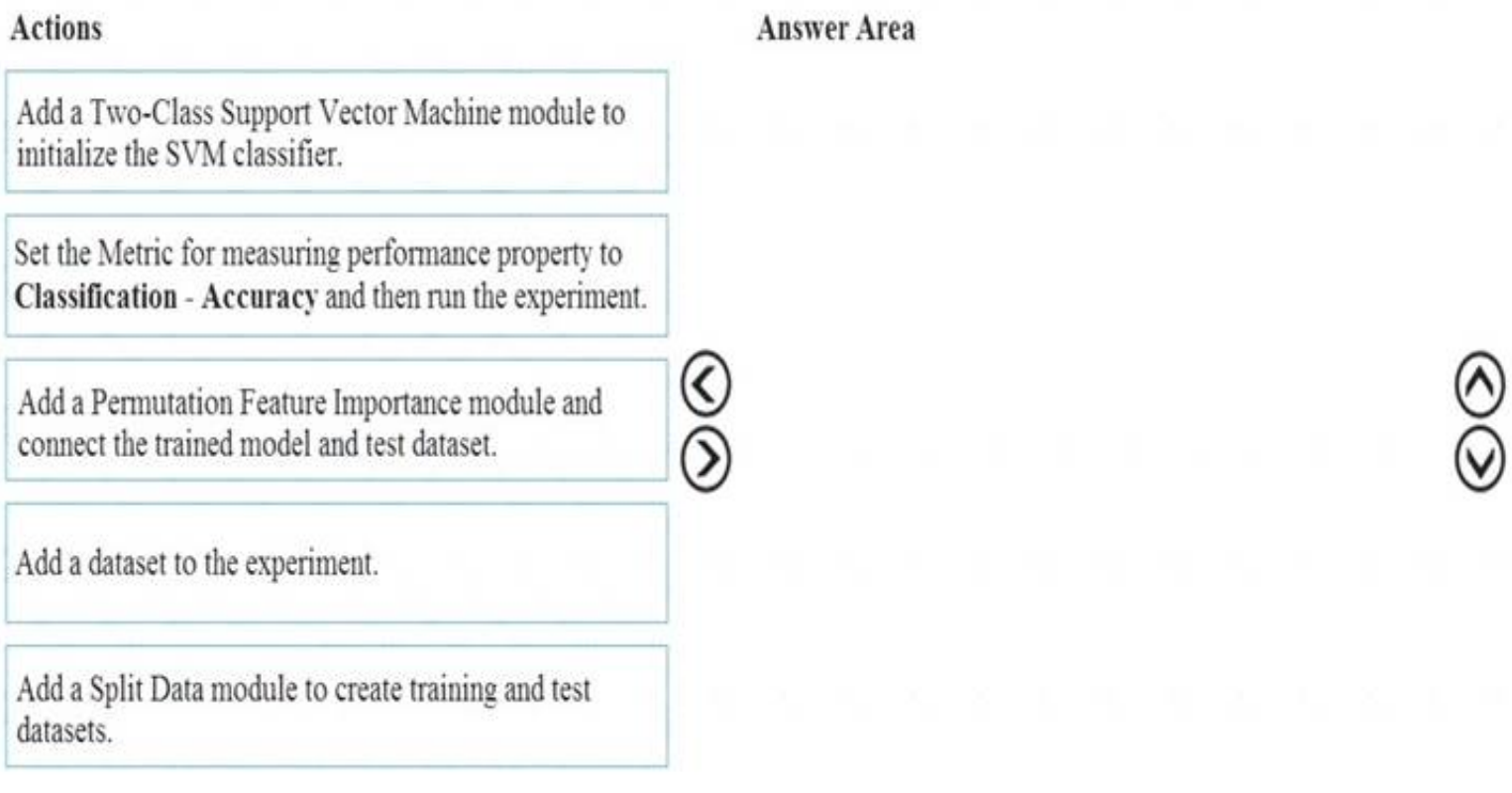
**Answer:** A

**Explanation:**

| Resource type | Option |
|---|---|
| Run code to submit the experiment | the Microsoft Surface device<br>the ds-workstation notebook VM<br>the cpu-cluster compute target<br>the gpu-cluster compute target |
| Run the training script | the ds-workstation notebook VM<br>the cpu-compute target<br>the gpu-compute target<br>the Microsoft Surface device |

**NEW QUESTION 9**

- (Exam Topic 3)
You have a dataset that contains over 150 features. You use the dataset to train a Support Vector Machine (SVM) binary classifier.
You need to use the Permutation Feature Importance module in Azure Machine Learning Studio to compute a set of feature importance scores for the dataset.
In which order should you perform the actions? To answer, move all actions from the list of actions to the answer area and arrange them in the correct order.




A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Step 1: Add a Two-Class Support Vector Machine module to initialize the SVM classifier.
Step 2: Add a dataset to the experiment
Step 3: Add a Split Data module to create training and test dataset.
To generate a set of feature scores requires that you have an already trained model, as well as a test dataset. Step 4: Add a Permutation Feature Importance module and connect to the trained model and test dataset. Step 5: Set the Metric for measuring performance property to Classification - Accuracy and then run the experiment.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/two-class-support-vector-mac https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/permutation-feature-importan

**NEW QUESTION 10**
- (Exam Topic 3)
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 plan to use a Python script to run an Azure Machine Learning experiment. The script creates a reference to the experiment run context, loads data from a file, identifies the set of unique values for the label column, and completes the experiment run:

```
from azureml.core import Run
import pandas as pd

run = Run.get_context()
data = pd.read_csv('data.csv')
label_vals = data['label'].unique()
# Add code to record metrics here
run.complete()
```

The experiment must record the unique labels in the data as metrics for the run that can be reviewed later. You must add code to the script to record the unique label values as run metrics at the point indicated by the comment.
Solution: Replace the comment with the following code:
run.log_list('Label Values', label_vals) Does the solution meet the goal?

A. Yes
B. No

**Answer:** A

**Explanation:**
run.log_list log a list of values to the run with the given name using log_list. Example: run.log_list("accuracies", [0.6, 0.7, 0.87])
Note:
Data= pd.read_csv('data.csv')
Data is read into a pandas.DataFrame, which is a two-dimensional, size-mutable, potentially heterogeneous tabular data.

label_vals =data['label'].unique
label_vals contains a list of unique label values. Reference:
https://www.element61.be/en/resource/azure-machine-learning-services-complete-toolbox-ai https://docs.microsoft.com/en-us/python/api/azureml-core/azureml.core.run(class) https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.html

**NEW QUESTION 10**
- (Exam Topic 3)
You use the Two-Class Neural Network module in Azure Machine Learning Studio to build a binary classification model. You use the Tune Model Hyperparameters module to tune accuracy for the model.
You need to select the hyperparameters that should be tuned using the Tune Model Hyperparameters module. Which two hyperparameters should you use? Each correct answer presents part of the solution.
NOTE: Each correct selection is worth one point.

A. Number of hidden nodes
B. Learning Rate
C. The type of the normalizer
D. Number of learning iterations
E. Hidden layer specification

**Answer:** DE

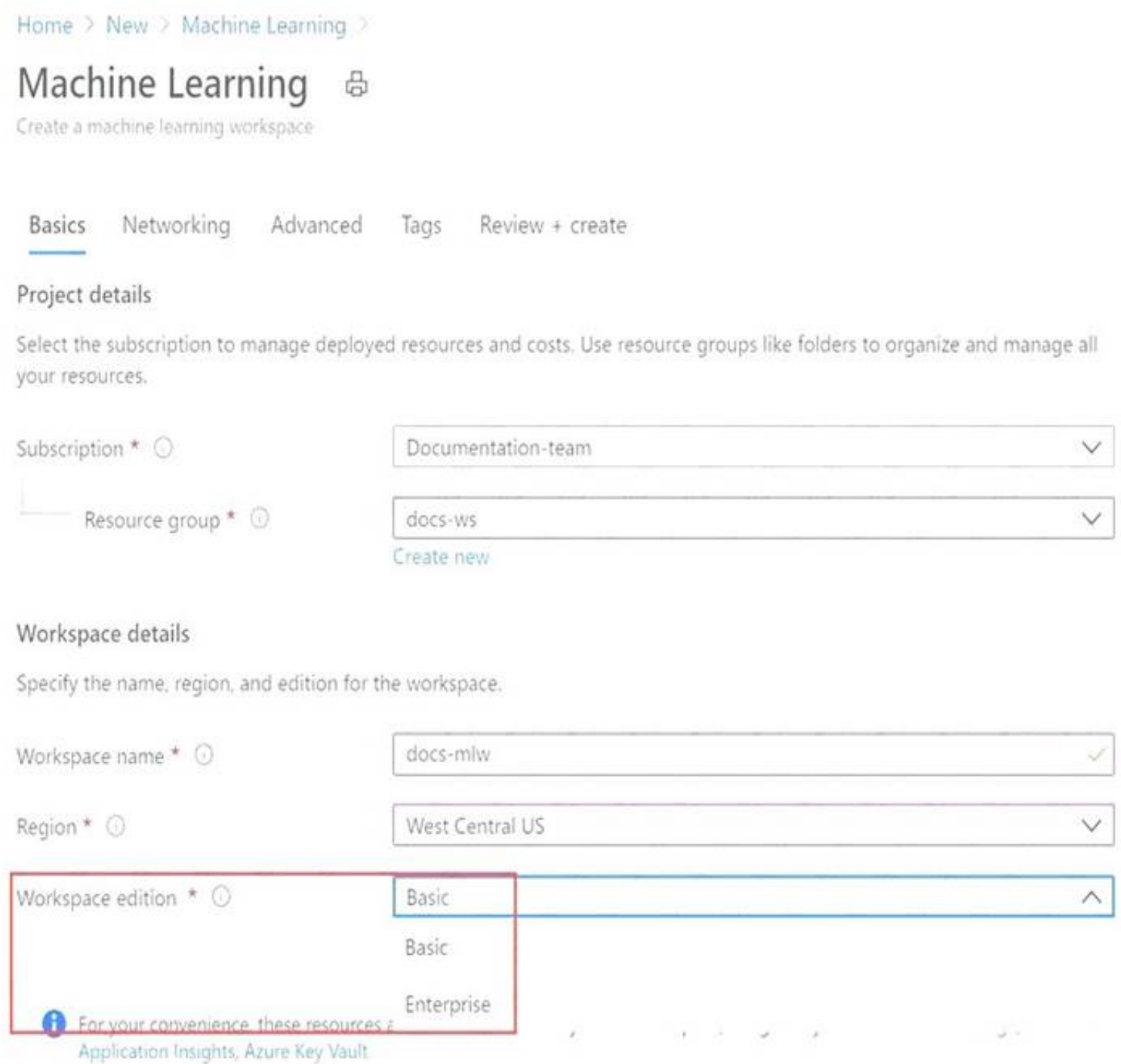
**Explanation:**
D: For Number of learning iterations, specify the maximum number of times the algorithm should process the training cases.
E: For Hidden layer specification, select the type of network architecture to create.
Between the input and output layers you can insert multiple hidden layers. Most predictive tasks can be accomplished easily with only one or a few hidden layers.
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/two-class-neural-network

**NEW QUESTION 15**
- (Exam Topic 3)
You create a new Azure subscription. No resources are provisioned in the subscription. You need to create an Azure Machine Learning workspace.
What are three possible ways to achieve this goal? Each correct answer presents a complete solution. NOTE: Each correct selection is worth one point.

A. Run Python code that uses the Azure ML SDK library and calls the Workspace.create method with name, subscription_id, resource_group, and location parameters.
B. Use an Azure Resource Management template that includes a Microsoft.MachineLearningServices/ workspaces resource and its dependencies.
C. Use the Azure Command Line Interface (CLI) with the Azure Machine Learning extension to call the az group create function with --name and --location parameters, and then the az ml workspace create function, specifying –w and –g parameters for the workspace name and resource group.
D. Navigate to Azure Machine Learning studio and create a workspace.
E. Run Python code that uses the Azure ML SDK library and calls the Workspace.get method with name, subscription_id, and resource_group parameters.

**Answer:** BCD

**Explanation:**
B: You can use an Azure Resource Manager template to create a workspace for Azure Machine Learning. Example:
{"type": "Microsoft.MachineLearningServices/workspaces",
…
C: You can create a workspace for Azure Machine Learning with Azure CLI Install the machine learning extension.
Create a resource group: az group create --name <resource-group-name> --location <location>
To create a new workspace where the services are automatically created, use the following command: az ml workspace create -w <workspace-name> -g <resource-group-name>
D: You can create and manage Azure Machine Learning workspaces in the Azure portal.
- Sign in to the Azure portal by using the credentials for your Azure subscription.
- In the upper-left corner of Azure portal, select + Create a resource.
- Use the search bar to find Machine Learning.
- Select Machine Learning.
- In the Machine Learning pane, select Create to begin.




Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-create-workspace-template https://docs.microsoft.com/en-us/azure/machine-learning/how-to-manage-workspace-cli https://docs.microsoft.com/en-us/azure/machine-learning/how-to-manage-workspace

**NEW QUESTION 18**
- (Exam Topic 3)
You run an experiment that uses an AutoMLConfig class to define an automated machine learning task with a maximum of ten model training iterations. The task will attempt to find the best performing model based on a metric named accuracy.
You submit the experiment with the following code:
You need to create Python code that returns the best model that is generated by the automated machine learning task. Which code segment should you use?
A)

```
best_model = automl_run.get_details()
```

B)

```
best_model = automl_run.get_output()[1]
```

C)

```
best_model = automl_run.get_file_names()[1]
```

D)

```
best_model = automl_run.get_metrics()
```

A. Option A
B. Option B
C. Option C
D. Option D

**Answer:** B

**Explanation:**
The get_output method returns the best run and the fitted model. Reference:
https://notebooks.azure.com/azureml/projects/azureml-getting-started/html/how-to-use-azureml/automated-mach

**NEW QUESTION 21**
- (Exam Topic 3)

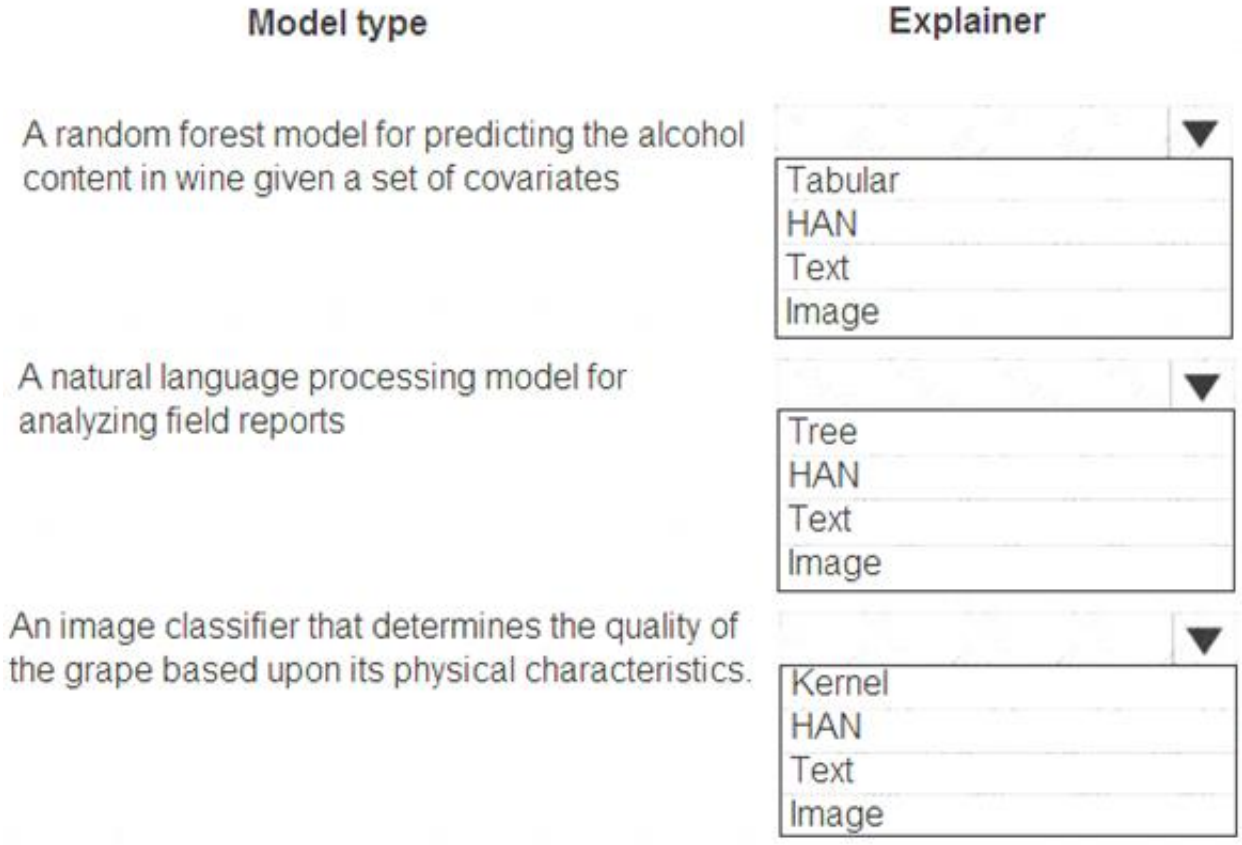
You plan to use a Deep Learning Virtual Machine (DLVM) to train deep learning models using Compute Unified Device Architecture (CUDA) computations.
You need to configure the DLVM to support CUDA. What should you implement?

A. Intel Software Guard Extensions (Intel SGX) technology
B. Solid State Drives (SSD)
C. Graphic Processing Unit (GPU)
D. Computer Processing Unit (CPU) speed increase by using overcloking
E. High Random Access Memory (RAM) configuration

**Answer:** C

**Explanation:**
A Deep Learning Virtual Machine is a pre-configured environment for deep learning using GPU instances.
References:
https://azuremarketplace.microsoft.com/en-au/marketplace/apps/microsoft-ads.dsvm-deep-learning

**NEW QUESTION 26**
- (Exam Topic 3)
You are hired as a data scientist at a winery. The previous data scientist used Azure Machine Learning. You need to review the models and explain how each model makes decisions.
Which explainer modules should you use? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.




A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
info based on
the given model and data sets. The meta explainers leverage all the libraries (SHAP, LIME, Mimic, etc.) that we have integrated or developed. The following are the meta explainers available in the SDK:
Tabular Explainer: Used with tabular datasets. Text Explainer: Used with text datasets. Image Explainer: Used with image datasets. Box 1: Tabular
Box 2: Text
Box 3: Image Reference:
https://medium.com/microsoftazure/automated-and-interpretable-machine-learning-d07975741298

**NEW QUESTION 31**
- (Exam Topic 3)
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: Remove the entire column that contains the missing data point. Does the solution meet the goal?

A. Yes
B. No

**Answer:** B

**Explanation:**
Use the Multiple Imputation by Chained Equations (MICE) method. 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

**NEW QUESTION 36**
- (Exam Topic 3)
You deploy a model in Azure Container Instance.
You must use the Azure Machine Learning SDK to call the model API.
You need to invoke the deployed model using native SDK classes and methods.
How should you complete the command? To answer, select the appropriate options in the answer areas.
NOTE: Each correct selection is worth one point.

```
from azureml.core import Workspace
[ ▼ ]
```
- from azureml.core.webservice import requests
- from azureml.core.webservice import Webservice
- from azureml.core.webservice import LocalWebservice

```
import json
ws = Workspace.from_config()
service_name = "mlmodel1-service"
service = Webservice(name=service_name, workspace=ws)
x_new = [[2,101.5,1,24,21], [1,89.7,4,41,21]]
input_json = json.dumps({"data": x_new})
[ ▼ ]
```
- predictions = service.run(input_json)
- predictions = requests.post(service.scoring_uri, input_json)
- predictions = service.deserialize(ws, input_json)

A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Box 1: from azureml.core.webservice import Webservice
The following code shows how to use the SDK to update the model, environment, and entry script for a web service to Azure Container Instances:
from azureml.core import Environment
from azureml.core.webservice import Webservice
from azureml.core.model import Model, InferenceConfig Box 2: predictions = service.run(input_json)
Example: The following code demonstrates sending data to the service: import json
test_sample = json.dumps({'data': [ [1, 2, 3, 4, 5, 6, 7, 8, 9, 10],
[10, 9, 8, 7, 6, 5, 4, 3, 2, 1]
]})
test_sample = bytes(test_sample, encoding='utf8') prediction = service.run(input_data=test_sample)
print(prediction) Reference:
https://docs.microsoft.com/bs-latn-ba/azure/machine-learning/how-to-deploy-azure-container-instance https://docs.microsoft.com/en-us/azure/machine-learning/how-to-troubleshoot-deployment

**NEW QUESTION 37**
- (Exam Topic 3)
You are creating a machine learning model. You need to identify outliers in the data.
Which two visualizations can you use? Each correct answer presents a complete solution.
NOTE: Each correct selection is worth one point. NOTE: Each correct selection is worth one point.

A. box plot
B. scatter
C. random forest diagram
D. Venn diagram
E. ROC curve

**Answer:** AB

**Explanation:**
The box-plot algorithm can be used to display outliers.
One other way to quickly identify Outliers visually is to create scatter plots. References:
https://blogs.msdn.microsoft.com/azuredev/2017/05/27/data-cleansing-tools-in-azure-machine-learning/

**NEW QUESTION 40**
- (Exam Topic 3)
You use the Azure Machine Learning service to create a tabular dataset named training.data. You plan to use this dataset in a training script.
You create a variable that references the dataset using the following code: training_ds = workspace.datasets.get("training_data")
You define an estimator to run the script.
You need to set the correct property of the estimator to ensure that your script can access the training.data dataset
Which property should you set?
A)

```
inputs = [training_ds.as_named_input('training_ds')]
```

B)

```
script_params = {"--training_ds":training_ds}
```

C)

```
environment_definition = {"training_data":training_ds}
```

D)

```
source_directory = training_ds
```

A. Option A
B. Option B
C. Option C
D. Option D

**Answer:** A

**Explanation:**
Example:
# Get the training dataset
diabetes_ds = ws.datasets.get("Diabetes Dataset")
# Create an estimator that uses the remote compute hyper_estimator = SKLearn(source_directory=experiment_folder, inputs=[diabetes_ds.as_named_input('diabetes')], # Pass the dataset as an input compute_target = cpu_cluster, conda_packages=['pandas','ipykernel','matplotlib'], pip_packages=['azureml-sdk','argparse','pyarrow'], entry_script='diabetes_training.py')
Reference:
https://notebooks.azure.com/GraemeMalcolm/projects/azureml-primers/html/04%20-%20Optimizing%20Model

**NEW QUESTION 45**
- (Exam Topic 3)
You are conducting feature engineering to prepuce data for further analysis. The data includes seasonal patterns on inventory requirements.
You need to select the appropriate method to conduct feature engineering on the data. Which method should you use?

