DP-100 Dumps

Designing and Implementing a Data Science Solution on Azure

https://www.certleader.com/DP-100-dumps.html

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.

Code Segment	Evaluation Statement
class_weight=balanced	▼
	Automatically select the performance metrics for the classification.
	Automatically adjust weights directly proportional to class frequencies in the input data.
	Automatically adjust weights inversely proportional to class frequences in the input data.
C parameter	▼
	Penalty parameter
	Degreee of polynomial kernel function
	Size of the kernel cache

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)
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.

Action		Answer area
Evaluate the probability function		
Remove duplicate rows		
Use the Filter Based Feature Selection module	ⓒ ⓓ	⋀ ⋁
Test the hypothesis using t-Test		
Compute linear correlation		
Build a counting transform		

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 5
- (Exam Topic 3)
You have a dataset that includes confidential data. You use the dataset to train a model.
You must use a differential privacy parameter to keep the data of individuals safe and private. You need to reduce the effect of user data on aggregated results.
What should you do?

A. Decrease the value of the epsilon parameter to reduce the amount of noise added to the data
B. Increase the value of the epsilon parameter to decrease privacy and increase accuracy
C. Decrease the value of the epsilon parameter to increase privacy and reduce accuracy
D. Set the value of the epsilon parameter to 1 to ensure maximum privacy

Answer: C

Explanation:
Differential privacy tries to protect against the possibility that a user can produce an indefinite number of reports to eventually reveal sensitive data. A value known as epsilon measures how noisy, or private, a report is. Epsilon has an inverse relationship to noise or privacy. The lower the epsilon, the more noisy (and private) the data is.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/concept-differential-privacy

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 underrepresented 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 are developing a deep learning model by using TensorFlow. You plan to run the model training workload on an Azure Machine Learning Compute Instance.
You must use CUDA-based model training. You need to provision the Compute Instance.
Which two virtual machines sizes can you use? To answer, select the appropriate virtual machine sizes in the answer area.
NOTE: Each correct selection is worth one point.

Virtual machine size

Search by name...

Name ↑	vCPUs	GPUs	RAM	Resource disk
BASIC_A0	1		0.75 GB	20 GB
STANDARD_D3_V2	4		14 GB	200 GB
STANDARD_E64_V3	64		432 GB	1,600 GB
STANDARD_M64LS	64		512 GB	2,000 GB
STANDARD_NC12	12	2	112 GB	680 GB
STANDARD_NC24	24	4	224 GB	1,440 GB

A. Mastered

B. Not Mastered

Answer: A

Explanation:
CUDA is a parallel computing platform and programming model developed by Nvidia for general computing on its own GPUs (graphics processing units). CUDA enables developers to speed up compute-intensive applications by harnessing the power of GPUs for the parallelizable part of the computation.
Reference:
https://www.infoworld.com/article/3299703/what-is-cuda-parallel-programming-for-gpus.html

NEW QUESTION 8
- (Exam Topic 3)
You train a model and register it in your Azure Machine Learning workspace. You are ready to deploy the model as a real-time web service.
You deploy the model to an Azure Kubernetes Service (AKS) inference cluster, but the deployment fails because an error occurs when the service runs the entry script that is associated with the model deployment.
You need to debug the error by iteratively modifying the code and reloading the service, without requiring a re-deployment of the service for each code update.
What should you do?

A. Register a new version of the model and update the entry script to load the new version of the model from its registered path.
B. Modify the AKS service deployment configuration to enable application insights and re-deploy to AKS.
C. Create an Azure Container Instances (ACI) web service deployment configuration and deploy the model on ACI.
D. Add a breakpoint to the first line of the entry script and redeploy the service to AKS.
E. Create a local web service deployment configuration and deploy the model to a local Docker container.

Answer: C

Explanation:
How to work around or solve common Docker deployment errors with Azure Container Instances (ACI) and Azure Kubernetes Service (AKS) using Azure Machine Learning.
The recommended and the most up to date approach for model deployment is via the Model.deploy() API using an Environment object as an input parameter. In this case our service will create a base docker image for you during deployment stage and mount the required models all in one call. The basic deployment tasks are:
* 1. Register the model in the workspace model registry.
* 2. Define Inference Configuration:
* a. Create an Environment object based on the dependencies you specify in the environment yaml file or use one of our procured environments.
* b. Create an inference configuration (InferenceConfig object) based on the environment and the scoring script.
* 3. Deploy the model to Azure Container Instance (ACI) service or to Azure Kubernetes Service (AKS).