A. Exponential Smoothing (ETS) function.
B. One Class Support Vector Machine module
C. Time Series Anomaly Detection module
D. Finite Impulse Response (FIR) Filter module.

**Answer:** D

**NEW QUESTION 47**
- (Exam Topic 3)
You have a dataset that includes home sales data for a city. The dataset includes the following columns.

| Name | Description |
|---|---|
| Price | The sales price for the house. |
| Bedrooms | The number of bedrooms in the house. |
| Size | The size of the house in square feet. |
| HasGarage | A binary value indicating whether or not the house has a garage. |
| HomeType | The category of home, for example, apartment, townhouse, single-family home. |

Each row in the dataset corresponds to an individual home sales transaction.
You need to use automated machine learning to generate the best model for predicting the sales price based on the features of the house.
Which values should you use? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.


A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Box 1: Regression
Regression is a supervised machine learning technique used to predict numeric values. Box 2: Price
Reference:
https://docs.microsoft.com/en-us/learn/modules/create-regression-model-azure-machine-learning-designer

**NEW QUESTION 52**
- (Exam Topic 3)
You are preparing to use the Azure ML SDK to run an experiment and need to create compute. You run the following code:

```
from azureml.core.compute import ComputeTarget, AmlCompute
from azureml.core.compute_target import ComputeTargetException
ws = Workspace.from_config()
cluster_name = 'aml-cluster'
try:
  training_compute = ComputeTarget(workspace=ws, name=cluster_name)
except ComputeTargetException:
  compute_config = AmlCompute.provisioning_configuration(vm_size='STANDARD_D2_V2', vm_priority='lowpriority',
max_nodes=4)
  training_compute = ComputeTarget.create(ws, cluster_name, compute_config)
  training_compute.wait_for_completion(show_output=True)
```

For each of the following statements, select Yes if the statement is true. Otherwise, select No.
NOTE: Each correct selection is worth one point.

| | Yes | No |
|---|---|---|
| If a training cluster named aml-cluster already exists in the workspace, it will be deleted and replaced. | O | O |
| The `wait_for_completion()` method will not return until the aml-cluster compute has four active nodes. | O | O |
| If the code creates a new aml-cluster compute target, it may be preempted due to capacity constraints. | O | O |
| The aml-cluster compute target is deleted from the workspace after the training experiment completes. | O | O |

A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Box 1: No
If a training cluster already exists it will be used. Box 2: Yes
The wait_for_completion method waits for the current provisioning operation to finish on the cluster. Box 3: Yes
Low Priority VMs use Azure's excess capacity and are thus cheaper but risk your run being pre-empted.
Box 4: No
Need to use training_compute.delete() to deprovision and delete the AmlCompute target. Reference:
https://notebooks.azure.com/azureml/projects/azureml-getting-started/html/how-to-use-azureml/training/train-on https://docs.microsoft.com/en-us/python/api/azureml-core/azureml.core.compute.computetarget

**NEW QUESTION 57**
- (Exam Topic 3)
You create a batch inference pipeline by using the Azure ML SDK. You run the pipeline by using the following code:
from azureml.pipeline.core import Pipeline
from azureml.core.experiment import Experiment
pipeline = Pipeline(workspace=ws, steps=[parallelrun_step]) pipeline_run = Experiment(ws, 'batch_pipeline').submit(pipeline) You need to monitor the progress of the pipeline execution.
What are two possible ways to achieve this goal? Each correct answer presents a complete solution.
NOTE: Each correct selection is worth one point.

A. Run the following code in a notebook:

```
from azureml.contrib.interpret.explanation.explanation_client import ExplanationClient
client = ExplanationClient.from_run(pipeline_run)
explanation = client.download_model_explanation()
explanation = client.download_model_explanation(top_k=4)
global_importance_values = explanation.get_ranked_global_values()
global_importance_names = explanation.get_ranked_global_names()
print('global importance values: {}'.format(global_importance_values))
print('global importance names: {}'.format(global_importance_names))
```

B. Use the Inference Clusters tab in Machine Learning Studio.

C. Use the Activity log in the Azure portal for the Machine Learning workspace.

D. Run the following code in a notebook:

```
from azureml.widgets import RunDetails
RunDetails(pipeline_run).show()
```

E. Run the following code and monitor the console output from the PipelineRun object:

```
pipeline_run.wait_for_completion(show_output=True)
```

A. Option A
B. Option B
C. Option C
D. Option D
E. Option E

**Answer:** DE

**Explanation:**
A batch inference job can take a long time to finish. This example monitors progress by using a Jupyter widget. You can also manage the job's progress by using:

- Azure Machine Learning Studio.
- Console output from the PipelineRun object. from azureml.widgets import RunDetails RunDetails(pipeline_run).show() pipeline_run.wait_for_completion(show_output=True) Reference: https://docs.microsoft.com/en-us/azure/machine-learning/how-to-use-parallel-run-step#monitor-the-parallel-run

**NEW QUESTION 58**
- (Exam Topic 3)
You use Azure Machine Learning designer to create a training pipeline for a regression model.
You need to prepare the pipeline for deployment as an endpoint that generates predictions asynchronously for a dataset of input data values.
What should you do?

A. Clone the training pipeline.
B. Create a batch inference pipeline from the training pipeline.
C. Create a real-time inference pipeline from the training pipeline.
D. Replace the dataset in the training pipeline with an Enter Data Manually module.

**Answer:** C

**Explanation:**
You must first convert the training pipeline into a real-time inference pipeline. This process removes training modules and adds web service inputs and outputs to handle requests.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/tutorial-designer-automobile-price-deploy https://docs.microsoft.com/en-us/azure/machine-learning/algorithm-module-reference/enter-data-manually

**NEW QUESTION 60**
- (Exam Topic 3)
You are building a binary classification model by using a supplied training set. The training set is imbalanced between two classes.
You need to resolve the data imbalance.
What are three possible ways to achieve this goal? Each correct answer presents a complete solution NOTE: Each correct selection is worth one point.

A. Penalize the classification
B. Resample the data set using under sampling or oversampling
C. Generate synthetic samples in the minority class.
D. Use accuracy as the evaluation metric of the model.
E. Normalize the training feature set.

**Answer:** ABD

**Explanation:**
References:
https://machinelearningmastery.com/tactics-to-combat-imbalanced-classes-in-your-machine-learning-dataset/

**NEW QUESTION 64**
- (Exam Topic 3)
You use Azure Machine Learning to train and register a model.
You must deploy the model into production as a real-time web service to an inference cluster named service-compute that the IT department has created in the Azure Machine Learning workspace.

Client applications consuming the deployed web service must be authenticated based on their Azure Active Directory service principal.
You need to write a script that uses the Azure Machine Learning SDK to deploy the model. The necessary modules have been imported.
How should you complete the code? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.

```
# Assume the necessary modules have been imported
deploy_target = [AksCompute | AmlCompute | RemoteCompute | BatchCompute] (ws, "service-compute")
deployment_config = [AksWebservice | AciWebservice | LocalWebService].deploy_configuration(cpu_cores=1, memory_gb=1,
[token_auth_enabled=True | token_auth_enabled=False | auth_enabled=True | auth_enabled=False])
service = Model.deploy(ws, "ml-service",
 [model], inference_config, deployment_config, deploy_target)
service.wait_for_deployment(show_output = True)
```

A. Mastered
B. Not Mastered

**Answer:** A

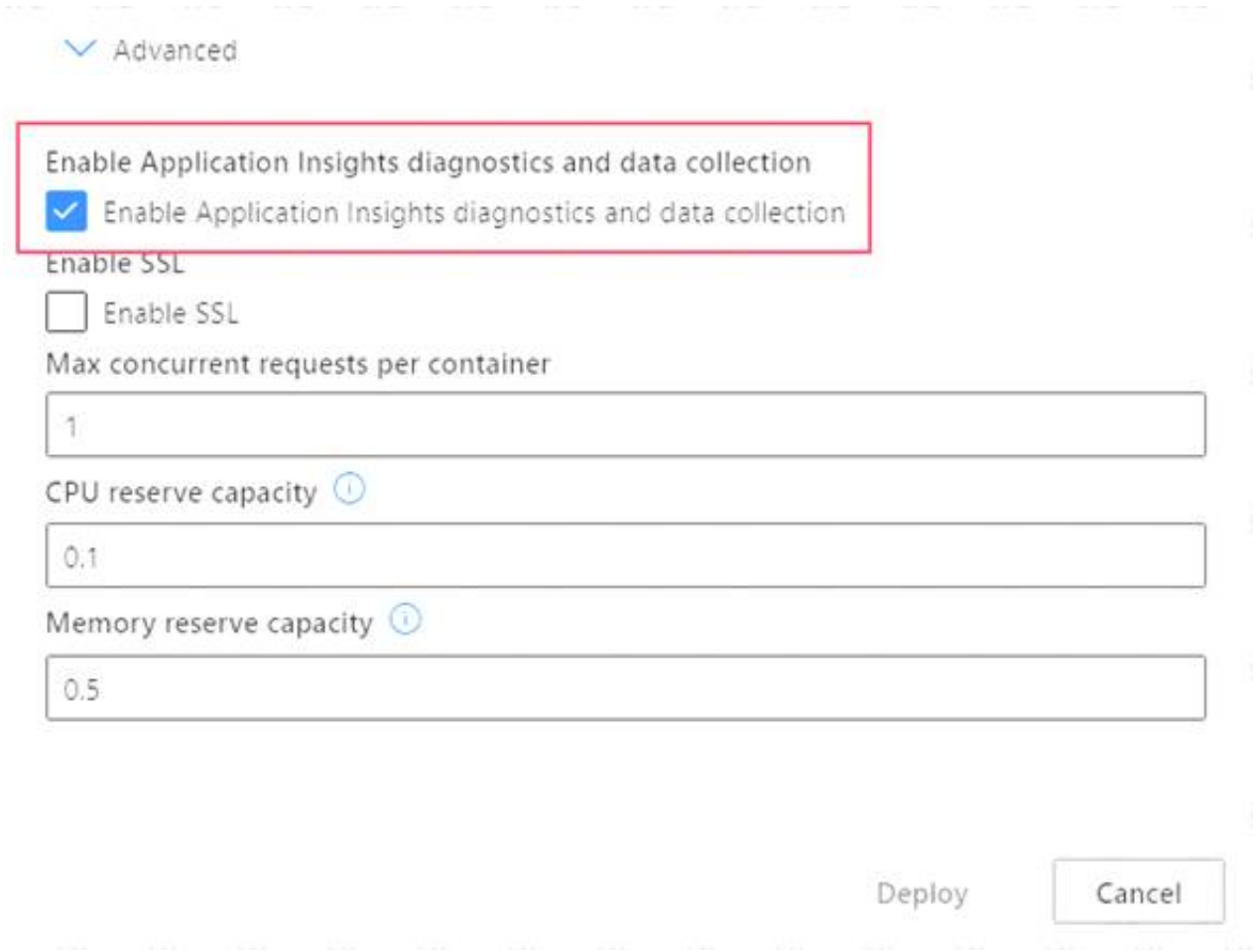
**Explanation:**
Box 1: AksCompute Example:
aks_target = AksCompute(ws,"myaks")
# If deploying to a cluster configured for dev/test, ensure that it was created with enough
# cores and memory to handle this deployment configuration. Note that memory is also used by
# things such as dependencies and AML components.
deployment_config = AksWebservice.deploy_configuration(cpu_cores = 1, memory_gb = 1)
service = Model.deploy(ws, "myservice", [model], inference_config, deployment_config, aks_target) Box 2: AksWebservice
Box 3: token_auth_enabled=Yes
Whether or not token auth is enabled for the Webservice.
Note: A Service principal defined in Azure Active Directory (Azure AD) can act as a principal on which authentication and authorization policies can be enforced in Azure Databricks.
The Azure Active Directory Authentication Library (ADAL) can be used to programmatically get an Azure AD access token for a user.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-deploy-azure-kubernetes-service https://docs.microsoft.com/en-us/azure/databricks/dev-tools/api/latest/aad/service-prin-aad-token

**NEW QUESTION 67**
- (Exam Topic 3)
You deploy a real-time inference service for a trained model.
The deployed model supports a business-critical application, and it is important to be able to monitor the data submitted to the web service and the predictions the data generates.
You need to implement a monitoring solution for the deployed model using minimal administrative effort. What should you do?

A. View the explanation for the registered model in Azure ML studio.
B. Enable Azure Application Insights for the service endpoint and view logged data in the Azure portal.
C. Create an ML Flow tracking URI that references the endpoint, and view the data logged by ML Flow.
D. View the log files generated by the experiment used to train the model.

**Answer:** B

**Explanation:**
Configure logging with Azure Machine Learning studio
You can also enable Azure Application Insights from Azure Machine Learning studio. When you're ready to deploy your model as a web service, use the following steps to enable Application Insights:
* 1. Sign in to the studio at https://ml.azure.com.
* 2. Go to Models and select the model you want to deploy.
* 3. Select +Deploy.
* 4. Populate the Deploy model form.
* 5. Expand the Advanced menu.
* 6. Select Enable Application Insights diagnostics and data collection.




Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-enable-app-insights

**NEW QUESTION 72**
- (Exam Topic 3)
You use the following code to run a script as an experiment in Azure Machine Learning:

```
from azureml.core import Workspace, Experiment, Run
from azureml.core import RunConfig, ScriptRunConfig
ws = Workspace.from_config()
run_config = RunConfiguration()
run_config.target='local'
script_config = ScriptRunConfig(source_directory='./script', script='experiment.py', run_config=run_config)
experiment = Experiment(workspace=ws, name='script experiment')
run = experiment.submit(config=script_config)
run.wait_for_completion()
```

You must identify the output files that are generated by the experiment run. You need to add code to retrieve the output file names.
Which code segment should you add to the script?

A. files = run.get_properties()
B. files= run.get_file_names()
C. files = run.get_details_with_logs()
D. files = run.get_metrics()
E. files = run.get_details()

**Answer:** B

**Explanation:**
You can list all of the files that are associated with this run record by called run.get_file_names() Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-track-experiments

**NEW QUESTION 74**
- (Exam Topic 3)
You plan to create a speech recognition deep learning model. The model must support the latest version of Python.
You need to recommend a deep learning framework for speech recognition to include in the Data Science Virtual Machine (DSVM).
What should you recommend?

A. Apache Drill
B. Tensorflow
C. Rattle
D. Weka

**Answer:** B

**Explanation:**
TensorFlow is an open source library for numerical computation and large-scale machine learning. It uses Python to provide a convenient front-end API for building applications with the framework
TensorFlow can train and run deep neural networks for handwritten digit classification, image recognition, word embeddings, recurrent neural networks, sequence-to-sequence models for machine translation, natural language processing, and PDE (partial differential equation) based simulations.
References:
https://www.infoworld.com/article/3278008/what-is-tensorflow-the-machine-learning-library-explained.html

**NEW QUESTION 77**
- (Exam Topic 3)
You plan to deliver a hands-on workshop to several students. The workshop will focus on creating data visualizations using Python. Each student will use a device that has internet access.
Student devices are not configured for Python development. Students do not have administrator access to install software on their devices. Azure subscriptions are not available for students.
You need to ensure that students can run Python-based data visualization code. Which Azure tool should you use?

A. Anaconda Data Science Platform
B. Azure BatchAI
C. Azure Notebooks
D. Azure Machine Learning Service

**Answer:** C

**Explanation:**
References:
https://notebooks.azure.com/

**NEW QUESTION 82**
- (Exam Topic 3)
You plan to run a Python script as an Azure Machine Learning experiment.
The script must read files from a hierarchy of folders. The files will be passed to the script as a dataset argument.
You must specify an appropriate mode for the dataset argument.
Which two modes can you use? Each correct answer presents a complete solution.
NOTE: Each correct selection is worth one point.

A. to_pandas_dataframe ()
B. as_download()
C. as_upload()
D. as mount ()

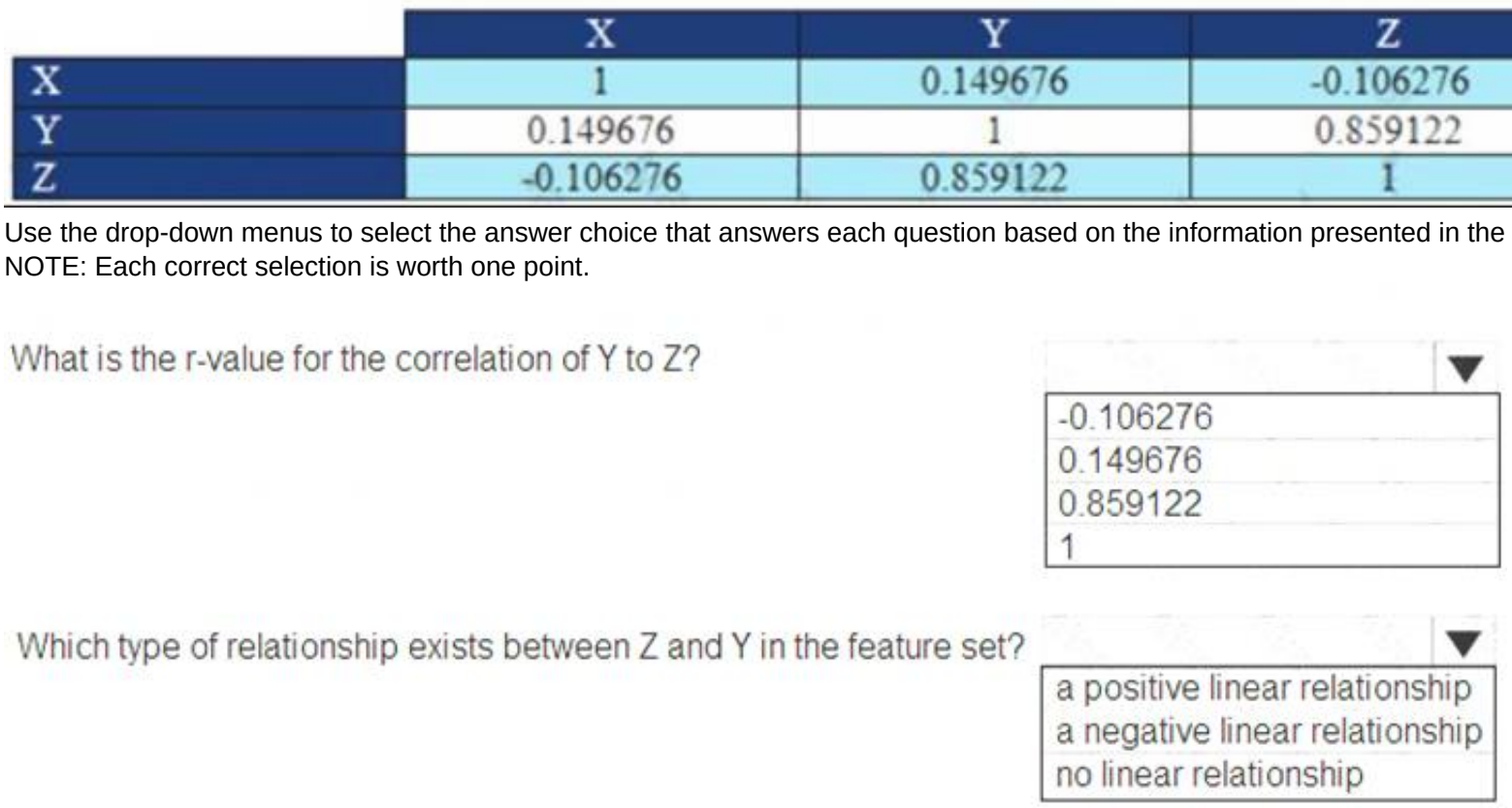
**Answer:** B

**Explanation:**
Reference:
https://docs.microsoft.com/en-us/python/api/azureml-core/azureml.data.filedataset?view=azure-ml-py

**NEW QUESTION 85**
- (Exam Topic 3)
You have a feature set containing the following numerical features: X, Y, and Z.
The Poisson correlation coefficient (r-value) of X, Y, and Z features is shown in the following image:

| | X | Y | Z |
|---|---|---|---|
| X | 1 | 0.149676 | -0.106276 |
| Y | 0.149676 | 1 | 0.859122 |
| Z | -0.106276 | 0.859122 | 1 |

Use the drop-down menus to select the answer choice that answers each question based on the information presented in the graphic.
NOTE: Each correct selection is worth one point.



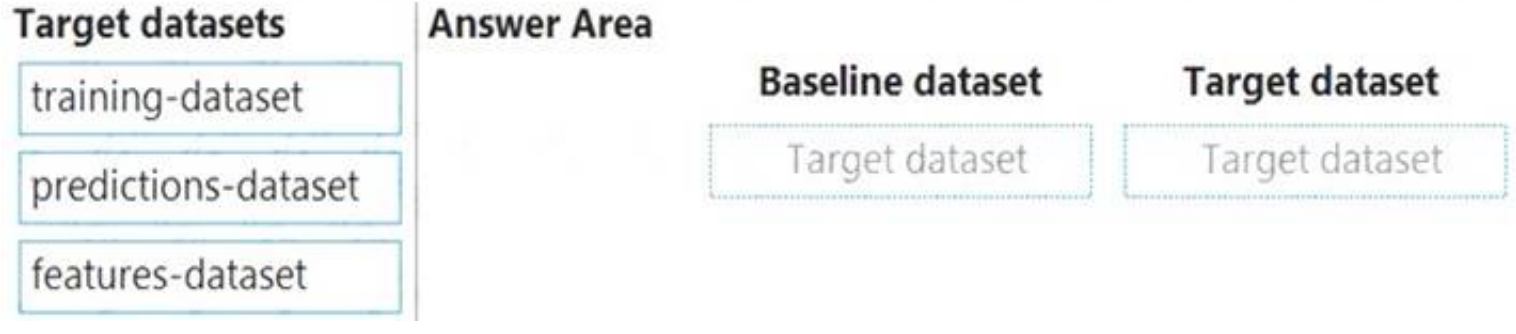

A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Box 1: 0.859122
Box 2: a positively linear relationship
+1 indicates a strong positive linear relationship
-1 indicates a strong negative linear correlation
0 denotes no linear relationship between the two variables. References:

https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/compute-linear-correlation

**NEW QUESTION 88**
- (Exam Topic 3)
You previously deployed a model that was trained using a tabular dataset named training-dataset, which is based on a folder of CSV files.
Over time, you have collected the features and predicted labels generated by the model in a folder containing a CSV file for each month. You have created two tabular datasets based on the folder containing the inference data: one named predictions-dataset with a schema that matches the training data exactly, including the predicted label; and another named features-dataset with a schema containing all of the feature columns and a timestamp column based on the filename, which includes the day, month, and year.
You need to create a data drift monitor to identify any changing trends in the feature data since the model was trained. To accomplish this, you must define the required datasets for the data drift monitor.
Which datasets should you use to configure the data drift monitor? To answer, drag the appropriate datasets to the correct data drift monitor options. Each source may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.
NOTE: Each correct selection is worth one point.

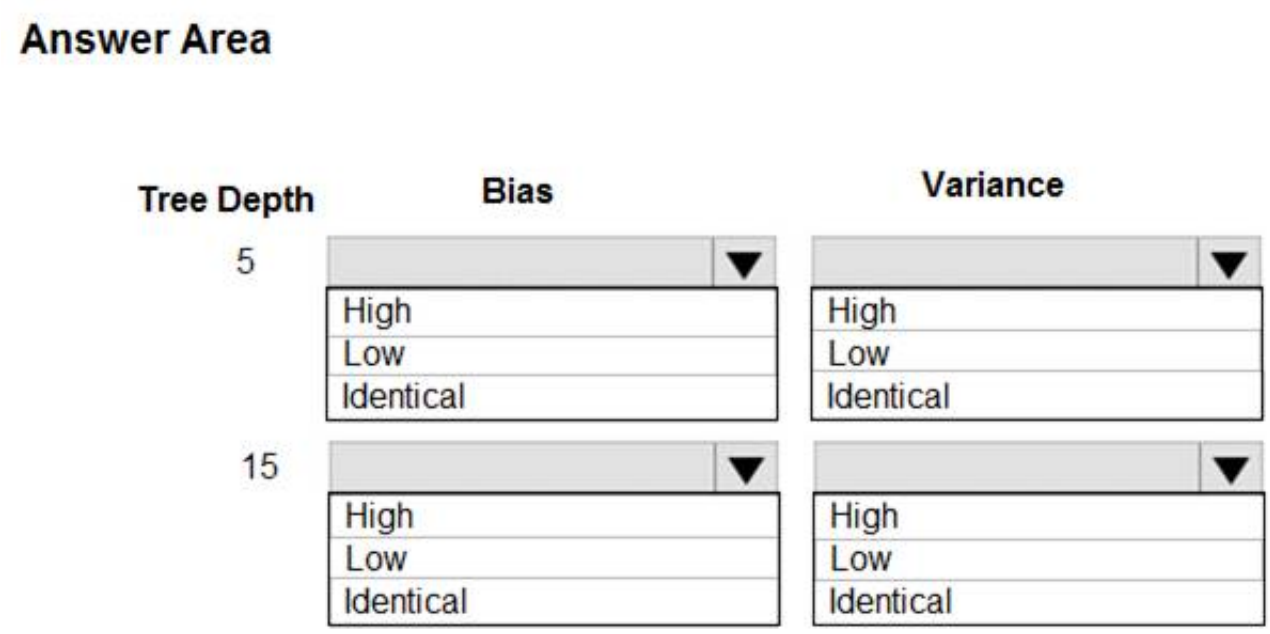



A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Text Description automatically generated with medium confidence
Box 1: training-dataset
Baseline dataset - usually the training dataset for a model. Box 2: predictions-dataset
Target dataset - usually model input data - is compared over time to your baseline dataset. This comparison means that your target dataset must have a timestamp column specified.
The monitor will compare the baseline and target datasets. Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-monitor-datasets

**NEW QUESTION 93**
- (Exam Topic 3)
You are using a decision tree algorithm. You have trained a model that generalizes well at a tree depth equal to 10.
You need to select the bias and variance properties of the model with varying tree depth values.
Which properties should you select for each tree depth? To answer, select the appropriate options in the answer area.




A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
In decision trees, the depth of the tree determines the variance. A complicated decision tree (e.g. deep) has low bias and high variance.
Note: In statistics and machine learning, the bias–variance tradeoff is the property of a set of predictive models whereby models with a lower bias in parameter estimation have a higher variance of the parameter estimates across samples, and vice versa. Increasing the bias will decrease the variance. Increasing the variance will decrease the bias.
References:
https://machinelearningmastery.com/gentle-introduction-to-the-bias-variance-trade-off-in-machine-learning/

**NEW QUESTION 98**

- (Exam Topic 3)
You use an Azure Machine Learning workspace.
You have a trained model that must be deployed as a web service. Users must authenticate by using Azure Active Directory.
What should you do?

A. Deploy the model to Azure Kubernetes Service (AKS). During deployment, set the token_auth_enabled parameter of the target configuration object to true
B. Deploy the model to Azure Container Instance
C. During deployment, set the auch_enabled parameter of the target configuration object to true
D. Deploy the model to Azure Container Instance
E. During deployment, set the coken_auch_enabled parameter of the target configuration object to true
F. Deploy the model to Azure Kubernetes Service (AKS). During deployment, set the auc
G. enabled parameter of the target configuration object to true

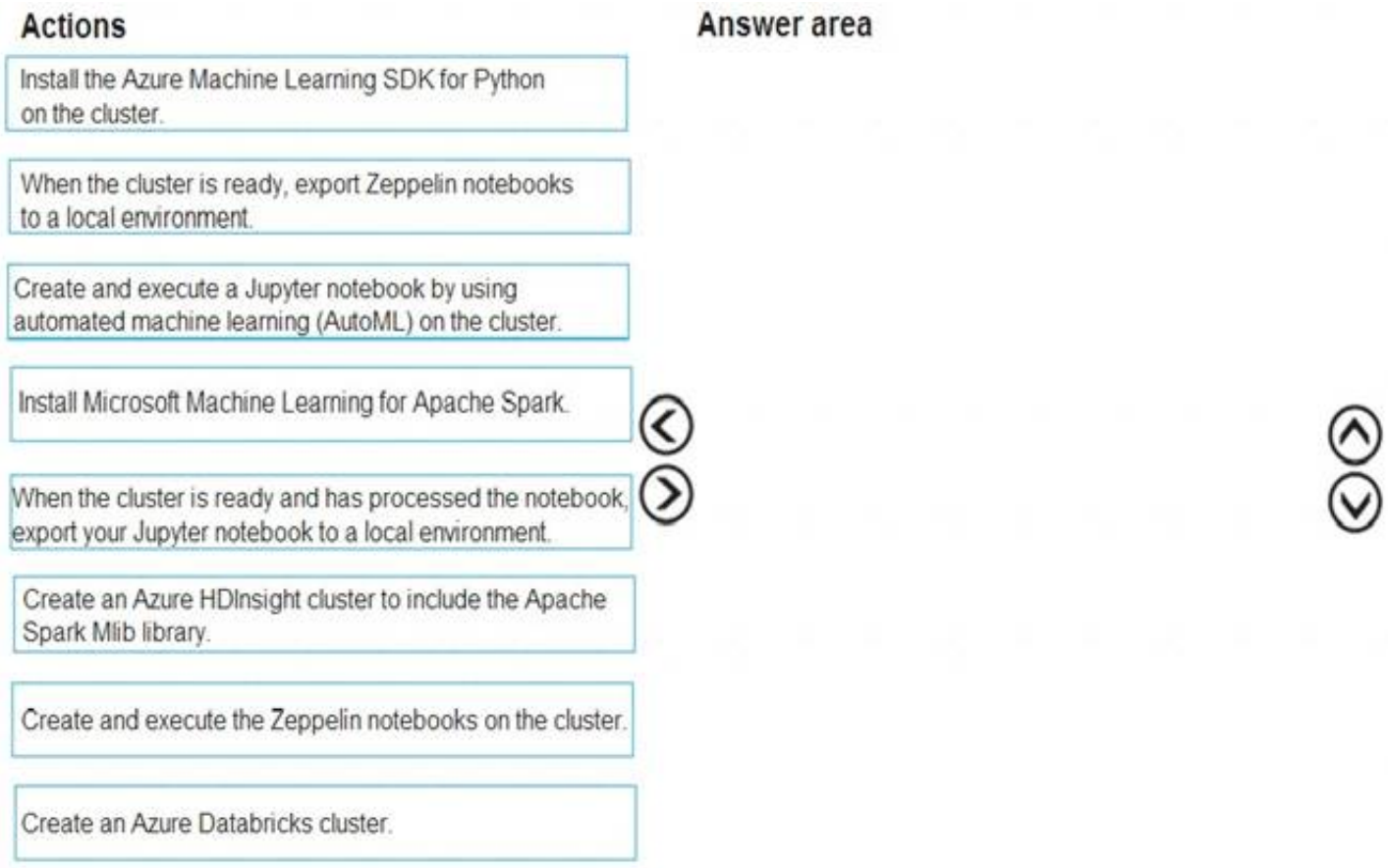
**Answer:** A

**Explanation:**
To control token authentication, use the token_auth_enabled parameter when you create or update a deployment
Token authentication is disabled by default when you deploy to Azure Kubernetes Service.
Note: The model deployments created by Azure Machine Learning can be configured to use one of two authentication methods:
key-based: A static key is used to authenticate to the web service.
token-based: A temporary token must be obtained from the Azure Machine Learning workspace (using Azure Active Directory) and used to authenticate to the web service.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-authenticate-web-service

**NEW QUESTION 102**
- (Exam Topic 3)
You are building an intelligent solution using machine learning models. The environment must support the following requirements:

- Data scientists must build notebooks in a cloud environment
- Data scientists must use automatic feature engineering and model building in machine learning pipelines.
- Notebooks must be deployed to retrain using Spark instances with dynamic worker allocation.
- Notebooks must be exportable to be version controlled locally.

You need to create the environment.
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.




A. Mastered
B. Not Mastered

**Answer:** A

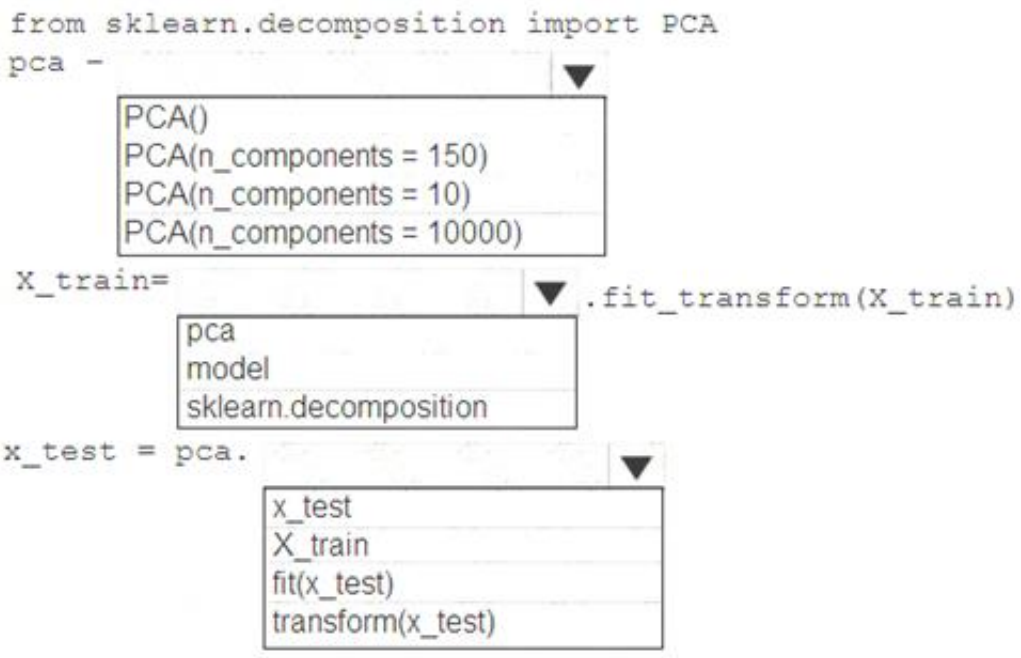
**Explanation:**
Step 1: Create an Azure HDInsight cluster to include the Apache Spark Mlib library
Step 2: Install Microsot Machine Learning for Apache Spark You install AzureML on your Azure HDInsight cluster.
Microsoft Machine Learning for Apache Spark (MMLSpark) provides a number of deep learning and data science tools for Apache Spark, including seamless integration of Spark Machine Learning pipelines with Microsoft Cognitive Toolkit (CNTK) and OpenCV, enabling you to quickly create powerful, highly-scalable predictive and analytical models for large image and text datasets.
Step 3: Create and execute the Zeppelin notebooks on the cluster
Step 4: When the cluster is ready, export Zeppelin notebooks to a local environment. Notebooks must be exportable to be version controlled locally.
References:
https://docs.microsoft.com/en-us/azure/hdinsight/spark/apache-spark-zeppelin-notebook https://azuremlbuild.blob.core.windows.net/pysparkapi/intro.html

**NEW QUESTION 106**

- (Exam Topic 3)
You have a dataset created for multiclass classification tasks that contains a normalized numerical feature set with 10,000 data points and 150 features.
You use 75 percent of the data points for training and 25 percent for testing. You are using the scikit-learn machine learning library in Python. You use X to denote the feature set and Y to denote class labels.
You create the following Python data frames:

| Name | Description |
|---|---|
| X_train | training feature set |
| Y_train | training class labels |
| x_train | testing feature set |
| y_train | testing class labels |

You need to apply the Principal Component Analysis (PCA) method to reduce the dimensionality of the feature set to 10 features in both training and testing sets.
How should you complete the code segment? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.

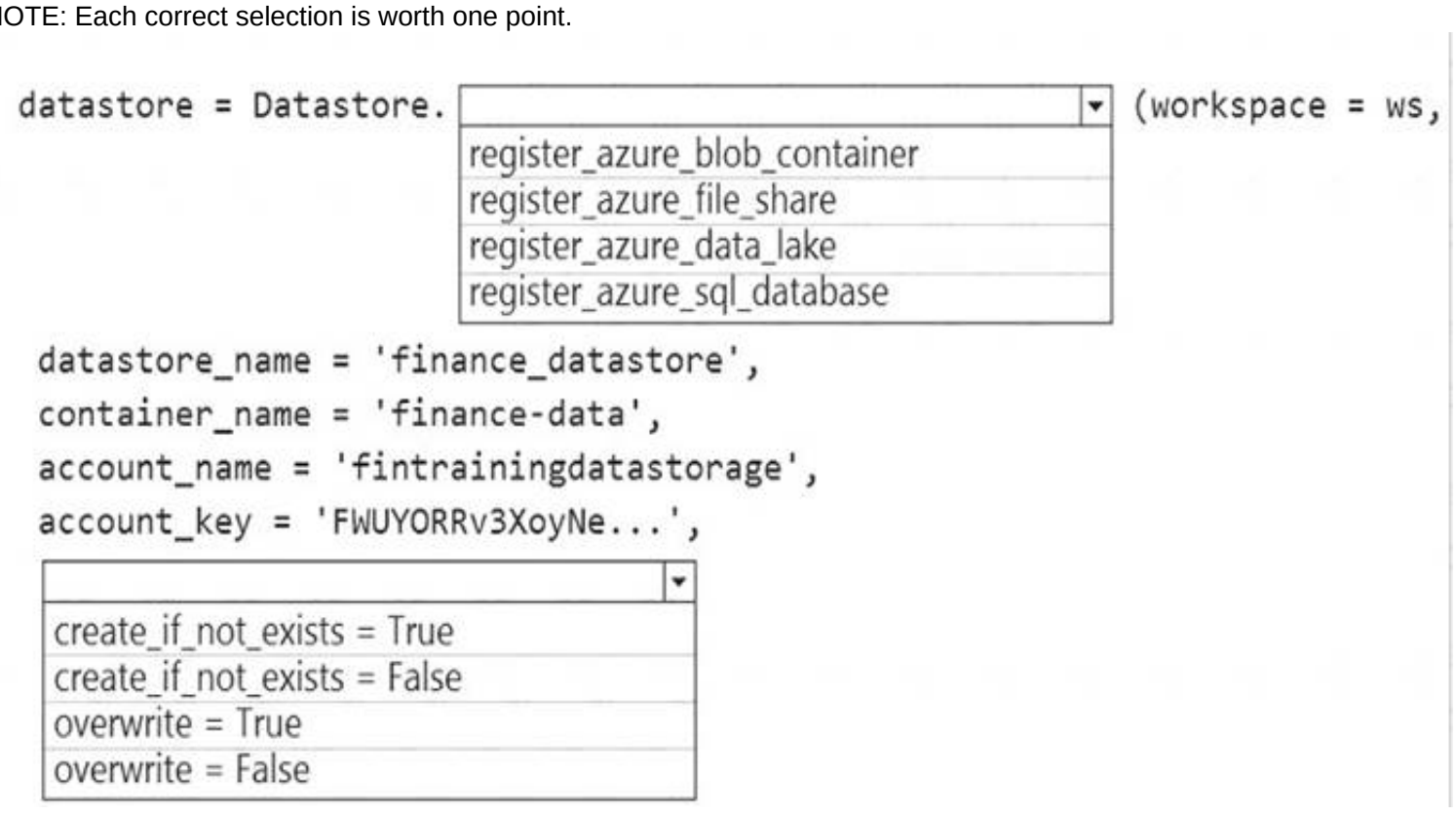



A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Box 1: PCA(n_components = 10)
Need to reduce the dimensionality of the feature set to 10 features in both training and testing sets. Example:
from sklearn.decomposition import PCA pca = PCA(n_components=2) ;2 dimensions principalComponents = pca.fit_transform(x)
Box 2: pca
fit_transform(X[, y])fits the model with X and apply the dimensionality reduction on X. Box 3: transform(x_test)
transform(X) applies dimensionality reduction to X. References:
https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html

**NEW QUESTION 110**
- (Exam Topic 3)
The finance team asks you to train a model using data in an Azure Storage blob container named finance-data. You need to register the container as a datastore in an Azure Machine Learning workspace and ensure that an
error will be raised if the container does not exist.
How should you complete the code? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.



A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Box 1: register_azure_blob_container
Register an Azure Blob Container to the datastore.
Box 2: create_if_not_exists = False
Create the file share if it does not exists, defaults to False. Reference:
https://docs.microsoft.com/en-us/python/api/azureml-core/azureml.core.datastore.datastore

**NEW QUESTION 114**
- (Exam Topic 3)
You create a machine learning model by using the Azure Machine Learning designer. You publish the model as a real-time service on an Azure Kubernetes Service (AKS) inference compute cluster. You make no changes to the deployed endpoint configuration.
You need to provide application developers with the information they need to consume the endpoint.
Which two values should you provide to application developers? Each correct answer presents part of the solution.
NOTE: Each correct selection is worth one point.

A. The name of the AKS cluster where the endpoint is hosted.
B. The name of the inference pipeline for the endpoint.
C. The URL of the endpoint.
D. The run ID of the inference pipeline experiment for the endpoint.
E. The key for the endpoint.