NEW QUESTION 9
- (Exam Topic 3)
You use the designer to create a training pipeline for a classification model. The pipeline uses a dataset that includes the features and labels required for model training.
You create a real-time inference pipeline from the training pipeline. You observe that the schema for the generated web service input is based on the dataset and includes the label column that the model predicts. Client applications that use the service must not be required to submit this value.
You need to modify the inference pipeline to meet the requirement. What should you do?

A. Add a Select Columns in Dataset module to the inference pipeline after the dataset and use it to select all columns other than the label.
B. Delete the dataset from the training pipeline and recreate the real-time inference pipeline.
C. Delete the Web Service Input module from the inference pipeline.
D. Replace the dataset in the inference pipeline with an Enter Data Manually module that includes data for the feature columns but not the label column.

Answer: A

Explanation:
By default, the Web Service Input will expect the same data schema as the module output data which connects to the same downstream port as it. You can remove the target variable column in the inference pipeline using Select Columns in Dataset module. Make sure that the output of Select Columns in Dataset removing target variable column is connected to the same port as the output of the Web Service Intput module.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/tutorial-designer-automobile-price-deploy

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 have a Python script named train.py in a local folder named scripts. The script trains a regression model by using scikit-learn. The script includes code to load a training data file which is also located in the scripts folder.
You must run the script as an Azure ML experiment on a compute cluster named aml-compute.
You need to configure the run to ensure that the environment includes the required packages for model training. You have instantiated a variable named aml-compute that references the target compute cluster.
Solution: Run the following code:

```
from azureml.train.dnn import TensorFlow

sk_est = TensorFlow(source_directory='./scripts',
  compute_target=aml-compute,
  entry_script='train.py')
```

Does the solution meet the goal?

A. Yes
B. No

Answer: B

Explanation:
The scikit-learn estimator provides a simple way of launching a scikit-learn training job on a compute target. It is implemented through the SKLearn class, which can be used to support single-node CPU training.
Example:
from azureml.train.sklearn import SKLearn
}
estimator = SKLearn(source_directory=project_folder, compute_target=compute_target, entry_script='train_iris.py'
)
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-train-scikit-learn

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 11
- (Exam Topic 3)
You retrain an existing model.
You need to register the new version of a model while keeping the current version of the model in the registry.
What should you do?

A. Register a model with a different name from the existing model and a custom property named versionwith the value 2.
B. Register the model with the same name as the existing model.
C. Save the new model in the default datastore with the same name as the existing mode
D. Do not register the new model.
E. Delete the existing model and register the new one with the same name.

Answer: B

Explanation:
Model version: A version of a registered model. When a new model is added to the Model Registry, it is added as Version 1. Each model registered to the same model name increments the version number.
Reference:
https://docs.microsoft.com/en-us/azure/databricks/applications/mlflow/model-registry

NEW QUESTION 12
- (Exam Topic 3)

You are training machine learning models in Azure Machine Learning. You use Hyperdrive to tune the hyperparameters. In previous model training and tuning runs, many models showed similar performance. You need to select an early termination policy that meets the following requirements:
• accounts for the performance of all previous runs when evaluating the current run
• avoids comparing the current run with only the best performing run to date
Which two early termination policies should you use? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

A. Bandit
B. Median stopping
C. Default
D. Truncation selection

Answer: BC

Explanation:
The Median Stopping policy computes running averages across all runs and cancels runs whose best performance is worse than the median of the running averages.
If no policy is specified, the hyperparameter tuning service will let all training runs execute to completion. Reference:
https://docs.microsoft.com/en-us/python/api/azureml-train-core/azureml.train.hyperdrive.medianstoppingpolicy https://docs.microsoft.com/en-us/python/api/azureml-train-core/azureml.train.hyperdrive.truncationselectionpoli https://docs.microsoft.com/en-us/python/api/azureml-train-core/azureml.train.hyperdrive.banditpolicy

NEW QUESTION 16
- (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.

Home > New > Machine Learning >

Machine Learning 🖶

Create a machine learning workspace

Basics   Networking   Advanced   Tags   Review + create

Project details

Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

Subscription * ⓘ   Documentation-team   ∨

        Resource group * ⓘ   docs-ws   ∨
                                     Create new

Workspace details

Specify the name, region, and edition for the workspace.