**Answer:** CE

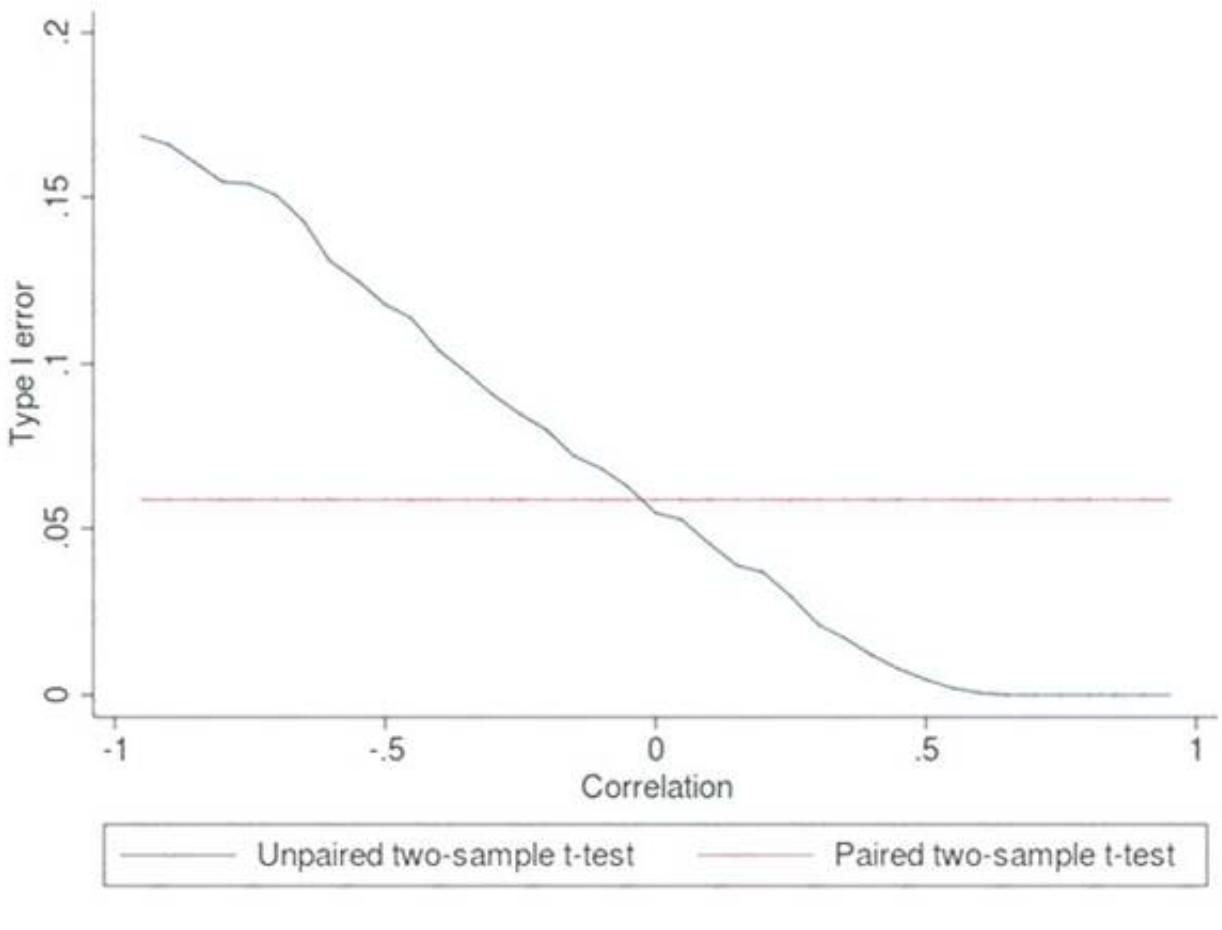
**Explanation:**
Deploying an Azure Machine Learning model as a web service creates a REST API endpoint. You can send data to this endpoint and receive the prediction returned by the model.
You create a web service when you deploy a model to your local environment, Azure Container Instances, Azure Kubernetes Service, or field-programmable gate arrays (FPGA). You retrieve the URI used to access the web service by using the Azure Machine Learning SDK. If authentication is enabled, you can also use the SDK to get the authentication keys or tokens.
Example:
# URL for the web service
scoring_uri = '<your web service URI>'
# If the service is authenticated, set the key or token key = '<your key or token>'
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-consume-web-service

**NEW QUESTION 118**
- (Exam Topic 3)
You are determining if two sets of data are significantly different from one another by using Azure Machine Learning Studio.
Estimated values in one set of data may be more than or less than reference values in the other set of data. You must produce a distribution that has a constant Type I error as a function of the correlation.
You need to produce the distribution.
Which type of distribution should you produce?

A. Paired t-test with a two-tail option
B. Unpaired t-test with a two tail option
C. Paired t-test with a one-tail option
D. Unpaired t-test with a one-tail option

**Answer:** A

**Explanation:**
Choose a one-tail or two-tail test. The default is a two-tailed test. This is the most common type of test, in which the expected distribution is symmetric around zero.
Example: Type I error of unpaired and paired two-sample t-tests as a function of the correlation. The simulated random numbers originate from a bivariate normal distribution with a variance of 1.




Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/test-hypothesis-using-t-test https://en.wikipedia.org/wiki/Student%27s_t-test

**NEW QUESTION 122**
- (Exam Topic 3)
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 create an Azure Machine Learning service datastore in a workspace. The datastore contains the following files:
• /data/2018/Q1 .csv
• /data/2018/Q2.csv
• /data/2018/Q3.csv
• /data/2018/Q4.csv
• /data/2019/Q1.csv
All files store data in the following format: id,M,f2,l
1,1,2,0
2,1,1,1
32,10
You run the following code:

```
data_store = Datastore.register_azure_blob_container(workspace=ws,
 datastore_name='data_store',
 container_name='quarterly_data',
 account_name='companydata',
 account_key='NRPxk8duxbM3…'
 create_if_not_exists=False)
```

You need to create a dataset named training_data and load the data from all files into a single data frame by using the following code:

```
data_frame = training_data.to_pandas_dataframe()
```

Solution: Run the following code:

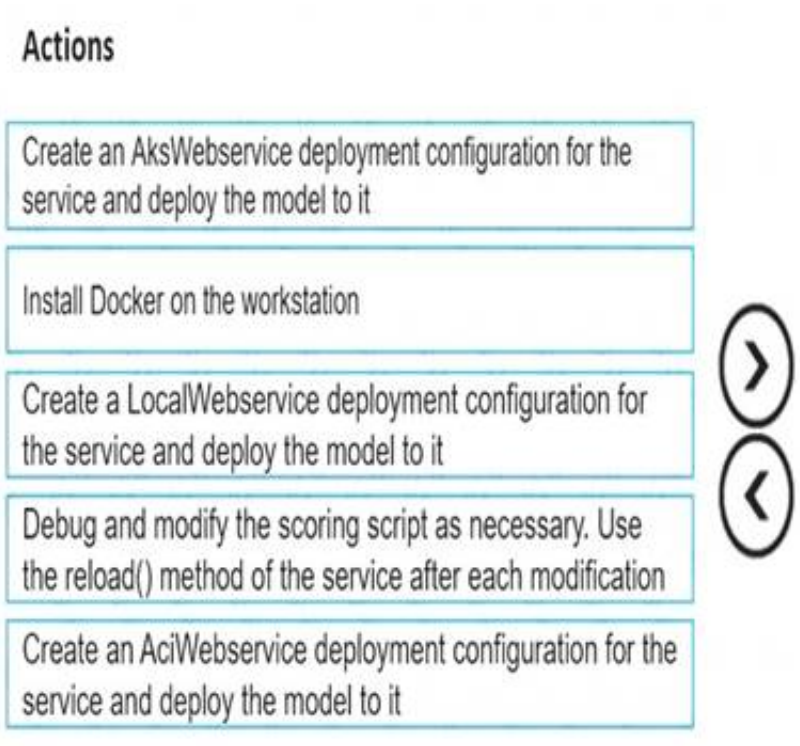
```
from azureml.core import Dataset
paths = (data_store,'data/*/*.csv')
training_data = Dataset.Tabular.from_delimited_files(paths)
```

Does the solution meet the goal?

A. Yes
B. No

**Answer:** A

**NEW QUESTION 127**
- (Exam Topic 3)
You train and register a model by using the Azure Machine Learning SDK on a local workstation. Python 3.6 and Visual Studio Code are installed on the workstation.
When you try to deploy the model into production as an Azure Kubernetes Service (AKS)-based web service, you experience an error in the scoring script that causes deployment to fail.
You need to debug the service on the local workstation before deploying the service to production.
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.




A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Graphical user interface, text, application, email Description automatically generated
Step 1: Install Docker on the workstation
Prerequisites include having a working Docker installation on your local system. Build or download the dockerfile to the compute node.
Step 2: Create an AksWebservice deployment configuration and deploy the model to it
To deploy a model to Azure Kubernetes Service, create a deployment configuration that describes the compute resources needed.
# If deploying to a cluster configured for dev/test, ensure that it was created with enough
# cores and memory to handle this deployment configuration. Note that memory is also used by
# things such as dependencies and AML components.
deployment_config = AksWebservice.deploy_configuration(cpu_cores = 1, memory_gb = 1)
service = Model.deploy(ws, "myservice", [model], inference_config, deployment_config, aks_target) service.wait_for_deployment(show_output = True)
print(service.state) print(service.get_logs())
Step 3: Create a LocalWebservice deployment configuration for the service and deploy the model to it
To deploy locally, modify your code to use LocalWebservice.deploy_configuration() to create a deployment configuration. Then use Model.deploy() to deploy the service.
Step 4: Debug and modify the scoring script as necessary. Use the reload() method of the service after each modification.
During local testing, you may need to update the score.py file to add logging or attempt to resolve any problems that you've discovered. To reload changes to the score.py file, use reload(). For example, the following code reloads the script for the service, and then sends data to it.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-deploy-azure-kubernetes-service https://docs.microsoft.com/en-us/azure/machine-learning/how-to-troubleshoot-deployment-local

**NEW QUESTION 129**
- (Exam Topic 3)
You are building a recurrent neural network to perform a binary classification. You review the training loss, validation loss, training accuracy, and validation accuracy for each training epoch.
You need to analyze model performance.
Which observation indicates that the classification model is over fitted?

A. The training loss .stays constant and the validation loss stays on a constant value and close to the training loss value when training the model.
B. The training loss increases while the validation loss decreases when training the model.
C. The training loss decreases while the validation loss increases when training the model.
D. The training loss stays constant and the validation loss decreases when training the model.

**Answer:** B

**NEW QUESTION 131**
- (Exam Topic 3)
You deploy a model as an Azure Machine Learning real-time web service using the following code.

```
# ws, model, interference_config, and deployment_config defined previously
service = Model.deploy(ws, 'classification-service', [model], inference_config, deployment_config)
service.wait_for_deployment(True)
```

The deployment fails.
You need to troubleshoot the deployment failure by determining the actions that were performed during deployment and identifying the specific action that failed.
Which code segment should you run?

A. service.get_logs()
B. service.state
C. service.serialize()
D. service.update_deployment_state()

**Answer:** A

**Explanation:**
You can print out detailed Docker engine log messages from the service object. You can view the log for ACI, AKS, and Local deployments. The following example demonstrates how to print the logs.
# if you already have the service object handy print(service.get_logs())
# if you only know the name of the service (note there might be multiple services with the same name but different version number)
print(ws.webservices['mysvc'].get_logs()) Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-troubleshoot-deployment

**NEW QUESTION 134**
- (Exam Topic 3)
You plan to provision an Azure Machine Learning Basic edition workspace for a data science project. You need to identify the tasks you will be able to perform in the workspace.
Which three tasks will you be able to perform? Each correct answer presents a complete solution.
NOTE: Each correct selection is worth one point. D

A. Create a Compute Instance and use it to run code in Jupyter notebooks.
B. Create an Azure Kubernetes Service (AKS) inference cluster.
C. Use the designer to train a model by dragging and dropping pre-defined modules.
D. Create a tabular dataset that supports versioning.
E. Use the Automated Machine Learning user interface to train a model.

**Answer:** ABD

**Explanation:**
Reference:
https://azure.microsoft.com/en-us/pricing/details/machine-learning/

**NEW QUESTION 139**
- (Exam Topic 3)
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 using Azure Machine Learning to run an experiment that trains a classification model.
You want to use Hyperdrive to find parameters that optimize the AUC metric for the model. You configure a HyperDriveConfig for the experiment by running the following code:

```
hyperdrive = HyperDriveConfig(estimator=your_estimator,
 hyperparameter_sampling=your_params,
 policy=policy,
 primary_metric_name='AUC',
 primary_metric_goal=PrimaryMetricGoal.MAXIMIZE,
 max_total_runs=6,
 max_concurrent_runs=4)
```

You plan to use this configuration to run a script that trains a random forest model and then tests it with validation data. The label values for the validation data are stored in a variable named y_test variable, and the predicted probabilities from the model are stored in a variable named y_predicted.
You need to add logging to the script to allow Hyperdrive to optimize hyperparameters for the AUC metric. Solution: Run the following code:

```
import json, os
from sklearn.metrics import roc_auc_score
# code to train model omitted
auc = roc_auc_score(y_test, y_predicted)
os.makedirs("outputs", exist_ok = True)
with open("outputs/AUC.txt", "w") as file_cur:
   file_cur.write(auc)
```

Does the solution meet the goal?

A. Yes
B. No

**Answer:** B

**Explanation:**
Use a solution with logging.info(message) instead. Note: Python printing/logging example: logging.info(message)
Destination: Driver logs, Azure Machine Learning designer Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-debug-pipelines

**NEW QUESTION 143**
- (Exam Topic 3)
You have a Python data frame named salesData in the following format:

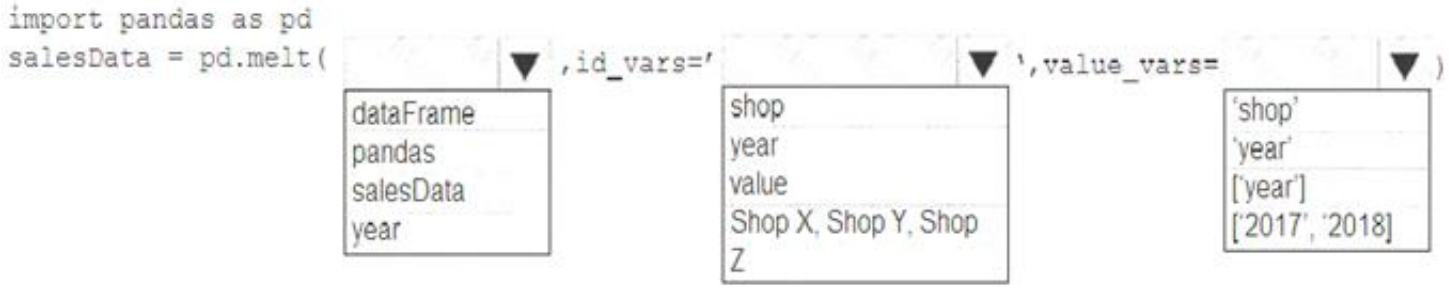
| | shop | 2017 | 2018 |
|---|---|---|---|
| 0 | Shop X | 34 | 25 |
| 1 | Shop Y | 65 | 76 |
| 2 | Shop Z | 48 | 55 |

The data frame must be unpivoted to a long data format as follows:

| | shop | year | value |
|---|---|---|---|
| 0 | Shop X | 2017 | 34 |
| 1 | Shop Y | 2017 | 65 |
| 2 | Shop Z | 2017 | 48 |
| 3 | Shop X | 2018 | 25 |
| 4 | Shop Y | 2018 | 76 |
| 5 | Shop Z | 2018 | 55 |

You need to use the pandas.melt() function in Python to perform the transformation.
How should you complete the code segment? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.

**Answer Area**

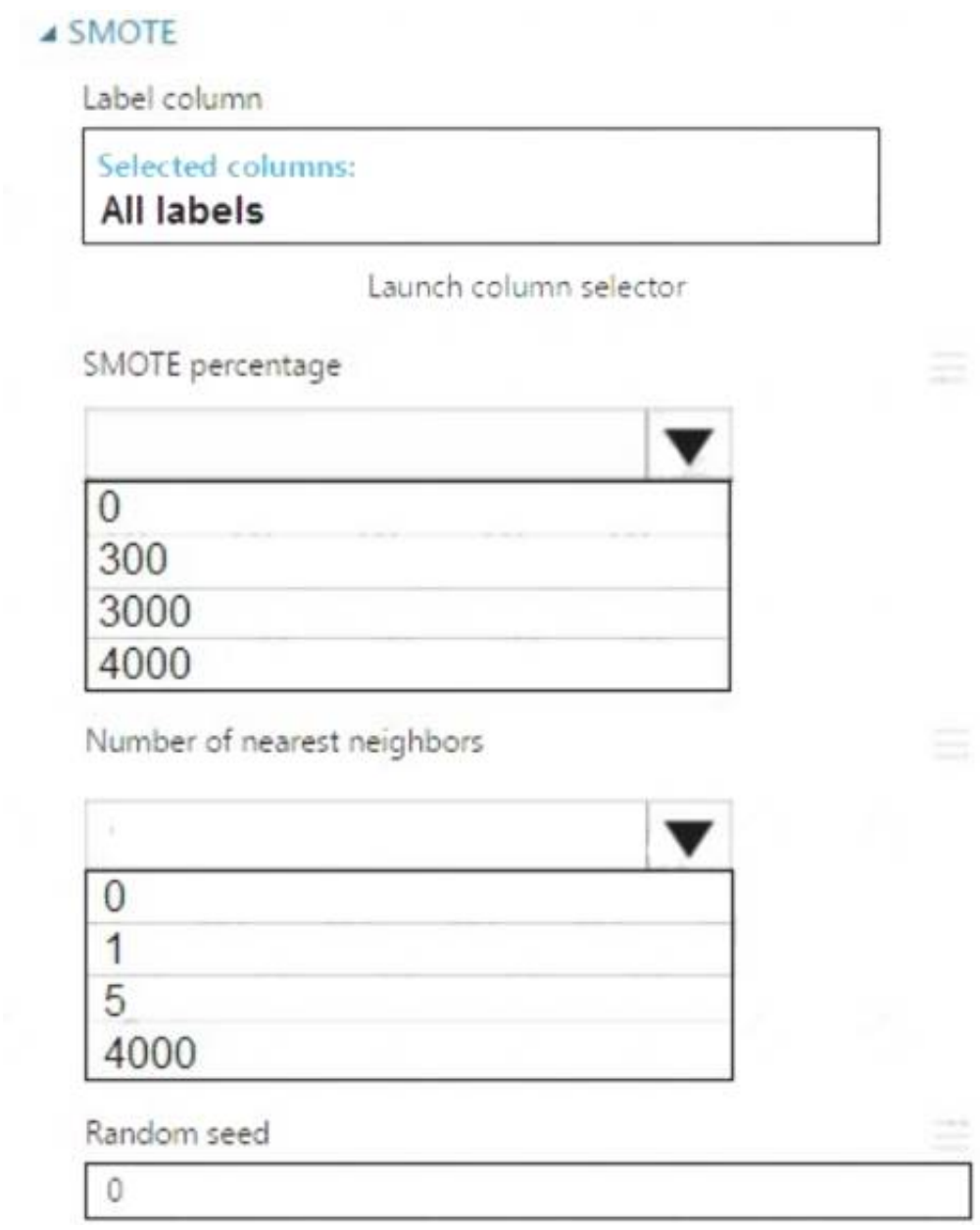



A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Box 1: dataFrame
Syntax: pandas.melt(frame, id_vars=None, value_vars=None, var_name=None, value_name='value', col_level=None)[source]
Where frame is a DataFrame Box 2: shop
Paramter id_vars id_vars : tuple, list, or ndarray, optional Column(s) to use as identifier variables.
Box 3: ['2017','2018']
value_vars : tuple, list, or ndarray, optional
Column(s) to unpivot. If not specified, uses all columns that are not set as id_vars. Example:
df = pd.DataFrame({'A': {0: 'a', 1: 'b', 2: 'c'},
'B': {0: 1, 1: 3, 2: 5},
'C': {0: 2, 1: 4, 2: 6}})
pd.melt(df, id_vars=['A'], value_vars=['B', 'C']) A variable value
0 a B 1
1 b B 3
2 c B 5
3 a C 2
4 b C 4
5 c C 6
References:
https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.melt.html

**NEW QUESTION 148**
- (Exam Topic 3)
You create an experiment in Azure Machine Learning Studio. You add a training dataset that contains 10,000 rows. The first 9,000 rows represent class 0 (90 percent).
The remaining 1,000 rows represent class 1 (10 percent).
The training set is imbalances between two classes. You must increase the number of training examples for class 1 to 4,000 by using 5 data rows. You add the Synthetic Minority Oversampling Technique (SMOTE) module to the experiment.
You need to configure the module.
Which values should you use? To answer, select the appropriate options in the dialog box in the answer area. NOTE: Each correct selection is worth one point.