Workspace name * ⓘ   docs-mlw   ✓

Region * ⓘ   West Central US   ∨

Workspace edition * ⓘ   Basic   ∧

                              Basic

                              Enterprise

ⓘ For your convenience, these resources
Application Insights, Azure Key Vault

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 19
- (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 22
- (Exam Topic 3) You are solving a classification task. The dataset is imbalanced.
You need to select an Azure Machine Learning Studio module to improve the classification accuracy. Which module should you use?

A. Fisher Linear Discriminant Analysis.
B. Filter Based Feature Selection
C. Synthetic Minority Oversampling Technique (SMOTE)
D. Permutation Feature Importance

Answer: C

Explanation:
Use the SMOTE module in Azure Machine Learning Studio (classic) 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.
You connect the SMOTE module to a dataset that is imbalanced. There are many reasons why a dataset might be imbalanced: the category you are targeting might be very rare in the population, or the data might simply be difficult to collect. Typically, you use SMOTE when the class you want to analyze is under-

represented.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/smote

NEW QUESTION 23
- (Exam Topic 3)
You are evaluating a completed binary classification machine learning model. You need to use the precision as the valuation metric.
Which visualization should you use?

A. Binary classification confusion matrix
B. box plot
C. Gradient descent
D. coefficient of determination

Answer: A

Explanation:
 References:
https://machinelearningknowledge.ai/confusion-matrix-and-performance-metrics-machine-learning/

NEW QUESTION 27
- (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 32
- (Exam Topic 3)
You create an Azure Machine Learning workspace and a new Azure DevOps organization. You register a
model in the workspace and deploy the model to the target environment.
All new versions of the model registered in the workspace must automatically be deployed to the target environment.

You need to configure Azure Pipelines to deploy the model.
Which four actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

Actions

| Answer Area |

Create a service connection

Create a release pipeline

Create a build pipeline

Create an Azure DevOps project

Install the Machine Learning extension for Azure Pipelines

A. Mastered
B. Not Mastered

Answer: A

Explanation:
Graphical user interface, text, application, email Description automatically generated
Step 1: Create an Azure DevOps project
Step 2: Create a release pipeline

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

≫ Go to Pipelines, and then select New pipeline.
Step 3: Install the Machine Learning extension for Azure Pipelines
You must install and configure the Azure CLI and ML extension.
Step 4: Create a service connection
How to set up your service connection
Graphical user interface, text, application, email Description automatically generated

Select AzureMLWorkspace for the scope level, then fill in the following subsequent parameters. Graphical user interface, text, application Description automatically generated
Note: How to enable model triggering in a release pipeline

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

≫ Enable the deployment trigger on your model artifact as shown here. Every time a new version of that model is registered, a release pipeline will be triggered.
Reference:
https://marketplace.visualstudio.com/items?itemName=ms-air-aiagility.vss-services-azureml https://docs.microsoft.com/en-us/azure/devops/pipelines/targets/azure-machine-learning

NEW QUESTION 34

- (Exam Topic 3)

You are the owner of an Azure Machine Learning workspace.

You must prevent the creation or deletion of compute resources by using a custom role. You must allow all other operations inside the workspace.

You need to configure the custom role.

How should you complete the configuration? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

```
{
  "Name": "Data Scientist Custom",
  "IsCustom": true,
  "Description": "Description"
  "Actions": [
```

▼

```
    Microsoft.MachineLearningServices/workspaces/*/read
    Microsoft.MachineLearningServices/workspaces/computes/*/write
    Microsoft.MachineLearningServices/workspaces/delete
```

▼

```
    Microsoft.MachineLearningServices/workspaces/*/write
    Microsoft.MachineLearningServices/workspaces/computes/*/write
    Microsoft.MachineLearningServices/workspaces/delete
```

```
  ],
  "NotActions": [
```

▼

```
    Microsoft.MachineLearningServices/workspaces/*/read
    Microsoft.MachineLearningServices/workspaces/*/write
    Microsoft.MachineLearningServices/workspaces/computes/*/delete
```

▼

```
    Microsoft.MachineLearningServices/workspaces/*/read
    Microsoft.MachineLearningServices/workspaces/*/write
    Microsoft.MachineLearningServices/workspaces/computes/*/write
```

```
  ],
  "AssignableScopes": [
    "/subscriptions/<subscription_id>"
  ]
}
```

A. Mastered
B. Not Mastered

Answer: A

Explanation:

Graphical user interface, application Description automatically generated

Graphical user interface, application Description automatically generated

Box 1: Microsoft.MachineLearningServices/workspaces/*/read

Reader role: Read-only actions in the workspace. Readers can list and view assets, including datastore credentials, in a workspace. Readers can't create or update these assets.