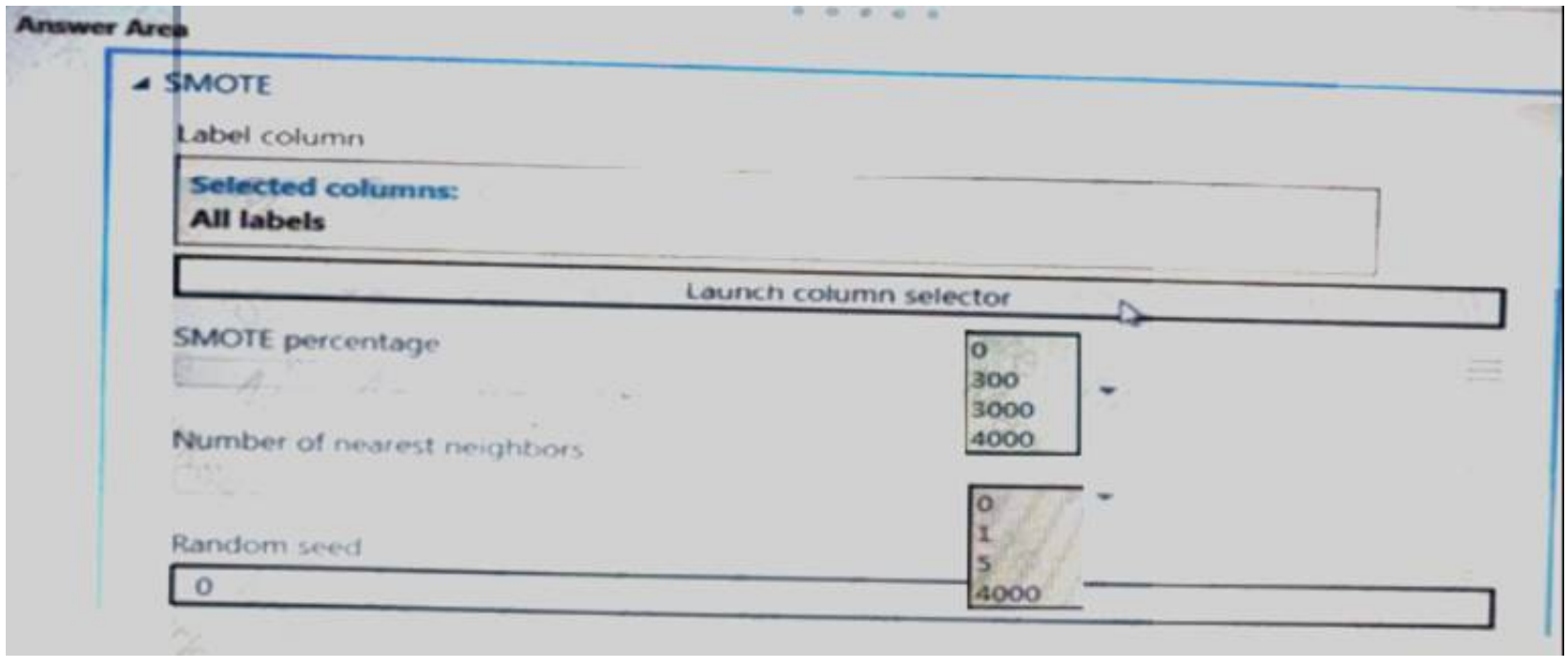



A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Box 1: 300
You type 300 (%), the module triples the percentage of minority cases (3000) compared to the original dataset (1000).
Box 2: 5
We should use 5 data rows.
Use the Number of nearest neighbors option to determine the size of the feature space that the SMOTE algorithm uses when in building new cases. A nearest neighbor is a row of data (a case) that is very similar to some target case. The distance between any two cases is measured by combining the weighted vectors of all features.
By increasing the number of nearest neighbors, you get features from more cases.
By keeping the number of nearest neighbors low, you use features that are more like those in the original sample.
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/smote

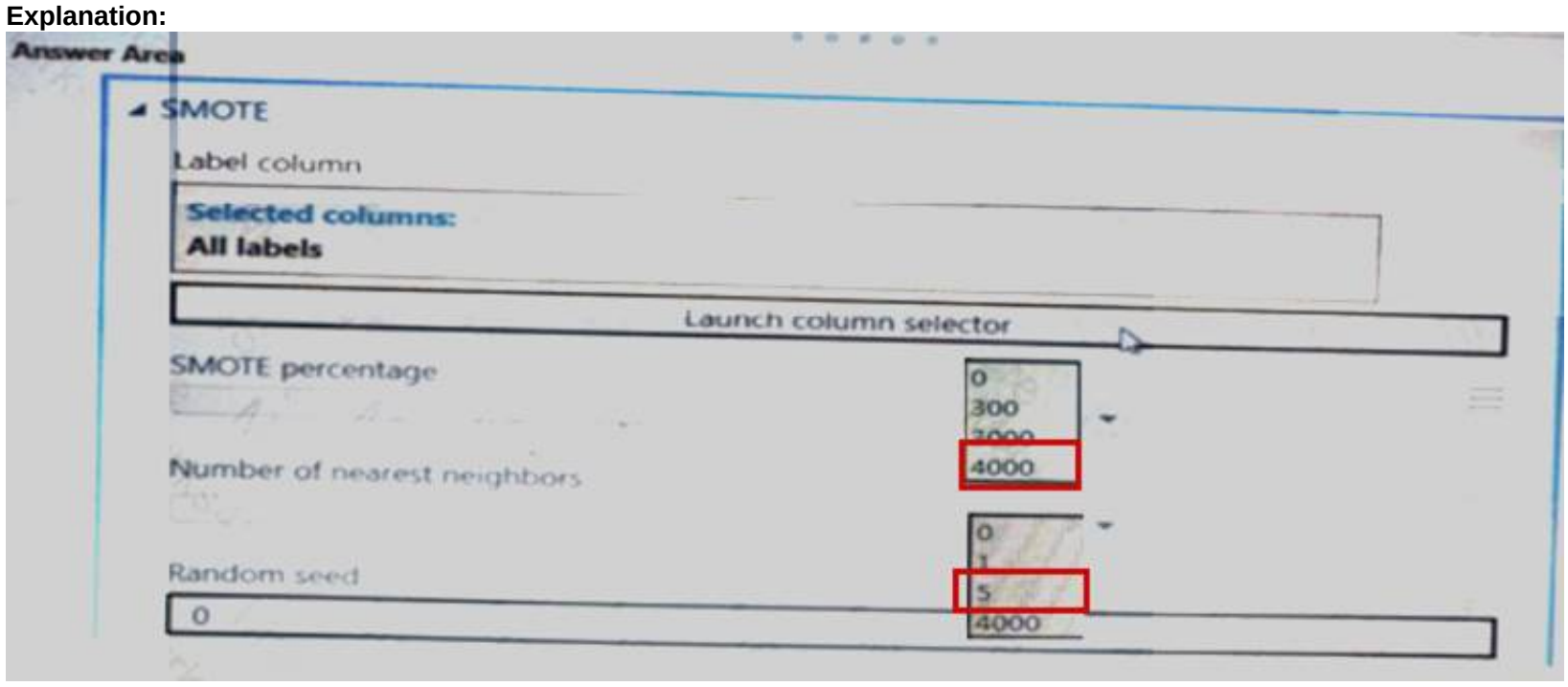
**NEW QUESTION 149**
- (Exam Topic 3)
You create an experiment in Azure Machine Learning Studio- You add a training dataset that contains 10.000 rows. The first 9.000 rows represent class 0 (90 percent). The first 1.000 rows represent class 1 (10 percent).
The training set is unbalanced between two Classes. You must increase the number of training examples for class 1 to 4,000 by using data rows. You add the Synthetic Minority Oversampling Technique (SMOTE) module to the experiment.
You need to configure the module.
Which values should you use? To answer, select the appropriate options in the dialog box in the answer area. NOTE: Each correct selection is worth one point.



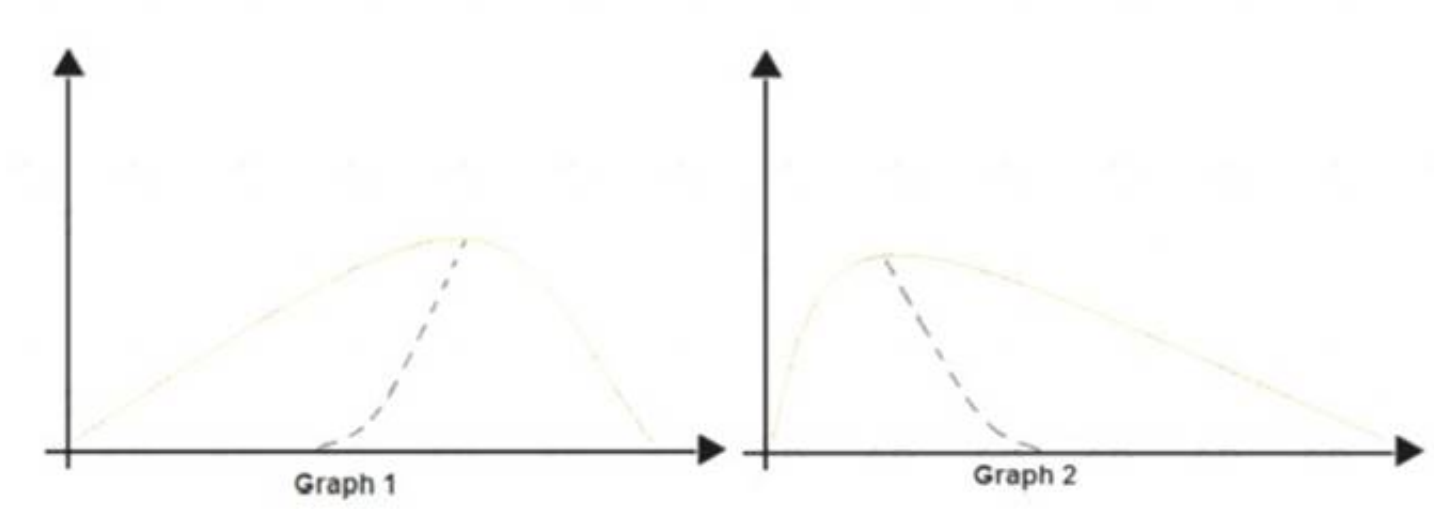

A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**




**NEW QUESTION 151**
- (Exam Topic 3)
You are analyzing the asymmetry in a statistical distribution.
The following image contains two density curves that show the probability distribution of two datasets.




Use the drop-down menus to select the answer choice that answers each question based on the information presented in the graphic.
NOTE: Each correct selection is worth one point.

| Question | Answer choice |
|---|---|
| Which type of distribution is shown for the dataset density curve of Graph 1? | Negative skew<br>Positive skew<br>Normal distribution<br>Bimodal distribution |
| Which type of distribution is shown for the dataset density curve of Graph 2? | Negative skew<br>Positive skew<br>Normal distribution<br>Bimodal distribution |

A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Box 1: Positive skew
Positive skew values means the distribution is skewed to the right. Box 2: Negative skew
Negative skewness values mean the distribution is skewed to the left. References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/compute-elementary-statistic

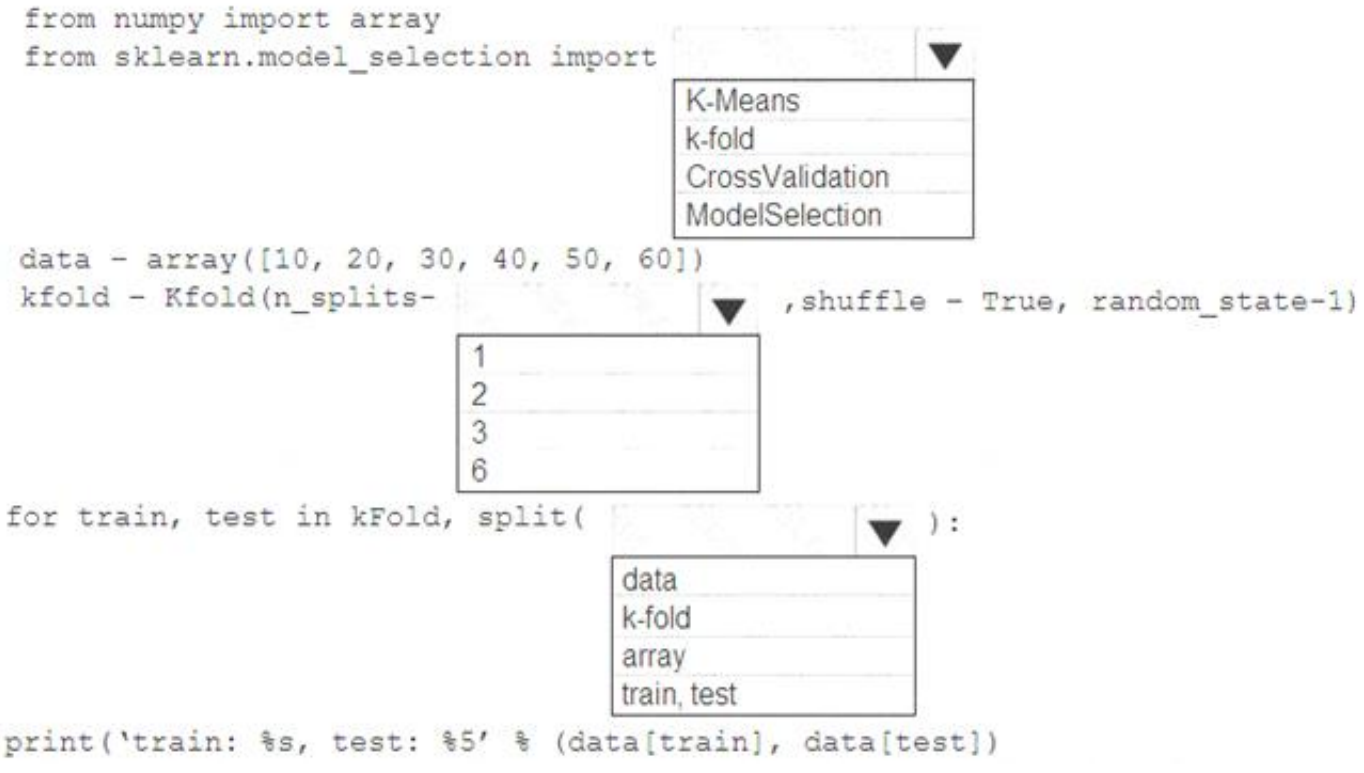
**NEW QUESTION 154**
- (Exam Topic 3)
You develop and train a machine learning model to predict fraudulent transactions for a hotel booking website. Traffic to the site varies considerably. The site experiences heavy traffic on Monday and Friday and much lower traffic on other days. Holidays are also high web traffic days. You need to deploy the model as an Azure Machine Learning real-time web service endpoint on compute that can dynamically scale up and down to support demand. Which deployment compute option should you use?

A. attached Azure Databricks cluster
B. Azure Container Instance (ACI)
C. Azure Kubernetes Service (AKS) inference cluster
D. Azure Machine Learning Compute Instance
E. attached virtual machine in a different region

**Answer:** D

**Explanation:**
Azure Machine Learning compute cluster is a managed-compute infrastructure that allows you to easily create a single or multi-node compute. The compute is created within your workspace region as a resource that can be shared with other users in your workspace. The compute scales up automatically when a job is submitted, and can be put in an Azure Virtual Network.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-create-attach-compute-sdk

**NEW QUESTION 157**
- (Exam Topic 3)
You are evaluating a Python NumPy array that contains six data points defined as follows: data = [10, 20, 30, 40, 50, 60]
You must generate the following output by using the k-fold algorithm implantation in the Python Scikit-learn machine learning library:
train: [10 40 50 60], test: [20 30]
train: [20 30 40 60], test: [10 50]
train: [10 20 30 50], test: [40 60]
You need to implement a cross-validation to generate the output.
How should you complete the code segment? To answer, select the appropriate code segment in the dialog box in the answer area.
NOTE: Each correct selection is worth one point.




A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Box 1: k-fold
Box 2: 3
K-F olds cross-validator provides train/test indices to split data in train/test sets. Split dataset into k consecutive folds (without shuffling by default).
The parameter n_splits ( int, default=3) is the number of folds. Must be at least 2. Box 3: data
Example: Example:
>>>
>>> from sklearn.model_selection import KFold
>>> X = np.array([[1, 2], [3, 4], [1, 2], [3, 4]])
>>> y = np.array([1, 2, 3, 4])
>>> kf = KFold(n_splits=2)
>>> kf.get_n_splits(X) 2
>>> print(kf)
KFold(n_splits=2, random_state=None, shuffle=False)
>>> for train_index, test_index in kf.split(X): print("TRAIN:", train_index, "TEST:", test_index) X_train, X_test = X[train_index], X[test_index] y_train, y_test = y[train_index], y[test_index] TRAIN: [2 3] TEST: [0 1]
TRAIN: [0 1] TEST: [2 3]
References:
https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.KFold.html

**NEW QUESTION 159**
- (Exam Topic 3)
Your team is building a data engineering and data science development environment. The environment must support the following requirements:
- support Python and Scala
- compose data storage, movement, and processing services into automated data pipelines
- the same tool should be used for the orchestration of both data engineering and data science
- support workload isolation and interactive workloads
- enable scaling across a cluster of machines You need to create the environment.

What should you do?

A. Build the environment in Apache Hive for HDInsight and use Azure Data Factory for orchestration.
B. Build the environment in Azure Databricks and use Azure Data Factory for orchestration.
C. Build the environment in Apache Spark for HDInsight and use Azure Container Instances for orchestration.
D. Build the environment in Azure Databricks and use Azure Container Instances for orchestration.

**Answer:** B

**Explanation:**
In Azure Databricks, we can create two different types of clusters.
- Standard, these are the default clusters and can be used with Python, R, Scala and SQL
- High-concurrency

Azure Databricks is fully integrated with Azure Data Factory.

**NEW QUESTION 164**
- (Exam Topic 3)
You plan to use the Hyperdrive feature of Azure Machine Learning to determine the optimal hyperparameter values when training a model.
You must use Hyperdrive to try combinations of the following hyperparameter values:
• learning_rate: any value between 0.001 and 0.1
• batch_size: 16, 32, or 64

You need to configure the search space for the Hyperdrive experiment.
Which two parameter expressions should you use? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

A. a choice expression for learning_rate
B. a uniform expression for learning_rate
C. a normal expression for batch_size
D. a choice expression for batch_size
E. a uniform expression for batch_size

**Answer:** BD

**Explanation:**
B: Continuous hyperparameters are specified as a distribution over a continuous range of values. Supported distributions include:
- uniform(low, high) - Returns a value uniformly distributed between low and high

D: Discrete hyperparameters are specified as a choice among discrete values. choice can be:
- one or more comma-separated values
- a range object
- any arbitrary list object Reference:

https://docs.microsoft.com/en-us/azure/machine-learning/how-to-tune-hyperparameters

**NEW QUESTION 169**
- (Exam Topic 3)
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 plan to use a Python script to run an Azure Machine Learning experiment. The script creates a reference to the experiment run context, loads data from a file, identifies the set of unique values for the label column, and completes the experiment run:
from azureml.core import Run import pandas as pd
run = Run.get_context()
data = pd.read_csv('data.csv') label_vals = data['label'].unique()
# Add code to record metrics here run.complete()
The experiment must record the unique labels in the data as metrics for the run that can be reviewed later.
You must add code to the script to record the unique label values as run metrics at the point indicated by the comment.
Solution: Replace the comment with the following code:
run.upload_file('outputs/labels.csv', './data.csv') Does the solution meet the goal?

A. Yes
B. No

**Answer:** B

**Explanation:**
label_vals has the unique labels (from the statement label_vals = data['label'].unique()), and it has to be logged.
Note:
Instead use the run_log function to log the contents in label_vals: for label_val in label_vals:
run.log('Label Values', label_val) Reference:
https://www.element61.be/en/resource/azure-machine-learning-services-complete-toolbox-ai

**NEW QUESTION 172**
- (Exam Topic 3)
You are creating a new experiment in Azure Machine Learning Studio. You have a small dataset that has missing values in many columns. The data does not require the application of predictors for each column. You plan to use the Clean Missing Data module to handle the missing data.
You need to select a data cleaning method. Which method should you use?

A. Synthetic Minority Oversampling Technique (SMOTE)
B. Replace using MICE
C. Replace using; Probabilistic PCA
D. Normalization

**Answer:** C

**Explanation:**
Replace using Probabilistic PCA: Compared to other options, such as Multiple Imputation using Chained Equations (MICE), this option has the advantage of not requiring the application of predictors for each column. Instead, it approximates the covariance for the full dataset. Therefore, it might offer better performance for datasets that have missing values in many columns.
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/clean-missing-data

**NEW QUESTION 177**
- (Exam Topic 3)
An organization uses Azure Machine Learning service and wants to expand their use of machine learning. You have the following compute environments. The organization does not want to create another compute
environment.

| Environment name | Compute type |
|---|---|
| nb_server | Compute Instance |
| aks_cluster | Azure Kubernetes Service |
| mlc_cluster | Machine Learning Compute |

You need to determine which compute environment to use for the following scenarios.
Which compute types should you use? To answer, drag the appropriate compute environments to the correct scenarios. Each compute environment may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.
NOTE: Each correct selection is worth one point.

**Environments**

nb_server
aks_cluster
mlc_cluster

**Answer Area**

| Scenario | Environment |
|---|---|
| Run an Azure Machine Learning Designer training pipeline. | Environment |
| Deploying a web service from the Azure Machine Learning designer. | Environment |

A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Box 1: nb_server

| Training targets | Automated ML | ML pipelines | Azure Machine Learning designer |
|---|---|---|---|
| Local computer | yes | | |
| Azure Machine Learning compute cluster | yes & hyperparameter tuning | yes | yes |
| Azure Machine Learning compute instance | yes & hyperparameter tuning | yes | yes |
| Remote VM | yes & hyperparameter tuning | yes | |
| Azure Databricks | yes (SDK local mode only) | yes | |
| Azure Data Lake Analytics | | yes | |
| Azure HDInsight | | yes | |
| Azure Batch | | yes | |

Box 2: mlc_cluster
With Azure Machine Learning, you can train your model on a variety of resources or environments, collectively referred to as compute targets. A compute target can be a local machine or a cloud resource, such as an Azure Machine Learning Compute, Azure HDInsight or a remote virtual machine.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/concept-compute-target https://docs.microsoft.com/en-us/azure/machine-learning/how-to-set-up-training-targets

**NEW QUESTION 178**
- (Exam Topic 3)
You plan to use the Hyperdrive feature of Azure Machine Learning to determine the optimal hyperparameter values when training a model.
You must use Hyperdrive to try combinations of the following hyperparameter values. You must not apply an early termination policy.
learning_rate: any value between 0.001 and 0.1
• batch_size: 16, 32, or 64
You need to configure the sampling method for the Hyperdrive experiment
Which two sampling methods can you use? Each correct answer is a complete solution. NOTE: Each correct selection is worth one point.

A. Grid sampling
B. No sampling
C. Bayesian sampling
D. Random sampling

**Answer:** CD

**Explanation:**
C: Bayesian sampling is based on the Bayesian optimization algorithm and makes intelligent choices on the hyperparameter values to sample next. It picks the sample based on how the previous samples performed, such that the new sample improves the reported primary metric.
Bayesian sampling does not support any early termination policy Example:
from azureml.train.hyperdrive import BayesianParameterSampling from azureml.train.hyperdrive import uniform, choice param_sampling = BayesianParameterSampling( { "learning_rate": uniform(0.05, 0.1),
"batch_size": choice(16, 32, 64, 128)
}
)
D: In random sampling, hyperparameter values are randomly selected from the defined search space. Random sampling allows the search space to include both discrete and continuous hyperparameters.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-tune-hyperparameters

**NEW QUESTION 183**
- (Exam Topic 3)
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 creating a model to predict the price of a student's artwork depending on the following variables: the student's length of education, degree type, and art form.
You start by creating a linear regression model. You need to evaluate the linear regression model.
Solution: Use the following metrics: Accuracy, Precision, Recall, F1 score and AUC. Does the solution meet the goal?

A. Yes
B. No

**Answer:** B

**Explanation:**
Those are metrics for evaluating classification models, instead use: Mean Absolute Error, Root Mean Absolute Error, Relative Absolute Error, Relative Squared Error, and the Coefficient of Determination.
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/evaluate-model

**NEW QUESTION 184**
- (Exam Topic 3)
You are creating a machine learning model in Python. The provided dataset contains several numerical columns and one text column. The text column represents a product's category. The product category will always be one of the following:
- Bikes
- Cars
- Vans
- Boats

You are building a regression model using the scikit-learn Python package.
You need to transform the text data to be compatible with the scikit-learn Python package.
How should you complete the code segment? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.

```
from sklearn import linear_model
import                                  ▼
```

| |
|---|
| pandas as df |
| numpy as df |
| scipy as df |

```
dataset = df.read_csv("data\\ProductSales.csv")
ProductCategoryMapping = {"Bikes":1, "Cars":2, "Boats": 3,
"Vans": 4}
dataset['ProductCategoryMapping'] =
dataset['ProductCategory'].                    ▼
```

| |
|---|
| map[ProductCategoryMapping] |
| reduce[ProductCategoryMapping] |
| transpose[ProductCategoryMapping] |

```
regr = linear_model.LinearRegression()
X_train = dataset[['ProductCategoryMapping', 'ProductSize',
'ProductCost']]
y_train = dataset[['Sales']]
regr.fit(X_train, y_train)
```

A. Mastered
B. Not Mastered

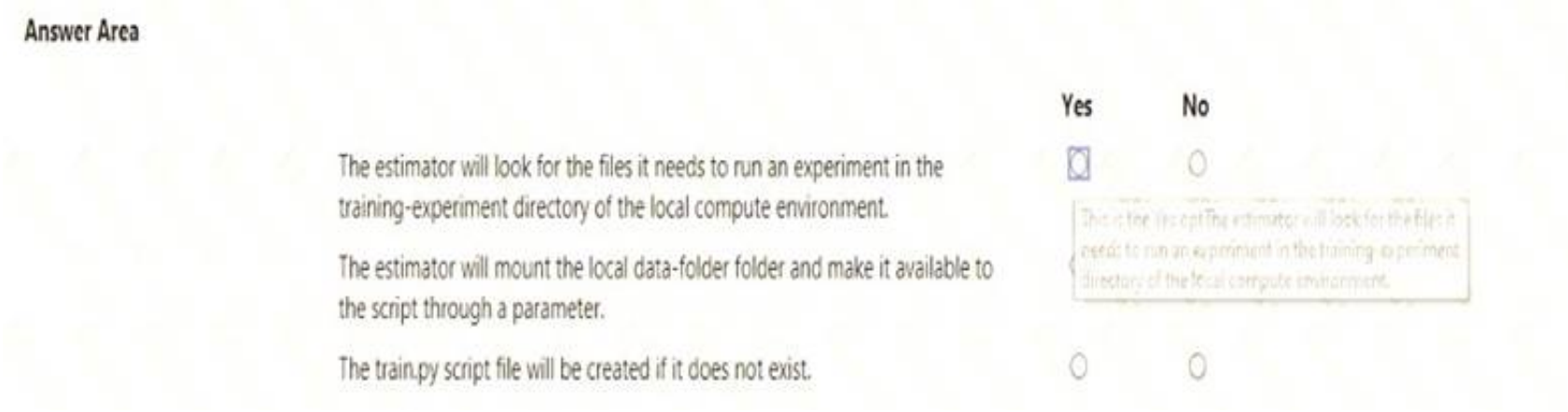
**Answer:** A

**Explanation:**
Box 1: pandas as df
Pandas takes data (like a CSV or TSV file, or a SQL database) and creates a Python object with rows and columns called data frame that looks very similar to table in a statistical software (think Excel or SPSS for example.
Box 2: transpose[ProductCategoryMapping] Reshape the data from the pandas Series to columns. Reference:
https://datascienceplus.com/linear-regression-in-python/

**NEW QUESTION 189**
- (Exam Topic 3)

```
train_cluster = ComputeTarget(workspace=work_space, name='train-cluster')
estimator = Estimator(source_directory =
 'training-experiment',
 script_params = {'--data-folder': data_source.as_mount(), '--regularization': 0.8},
 compute_target = train_cluster,
 entry_script = 'train.py',
 conda_packages = ['scikit-learn'])
```

For each of the following statements, select Yes if the statement is true. Otherwise, select No. NOTE: Each correct selection is worth one point.




A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**


**NEW QUESTION 194**
- (Exam Topic 3)
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 a data scientist using Azure Machine Learning Studio.
You need to normalize values to produce an output column into bins to predict a target column. Solution: Apply an Equal Width with Custom Start and Stop binning mode.
Does the solution meet the goal?

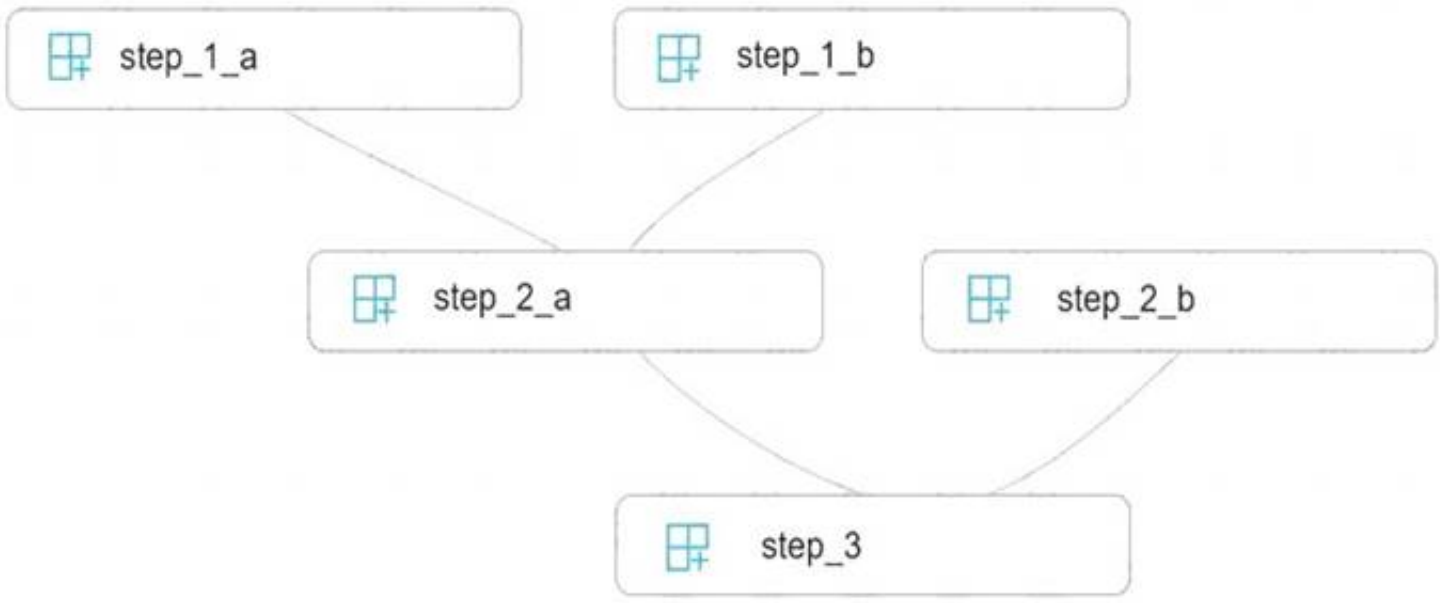
A. Yes
B. No

**Answer:** B

**Explanation:**
Use the Entropy MDL binning mode which has a target column. References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/group-data-into-bins

**NEW QUESTION 199**
- (Exam Topic 3)
You write five Python scripts that must be processed in the order specified in Exhibit A – which allows the same modules to run in parallel, but will wait for modules

with dependencies.
You must create an Azure Machine Learning pipeline using the Python SDK, because you want to script to create the pipeline to be tracked in your version control system. You have created five PythonScriptSteps and have named the variables to match the module names.




You need to create the pipeline shown. Assume all relevant imports have been done. Which Python code segment should you use?

A.
```
p = Pipeline(ws, steps=[[[[step_1_a, step_1_b], step_2_a],  step_2_b], step_3])
```

B.
```
pipeline_steps = {
    "Pipeline": {
        "run": step_3,
        "run_after": {[
            {"run": step_2_a,
                "run_after":
                    [{"run": step_1_a},
                     {"run": step_1_b}]
                },
            {"run": step_2_b}]
        }
    }
}
p = Pipeline(ws, steps=pipeline_steps)
```

C.
```
step_2_a.run_after(step_1_b)
step_2_a.run_after(step_1_a)
step_3.run_after(step_2_b)
step_3.run_after(step_2_a)
p = Pipeline(ws, steps=[step_3])
```

D.
```
p = Pipeline(ws, steps=[step_1_a, step_1_b, step_2_a, step_2_b, step_3])
```

A. Option A
B. Option B
C. Option C
D. Option D

**Answer:** A

**Explanation:**
The steps parameter is an array of steps. To build pipelines that have multiple steps, place the steps in order in this array.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-use-parallel-run-step

**NEW QUESTION 204**
- (Exam Topic 3)
You use Azure Machine Learning Studio to build a machine learning experiment. You need to divide data into two distinct datasets.
Which module should you use?

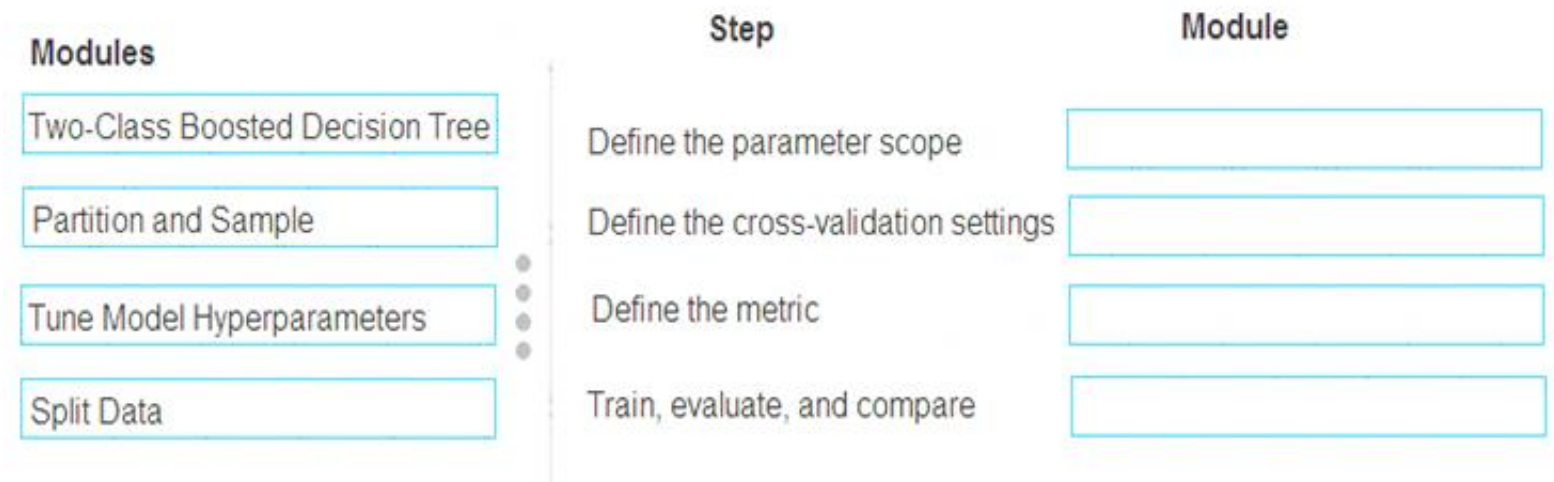
A. Partition and Sample
B. Assign Data to Clusters
C. Group Data into Bins
D. Test Hypothesis Using t-Test

**Answer:** A

**Explanation:**
Partition and Sample with the Stratified split option outputs multiple datasets, partitioned using the rules you specified.
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/partition-and-sample

**NEW QUESTION 205**
- (Exam Topic 3)
You have a model with a large difference between the training and validation error values. You must create a new model and perform cross-validation.
You need to identify a parameter set for the new model using Azure Machine Learning Studio.
Which module you should use for each step? To answer, drag the appropriate modules to the correct steps. Each module may be used once or more than once, or

not at all. You may need to drag the split bar between panes or scroll to view content.
NOTE: Each correct selection is worth one point.

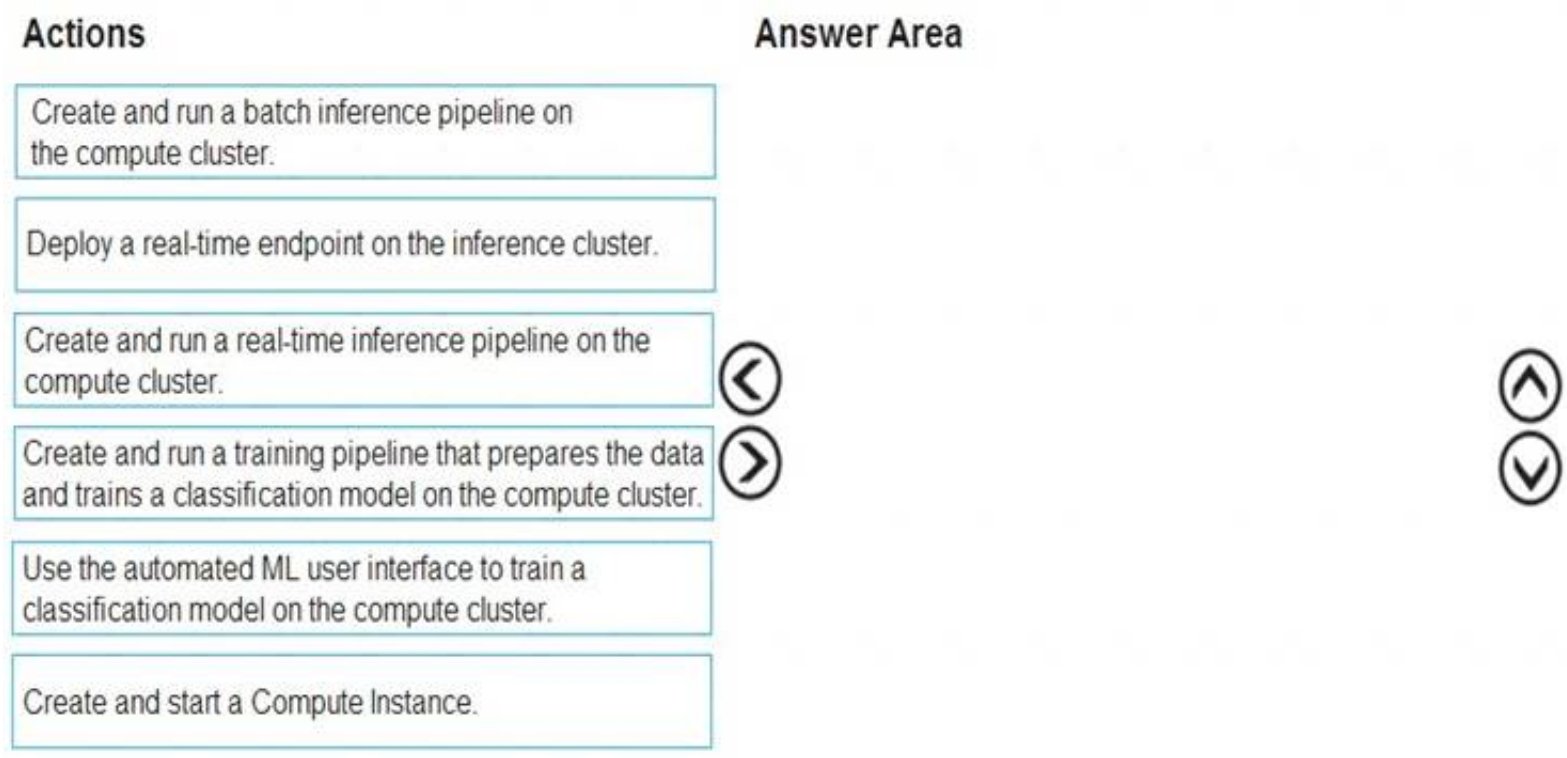



A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Box 1: Split data
Box 2: Partition and Sample
Box 3: Two-Class Boosted Decision Tree Box 4: Tune Model Hyperparameters
Integrated train and tune: You configure a set of parameters to use, and then let the module iterate over multiple combinations, measuring accuracy until it finds a "best" model. With most learner modules, you can choose which parameters should be changed during the training process, and which should remain fixed.
We recommend that you use Cross-Validate Model to establish the goodness of the model given the specified
parameters. Use Tune Model Hyperparameters to identify the optimal parameters. References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/partition-and-sample

**NEW QUESTION 208**
- (Exam Topic 3)
You have an Azure Machine Learning workspace that contains a CPU-based compute cluster and an Azure Kubernetes Services (AKS) inference cluster. You create a tabular dataset containing data that you plan to use to create a classification model.
You need to use the Azure Machine Learning designer to create a web service through which client applications can consume the classification model by submitting new data and getting an immediate prediction as a response.
Which three 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.




A. Mastered
B. Not Mastered

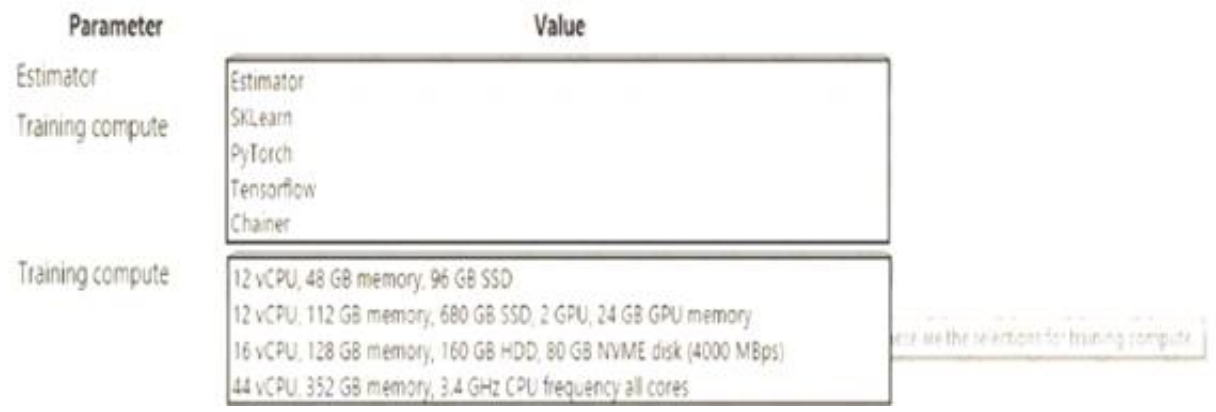
**Answer:** A

**Explanation:**
Step 1: Create and start a Compute Instance
To train and deploy models using Azure Machine Learning designer, you need compute on which to run the training process, test the model, and host the model in a deployed service.
There are four kinds of compute resource you can create:
Compute Instances: Development workstations that data scientists can use to work with data and models. Compute Clusters: Scalable clusters of virtual machines for on-demand processing of experiment code. Inference Clusters: Deployment targets for predictive services that use your trained models.
Attached Compute: Links to existing Azure compute resources, such as Virtual Machines or Azure Databricks clusters.
Step 2: Create and run a training pipeline..
After you've used data transformations to prepare the data, you can use it to train a machine learning model. Create and run a training pipeline
Step 3: Create and run a real-time inference pipeline
After creating and running a pipeline to train the model, you need a second pipeline that performs the same data transformations for new data, and then uses the

trained model to inference (in other words, predict) label values based on its features. This pipeline will form the basis for a predictive service that you can publish for applications to use.
Reference:
https://docs.microsoft.com/en-us/learn/modules/create-classification-model-azure-machine-learning-designer/

**NEW QUESTION 213**
- (Exam Topic 3)
You create an Azure Machine Learning workspace and set up a development environment. You plan to train a deep neural network (DNN) by using the Tensorflow framework and by using estimators to submit training scripts.
You must optimize computation speed for training runs.
You need to choose the appropriate estimator to use as well as the appropriate training compute target configuration.
Which values should you use? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.

**Answer Area**




A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Box 1: Tensorflow
TensorFlow represents an estimator for training in TensorFlow experiments. Box 2: 12 vCPU, 112 GB memory..,2 GPU,..
Use GPUs for the deep neural network. Reference:
https://docs.microsoft.com/en-us/python/api/azureml-train-core/azureml.train.dnn

**NEW QUESTION 216**
- (Exam Topic 3)
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 using Azure Machine Learning Studio to perform feature engineering on a dataset. You need to normalize values to produce a feature column grouped into bins.
Solution: Apply an Entropy Minimum Description Length (MDL) binning mode. Does the solution meet the goal?

A. Yes
B. No

**Answer:** A

**Explanation:**
Entropy MDL binning mode: This method requires that you select the column you want to predict and the column or columns that you want to group into bins. It then makes a pass over the data and attempts to determine the number of bins that minimizes the entropy. In other words, it chooses a number of bins that allows the data column to best predict the target column. It then returns the bin number associated with each row of your data in a column named <colname>quantized.
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/group-data-into-bins

**NEW QUESTION 220**
- (Exam Topic 3)
You use Azure Machine Learning designer to create a real-time service endpoint. You have a single Azure Machine Learning service compute resource. You train the model and prepare the real-time pipeline for deployment You need to publish the inference pipeline as a web service. Which compute type should you use?

A. HDInsight
B. Azure Databricks
C. Azure Kubernetes Services
D. the existing Machine Learning Compute resource
E. a new Machine Learning Compute resource

**Answer:** C

**Explanation:**
Azure Kubernetes Service (AKS) can be used real-time inference. Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/concept-compute-target

**NEW QUESTION 225**
- (Exam Topic 3)

You have a comma-separated values (CSV) file containing data from which you want to train a classification model.
You are using the Automated Machine Learning interface in Azure Machine Learning studio to train the classification model. You set the task type to Classification.
You need to ensure that the Automated Machine Learning process evaluates only linear models. What should you do?

A. Add all algorithms other than linear ones to the blocked algorithms list.
B. Set the Exit criterion option to a metric score threshold.
C. Clear the option to perform automatic featurization.
D. Clear the option to enable deep learning.
E. Set the task type to Regression.

**Answer:** C

**Explanation:**
Automatic featurization can fit non-linear models. Reference: https://econml.azurewebsites.net/spec/estimation/dml.html
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-use-automated-ml-for-ml-models

**NEW QUESTION 226**
- (Exam Topic 3)
You are building a regression model tot estimating the number of calls during an event.
You need to determine whether the feature values achieve the conditions to build a Poisson regression model. Which two conditions must the feature set contain?
I ach correct answer presents part of the solution. NOTE:
Each correct selection is worth one point.

A. The label data must be a negative value.
B. The label data can be positive or negative,
C. The label data must be a positive value
D. The label data must be non discrete.
E. The data must be whole numbers.

**Answer:** CE

**Explanation:**
Poisson regression is intended for use in regression models that are used to predict numeric values, typically counts. Therefore, you should use this module to create your regression model only if the values you are trying to predict fit the following conditions:

- The response variable has a Poisson distribution.
- Counts cannot be negative. The method will fail outright if you attempt to use it with negative labels.
- A Poisson distribution is a discrete distribution; therefore, it is not meaningful to use this method with non-whole numbers.

References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/poisson-regression

**NEW QUESTION 231**
- (Exam Topic 3)
You are performing a filter based feature selection for a dataset 10 build a multi class classifies by using Azure Machine Learning Studio.
The dataset contains categorical features that are highly correlated to the output label column.
You need to select the appropriate feature scoring statistical method to identify the key predictors. Which method should you use?

A. Chi-squared
B. Spearman correlation
C. Kendall correlation
D. Person correlation

**Answer:** D

**Explanation:**
Pearson’s correlation statistic, or Pearson’s correlation coefficient, is also known in statistical models as the r value. For any two variables, it returns a value that indicates the strength of the correlation
Pearson’s correlation coefficient is the test statistics that measures the statistical relationship, or association, between two continuous variables. It is known as the best method of measuring the association between variables of interest because it is based on the method of covariance. It gives information about the magnitude of the association, or correlation, as well as the direction of the relationship.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/filter-based-feature-selection https://www.statisticssolutions.com/pearsons-correlation-coefficient/

**NEW QUESTION 232**
- (Exam Topic 3)
You are developing a linear regression model in Azure Machine Learning Studio. You run an experiment to compare different algorithms.
The following image displays the results dataset output:

| Algorithm | Mean Absolute Error | Root Mean Squared Error | Relative Absolute Error | Relative Squared Error |
|---|---|---|---|---|
| Bayesian Liner | 3.276025 | 4.655442 | 0.511436 | 0.282138 |
| Neural Network | 2.676538 | 3.621476 | 0.417847 | 0.17073 |
| Boosted Decision Tree | 2.168847 | 2.878077 | 0.338589 | 0.107831 |
| Linear | 6.350005 | 8.720718 | 0.99133 | 0.99002 |
| Decision Forest | 2.390206 | 3.315 164 | 0.373146 | 0.14307 |

Use the drop-down menus to select the answer choice that answers each question based on the information presented in the image.
NOTE: Each correct selection is worth one point.

| Question | Answer choice |
|---|---|
| Which algorithm minimizes differences between actual and predicted values? | Bayesian Linear Regression<br>Neural Network Regression<br>Boosted Decision Tree Regression<br>Linear Regression<br>Decision Forest Regression |
| Which approach should you use to find the best parameters for a Linear Regression model for the Online Gradient Descent method? | Set the Decrease learning rate option to True.<br>Set the Decrease learning rate option to True.<br>Set the Create trainer mode option to Parameter Range.<br>Increase the number of epochs.<br>Decrease the number of epochs. |

A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Box 1: Boosted Decision Tree Regression
Mean absolute error (MAE) measures how close the predictions are to the actual outcomes; thus, a lower score is better.
Box 2:
Online Gradient Descent: If you want the algorithm to find the best parameters for you, set Create trainer mode option to Parameter Range. You can then specify multiple values for the algorithm to try.
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/evaluate-model https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/linear-regression

**NEW QUESTION 235**
- (Exam Topic 3)
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 create an Azure Machine Learning service datastore in a workspace. The datastore contains the following files:
• /data/2018/Q1 .csv
• /data/2018/Q2.csv
• /data/2018/Q3.csv
• /data/2018/Q4.csv
• /data/2019/Q1.csv
All files store data in the following format:
• id,f1,f2,l
• 1,1,2,0
• 2,1,1,1
• 3.2.1.0
You run the following code:

```
data_store = Datastore.register_azure_blob_container(workspace=ws,
 datastore_name='data_store',
 container_name='quarterly_data',
 account_name='companydata',
 account_key='NRPxk8duxbM3…'
 create_if_not_exists=False)
```

You need to create a dataset named training_data and load the data from all files into a single data frame by using the following code:

```
data_frame = training_data.to_pandas_dataframe()
```

Solution: Run the following code:

```
from azureml.core import Dataset
paths = [(data_store,'data/2018/*.csv'),(data_store,'data/2019/*.csv')]
training_data = Dataset.Tabular.from_delimited_files(paths)
```

Does the solution meet the goal?

A. Yes
B. No

**Answer:** A

**Explanation:**
Use two file paths.
Use Dataset.Tabular_from_delimeted as the data isn't cleansed. Note:
A TabularDataset represents data in a tabular format by parsing the provided file or list of files. This provides you with the ability to materialize the data into a pandas or Spark DataFrame so you can work with familiar data preparation and training libraries without having to leave your notebook. You can create a TabularDataset object from .csv, .tsv, .parquet, .jsonl files, and from SQL query results.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-create-register-datasets

**NEW QUESTION 238**
- (Exam Topic 3)
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 creating a model to predict the price of a student's artwork depending on the following variables: the student's length of education, degree type, and art form.
You start by creating a linear regression model. You need to evaluate the linear regression model.
Solution: Use the following metrics: Mean Absolute Error, Root Mean Absolute Error, Relative Absolute Error, Accuracy, Precision, Recall, F1 score, and AUC.
Does the solution meet the goal?

A. Yes
B. No

**Answer:** B

**Explanation:**
Accuracy, Precision, Recall, F1 score, and AUC are metrics for evaluating classification models. Note: Mean Absolute Error, Root Mean Absolute Error, Relative Absolute Error are OK for the linear
regression model.
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/evaluate-model

**NEW QUESTION 243**
- (Exam Topic 3)
You create an Azure Machine Learning compute resource to train models. The compute resource is configured as follows:

- Minimum nodes: 2
- Maximum nodes: 4

You must decrease the minimum number of nodes and increase the maximum number of nodes to the following values:

- Minimum nodes: 0
- Maximum nodes: 8

You need to reconfigure the compute resource.
What are three possible ways to achieve this goal? Each correct answer presents a complete solution.
NOTE: Each correct selection is worth one point.

A. Use the Azure Machine Learning studio.
B. Run the update method of the AmlCompute class in the Python SDK.
C. Use the Azure portal.
D. Use the Azure Machine Learning designer.
E. Run the refresh_state() method of the BatchCompute class in the Python SDK

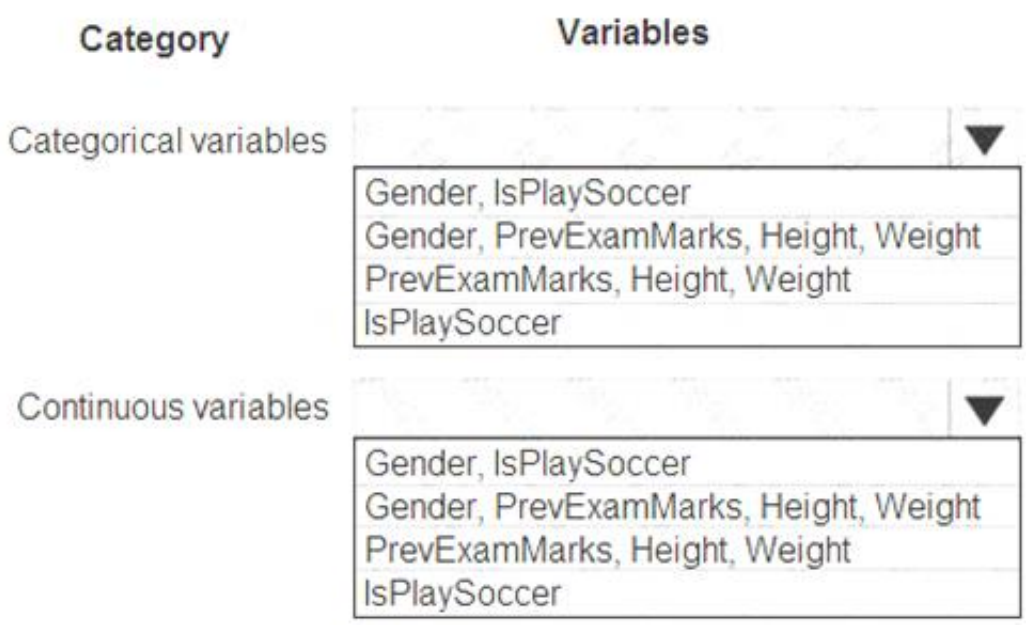
**Answer:** ABC

**Explanation:**
Reference:
https://docs.microsoft.com/en-us/python/api/azureml-core/azureml.core.compute.amlcompute(class)

**NEW QUESTION 246**
- (Exam Topic 3)
You are working on a classification task. You have a dataset indicating whether a student would like to play soccer and associated attributes. The dataset includes the following columns:

| Name | Description |
|---|---|
| IsPlaySoccer | Values can be 1 and 0. |
| Gender | Values can be M or F. |
| PrevExamMarks | Stores values from 0 to 100 |
| Height | Stores values in centimeters |
| Weight | Stores values in kilograms |

You need to classify variables by type.
Which variable should you add to each category? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.




A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
References: https://www.edureka.co/blog/classification-algorithms/

**NEW QUESTION 247**
- (Exam Topic 3)
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 a data scientist using Azure Machine Learning Studio.
You need to normalize values to produce an output column into bins to predict a target column. Solution: Apply a Quantiles binning mode with a PQuantile normalization.
Does the solution meet the goal?

A. Yes
B. No

**Answer:** B

**Explanation:**
Use the Entropy MDL binning mode which has a target column. References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/group-data-into-bins

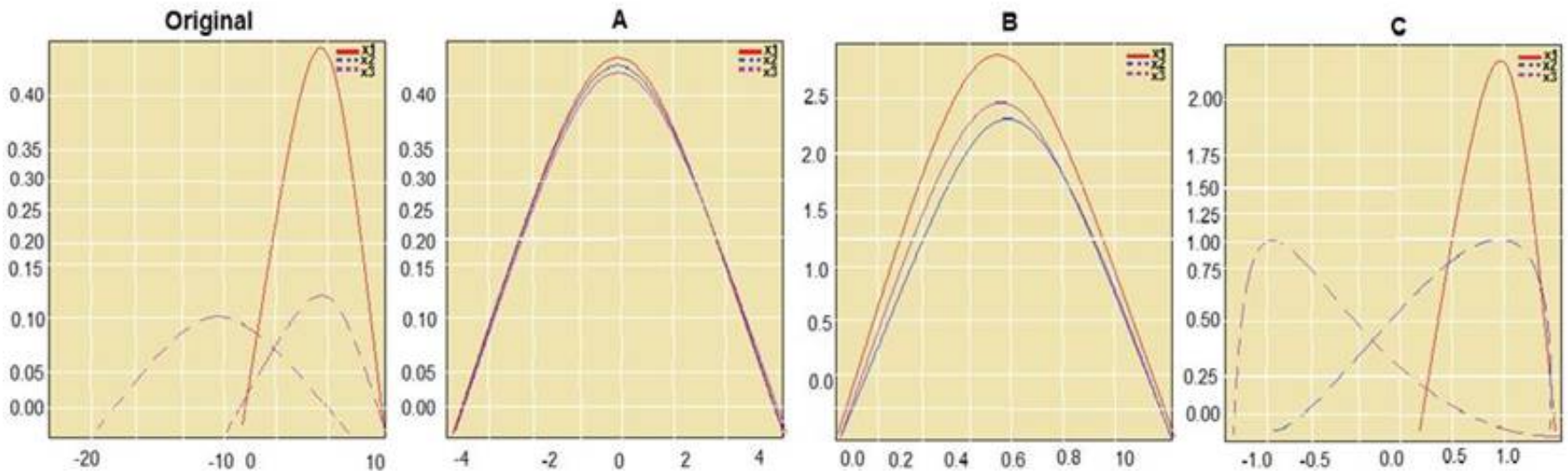
**NEW QUESTION 252**
- (Exam Topic 3)
You create a script that trains a convolutional neural network model over multiple epochs and logs the validation loss after each epoch. The script includes arguments for batch size and learning rate.
You identify a set of batch size and learning rate values that you want to try.
You need to use Azure Machine Learning to find the combination of batch size and learning rate that results in the model with the lowest validation loss.
What should you do?

A. Run the script in an experiment based on an AutoMLConfig object
B. Create a PythonScriptStep object for the script and run it in a pipeline
C. Use the Automated Machine Learning interface in Azure Machine Learning studio
D. Run the script in an experiment based on a ScriptRunConfig object
E. Run the script in an experiment based on a HyperDriveConfig object

**Answer:** E

**Explanation:**
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-tune-hyperparameters

**NEW QUESTION 255**
- (Exam Topic 3)
You are performing feature scaling by using the scikit-learn Python library for x.1 x2, and x3 features. Original and scaled data is shown in the following image.



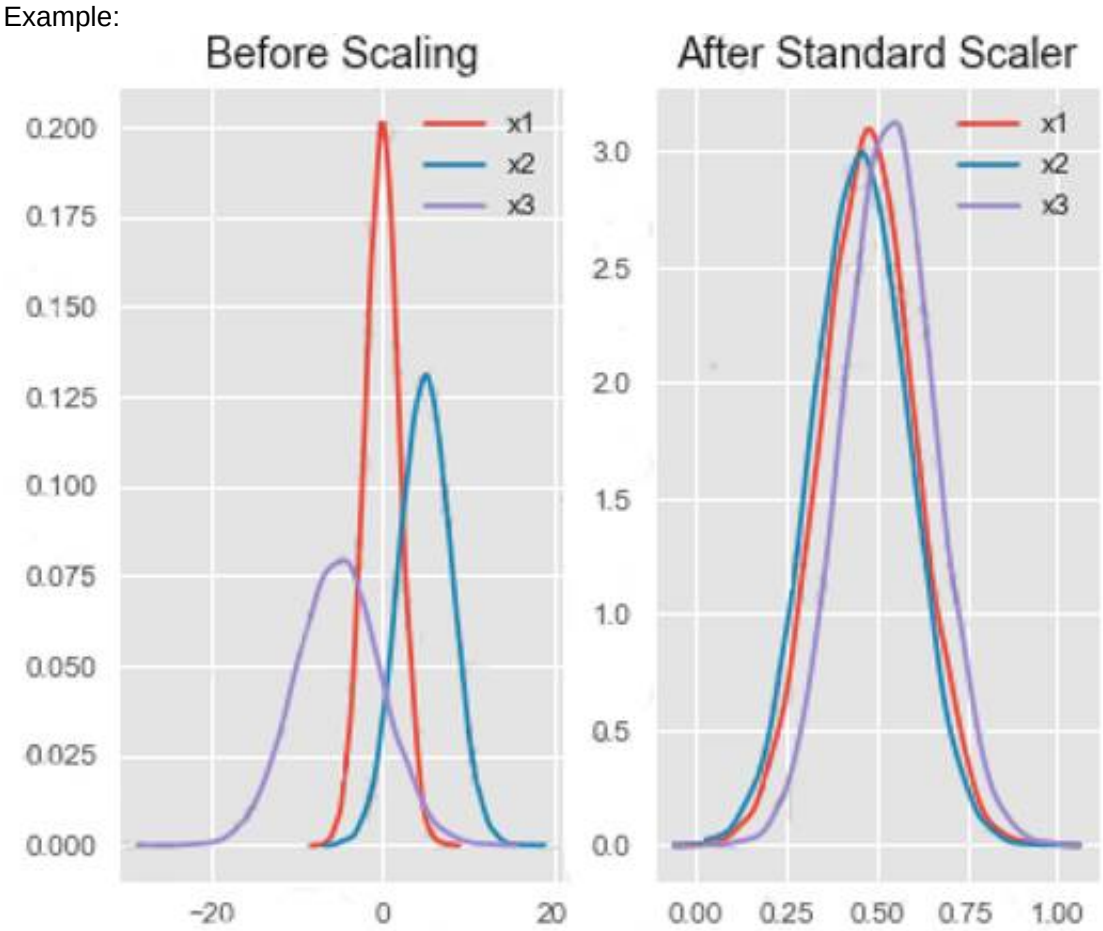

Use the drop-down menus to select the answer choice that answers each question based on the information presented in the graphic.
NOTE: Each correct selection is worth one point.

| Question | Answer choice |
|---|---|
| Which scaler is used in graph A? | Standard Scaler / Min Max Scale / Normalizer |
| Which scaler is used in graph B? | Standard Scaler / Min Max Scale / Normalizer |
| Which scaler is used in graph C? | Standard Scaler / Min Max Scale / Normalizer |

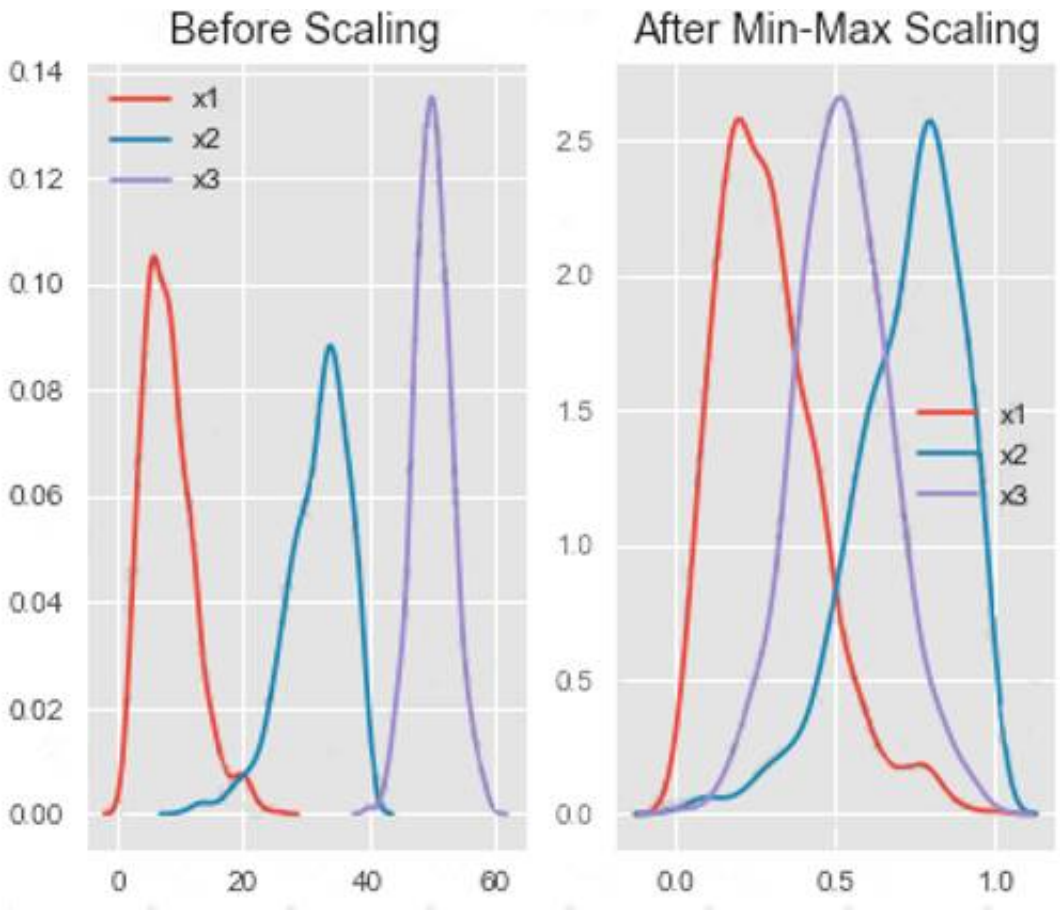
A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Box 1: StandardScaler
The StandardScaler assumes your data is normally distributed within each feature and will scale them such that the distribution is now centred around 0, with a standard deviation of 1.
Example:




All features are now on the same scale relative to one another. Box 2: Min Max Scaler




Notice that the skewness of the distribution is maintained but the 3 distributions are brought into the same scale so that they overlap.
Box 3: Normalizer
References:
http://benalexkeen.com/feature-scaling-with-scikit-learn/

**NEW QUESTION 256**
- (Exam Topic 3)
You use Azure Machine Learning to train a model based on a dataset named dataset1. You define a dataset monitor and create a dataset named dataset2 that contains new data.
You need to compare dataset1 and dataset2 by using the Azure Machine Learning SDK for Python. Which method of the DataDriftDetector class should you use?

A. run
B. get
C. backfill
D. update

**Answer:** C

**Explanation:**
A backfill run is used to see how data changes over time. Reference:
https://docs.microsoft.com/en-us/python/api/azureml-datadrift/azureml.datadrift.datadriftdetector.datadriftdetect

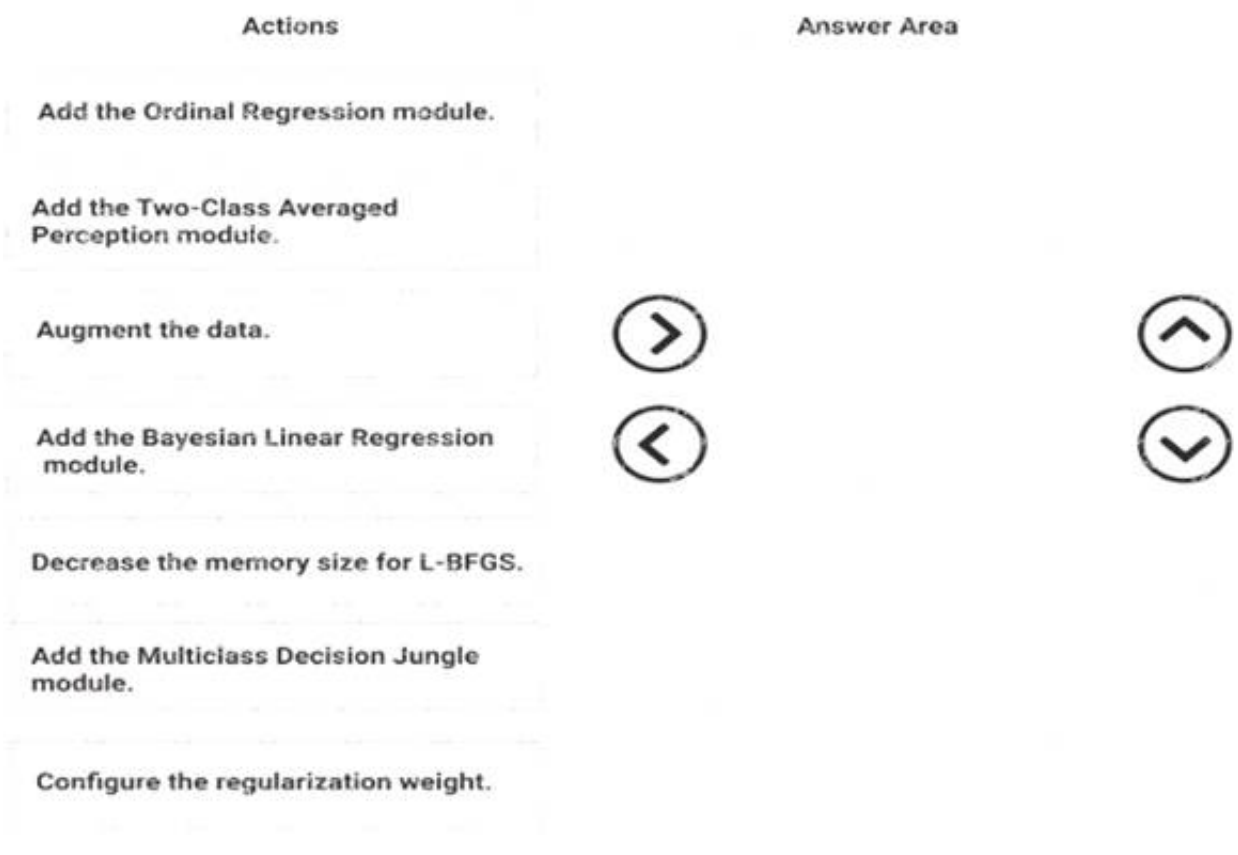
**NEW QUESTION 260**
- (Exam Topic 3)
An organization creates and deploys a multi-class image classification deep learning model that uses a set of labeled photographs.
The software engineering team reports there is a heavy inferencing load for the prediction web services during the summer. The production web service for the model fails to meet demand despite having a fully-utilized compute cluster where the web service is deployed.
You need to improve performance of the image classification web service with minimal downtime and minimal administrative effort.
What should you advise the IT Operations team to do?

A. Increase the minimum node count of the compute cluster where the web service is deployed.
B. Create a new compute cluster by using larger VM sizes for the nodes, redeploy the web service to that cluster, and update the DNS registration for the service endpoint to point to the new cluster.
C. Increase the VM size of nodes in the compute cluster where the web service is deployed.
D. Increase the node count of the compute cluster where the web service is deployed.

**Answer:** D

**Explanation:**
The Azure Machine Learning SDK does not provide support scaling an AKS cluster. To scale the nodes in the cluster, use the UI for your AKS cluster in the Azure Machine Learning studio. You can only change the node count, not the VM size of the cluster.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-create-attach-kubernetes

**NEW QUESTION 264**
- (Exam Topic 2)
You need to correct the model fit issue.
Which three 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.

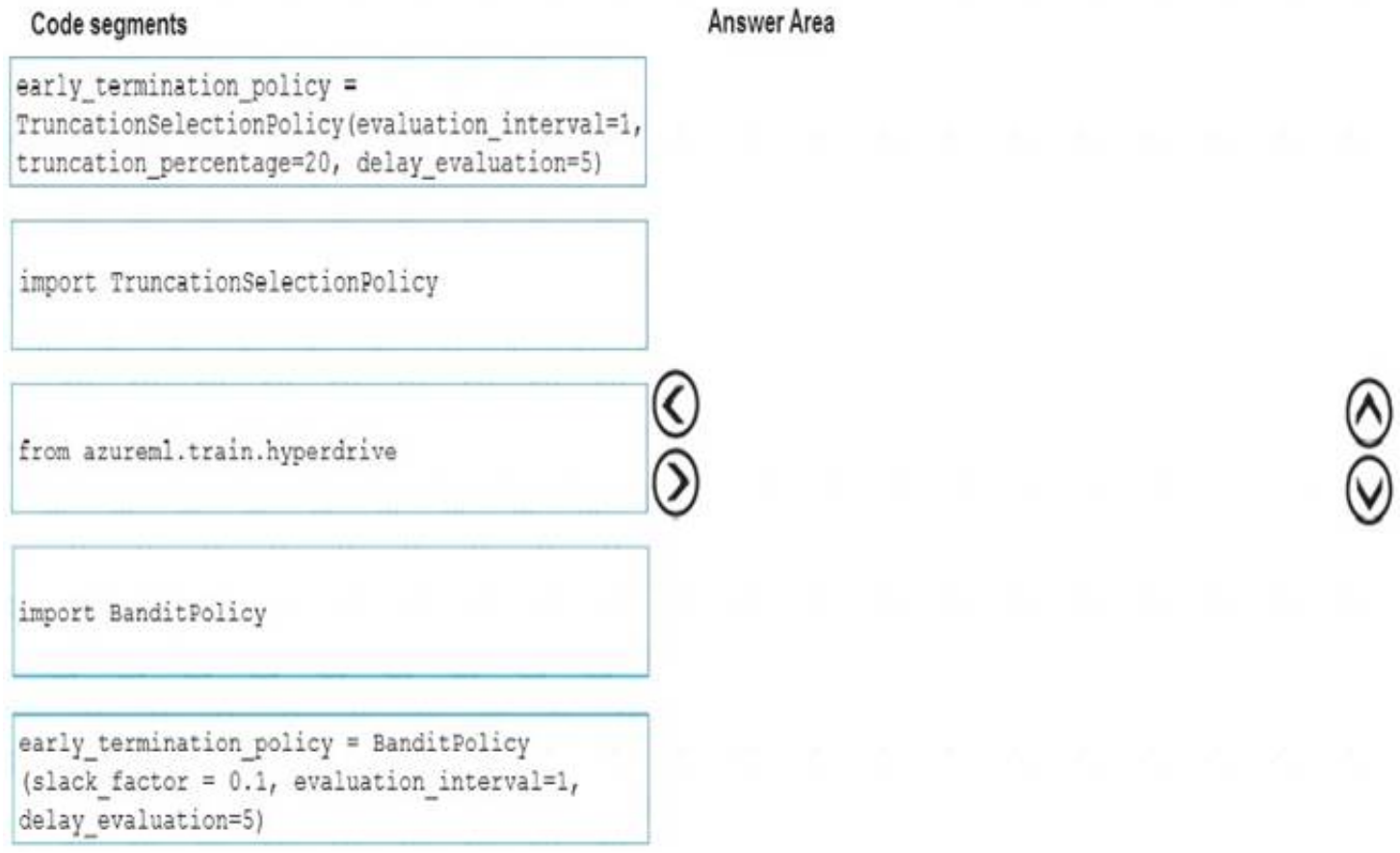



A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Step 1: Augment the data
Scenario: Columns in each dataset contain missing and null values. The datasets also contain many outliers.
Step 2: Add the Bayesian Linear Regression module.
Scenario: You produce a regression model to predict property prices by using the Linear Regression and Bayesian Linear Regression modules.
Step 3: Configure the regularization weight.
Regularization typically is used to avoid overfitting. For example, in L2 regularization weight, type the value to use as the weight for L2 regularization. We recommend that you use a non-zero value to avoid overfitting.
Scenario:
Model fit: The model shows signs of overfitting. You need to produce a more refined regression model that reduces the overfitting.

**NEW QUESTION 265**
- (Exam Topic 2)
You need to implement early stopping criteria as suited in the model training requirements.
Which three code segments should you use to develop the solution? To answer, move the appropriate code segments from the list of code segments to the answer area and arrange them in the correct order.
NOTE: More than one order of answer choices is correct. You will receive credit for any of the correct orders you select.




A. Mastered
B. Not Mastered

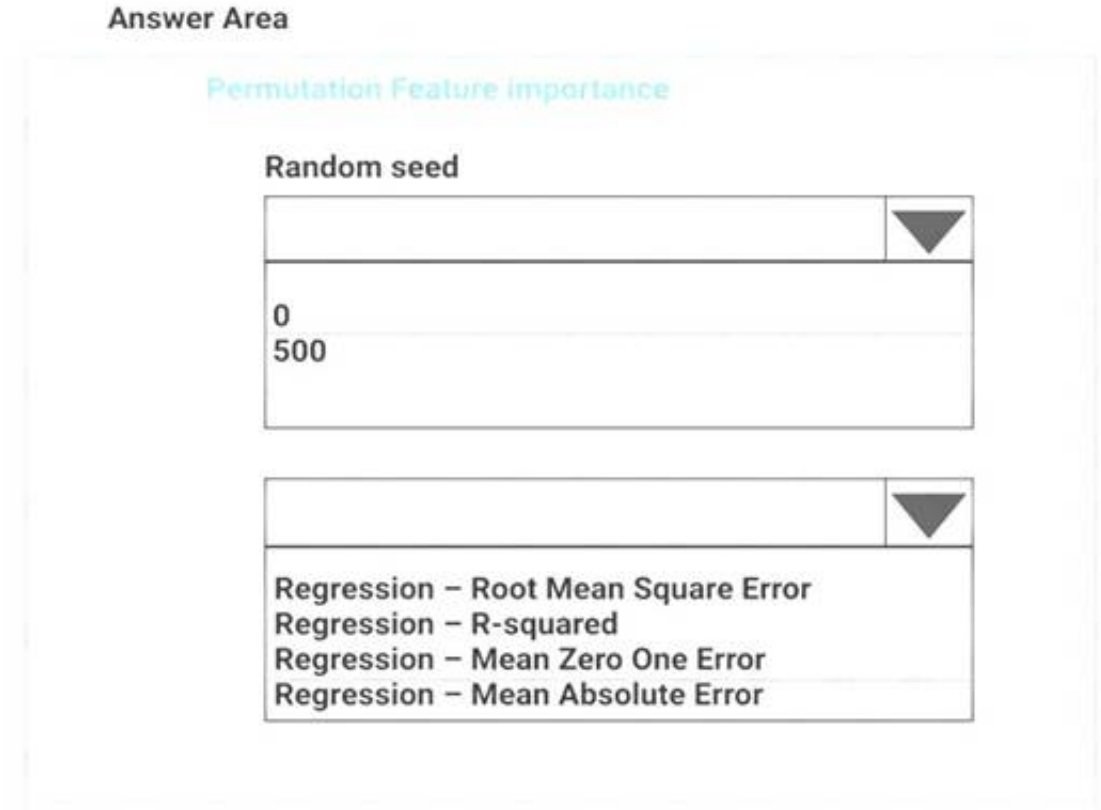
**Answer:** A

**Explanation:**

You need to implement an early stopping criterion on models that provides savings without terminating promising jobs.
Truncation selection cancels a given percentage of lowest performing runs at each evaluation interval. Runs are compared based on their performance on the primary metric and the lowest X% are terminated. Example:
from azureml.train.hyperdrive import TruncationSelectionPolicy
early_termination_policy = TruncationSelectionPolicy(evaluation_interval=1, truncation_percentage=20, delay_evaluation=5)

**NEW QUESTION 266**
- (Exam Topic 2)
You need to configure the Permutation Feature Importance module for the model training requirements. What should you do? To answer, select the appropriate options in the dialog box in the answer area. NOTE: Each correct selection is worth one point.

Answer Area




A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Box 1: 500
For Random seed, type a value to use as seed for randomization. If you specify 0 (the default), a number is generated based on the system clock.
A seed value is optional, but you should provide a value if you want reproducibility across runs of the same experiment.
Here we must replicate the findings. Box 2: Mean Absolute Error
Scenario: Given a trained model and a test dataset, you must compute the Permutation Feature Importance scores of feature variables. You need to set up the Permutation Feature Importance module to select the correct metric to investigate the model's accuracy and replicate the findings.
Regression. Choose one of the following: Precision, Recall, Mean Absolute Error , Root Mean Squared Error, Relative Absolute Error, Relative Squared Error, Coefficient of Determination
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/permutation-feature-importan

**NEW QUESTION 271**
- (Exam Topic 1)
You need to implement a feature engineering strategy for the crowd sentiment local models. What should you do?

A. Apply an analysis of variance (ANOVA).
B. Apply a Pearson correlation coefficient.
C. Apply a Spearman correlation coefficient.
D. Apply a linear discriminant analysis.

**Answer:** D

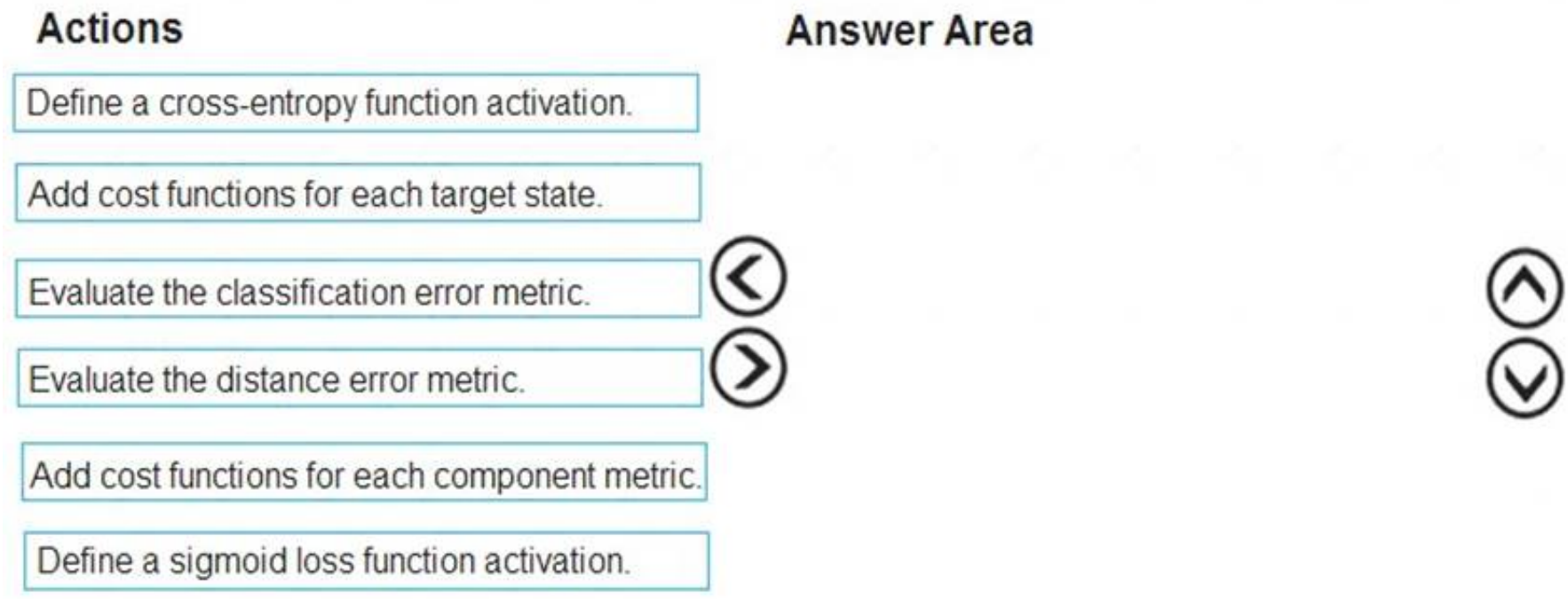
**Explanation:**
The linear discriminant analysis method works only on continuous variables, not categorical or ordinal variables.
Linear discriminant analysis is similar to analysis of variance (ANOVA) in that it works by comparing the means of the variables.
Scenario:
Data scientists must build notebooks in a local environment using automatic feature engineering and model building in machine learning pipelines.
Experiments for local crowd sentiment models must combine local penalty detection data. All shared features for local models are continuous variables.

**NEW QUESTION 273**
- (Exam Topic 1)
You need to resolve the local machine learning pipeline performance issue. What should you do?

A. Increase Graphic Processing Units (GPUs).
B. Increase the learning rate.
C. Increase the training iterations,
D. Increase Central Processing Units (CPUs).

**Answer:** A

**NEW QUESTION 276**
- (Exam Topic 1)
You need to define an evaluation strategy for the crowd sentiment models.
Which three 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.

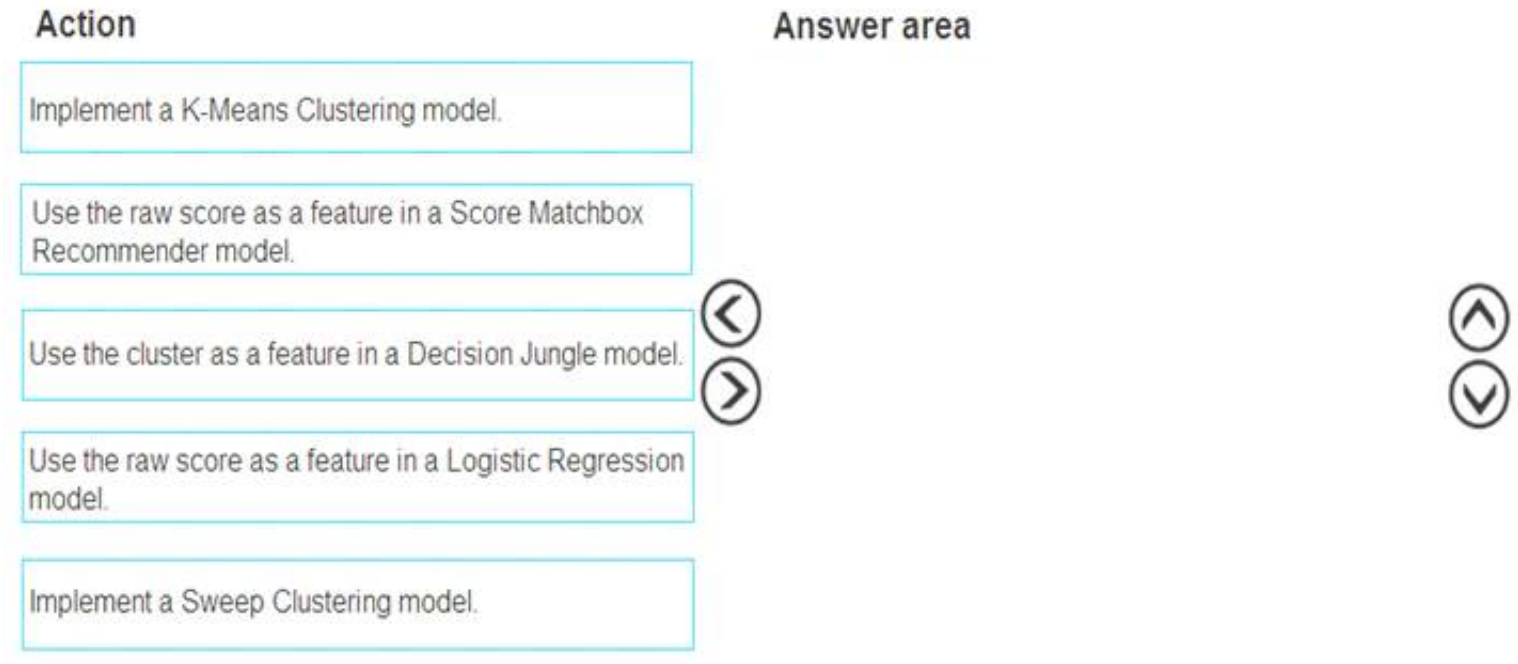



A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Step 1: Define a cross-entropy function activation
When using a neural network to perform classification and prediction, it is usually better to use cross-entropy error than classification error, and somewhat better to use cross-entropy error than mean squared error to evaluate the quality of the neural network.
Step 2: Add cost functions for each target state. Step 3: Evaluated the distance error metric. References:
https://www.analyticsvidhya.com/blog/2018/04/fundamentals-deep-learning-regularization-techniques/

**NEW QUESTION 280**
- (Exam Topic 1)
You need to define a modeling strategy for ad response.
Which three 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.




A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Step 1: Implement a K-Means Clustering model
Step 2: Use the cluster as a feature in a Decision jungle model.
Decision jungles are non-parametric models, which can represent non-linear decision boundaries. Step 3: Use the raw score as a feature in a Score Matchbox Recommender model
The goal of creating a recommendation system is to recommend one or more "items" to "users" of the system. Examples of an item could be a movie, restaurant, book, or song. A user could be a person, group of persons, or other entity with item preferences.
Scenario:

Ad response rated declined.
Ad response models must be trained at the beginning of each event and applied during the sporting event. Market segmentation models must optimize for similar ad response history.
Ad response models must support non-linear boundaries of features. References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/multiclass-decision-jungle https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/score-matchbox-recommende

**NEW QUESTION 285**
- (Exam Topic 1)
You need to use the Python language to build a sampling strategy for the global penalty detection models. How should you complete the code segment? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.


A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Box 1: import pytorch as deeplearninglib Box 2: ..DistributedSampler(Sampler).. DistributedSampler(Sampler):
Sampler that restricts data loading to a subset of the dataset.
It is especially useful in conjunction with class:`torch.nn.parallel.DistributedDataParallel`. In such case, each process can pass a DistributedSampler instance as a DataLoader sampler, and load a subset of the original dataset that is exclusive to it.
Scenario: Sampling must guarantee mutual and collective exclusively between local and global segmentation models that share the same features.
Box 3: optimizer = deeplearninglib.train. GradientDescentOptimizer(learning_rate=0.10)

**NEW QUESTION 289**
.....

# Thank You for Trying Our Product

## We offer two products:

1st - We have Practice Tests Software with Actual Exam Questions

2nd - Questons and Answers in PDF Format

## DP-100 Practice Exam Features:

* DP-100 Questions and Answers Updated Frequently

* DP-100 Practice Questions Verified by Expert Senior Certified Staff

* DP-100 Most Realistic Questions that Guarantee you a Pass on Your FirstTry

* DP-100 Practice Test Questions in Multiple Choice Formats and Updatesfor 1 Year

100% Actual & Verified — Instant Download, Please Click
Order The DP-100 Practice Test Here