Box 2: Microsoft.MachineLearningServices/workspaces/*/write

If the roles include Actions that have a wildcard (*), the effective permissions are computed by subtracting the NotActions from the allowed Actions.

Box 3: Box 2: Microsoft.MachineLearningServices/workspaces/computes/*/delete

Box 4: Microsoft.MachineLearningServices/workspaces/computes/*/write Reference:

https://docs.microsoft.com/en-us/azure/role-based-access-control/overview#how-azure-rbac-determines-if-a-use

NEW QUESTION 37

- (Exam Topic 3)

You are using a Git repository to track work in an Azure Machine Learning workspace. You need to authenticate a Git account by using SSH.

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.

Actions	Answer Area
Generate a public/private key pair	
Add the private key to the Git account	
Clone the Git repository by using an SSH repository URL	
Add the public key to the Git account	
Create a new Azure Key Vault resource	

A. Mastered
B. Not Mastered

Answer: A

Explanation:
Graphical user interface, text, application, chat or text message Description automatically generated
Authenticate your Git Account with SSH: Step 1: Generating a public/private key pair Generate a new SSH key
* 1. Open the terminal window in the Azure Machine Learning Notebook Tab.
* 2. Paste the text below, substituting in your email address. ssh-keygen -t rsa -b 4096 -C "your_email@example.com"
This creates a new ssh key, using the provided email as a label.
> Generating public/private rsa key pair.
Step 2: Add the public key to the Git Account
In your terminal window, copy the contents of your public key file. Step 3: Clone the Git repository by using an SSH repository URL
* 1. Copy the SSH Git clone URL from the Git repo.
* 2. Paste the url into the git clone command below, to use your SSH Git repo URL. This will look something like:
git clone git@example.com:GitUser/azureml-example.git Cloning into 'azureml-example'.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/concept-train-model-git-integration

NEW QUESTION 40
- (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 45
- (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: Replace each missing value using the Multiple Imputation by Chained Equations (MICE) method. Does the solution meet the goal?

A. Yes
B. NO

Answer: A

Explanation:
Replace using MICE: For each missing value, this option assigns a new value, which is calculated by using a method described in the statistical literature as "Multivariate Imputation using Chained Equations" or "Multiple Imputation by Chained Equations". With a multiple imputation method, each variable with missing data is modeled conditionally using the other variables in the data before filling in the missing values.
Note: Multivariate imputation by chained equations (MICE), sometimes called "fully conditional specification" or "sequential regression multiple imputation" has

emerged in the statistical literature as one principled method of addressing missing data. Creating multiple imputations, as opposed to single imputations, accounts for the statistical uncertainty in the imputations. In addition, the chained equations approach is very flexible and can handle variables of varying types (e.g., continuous or binary) as well as complexities such as bounds or survey skip patterns.
References: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3074241/
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/clean-missing-data

NEW QUESTION 48
- (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.

What is the r-value for the correlation of Y to Z?

-0.106276
0.149676
0.859122
1

Which type of relationship exists between Z and Y in the feature set?

a positive linear relationship
a negative linear relationship
no linear relationship

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 49
- (Exam Topic 3)
You need to select a pre built development environment for a series of data science experiments. You must use the R language for the experiments.
Which three environments can you use? Each correct answer presents a complete solution. NOTE: Each correct selection is worth one point.

A. MI.NET Library on a local environment
B. Azure Machine Learning Studio
C. Data Science Virtual Machine (OSVM)
D. Azure Data bricks
E. Azure Cognitive Services

Answer: ABD

NEW QUESTION 51
- (Exam Topic 3)
You train and register a machine learning model. You create a batch inference pipeline that uses the model to generate predictions from multiple data files.
You must publish the batch inference pipeline as a service that can be scheduled to run every night. You need to select an appropriate compute target for the inference service.
Which compute target should you use?

A. Azure Machine Learning compute instance
B. Azure Machine Learning compute cluster
C. Azure Kubernetes Service (AKS)-based inference cluster
D. Azure Container Instance (ACI) compute target

Answer: B

Explanation:
Azure Machine Learning compute clusters is used for Batch inference. Run batch scoring on serverless compute. Supports normal and low-priority VMs. No support for real-time inference.

Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/concept-compute-target

NEW QUESTION 53
- (Exam Topic 3)
You use the Azure Machine Learning SDK in a notebook to run an experiment using a script file in an experiment folder.
The experiment fails.
You need to troubleshoot the failed experiment.
What are two possible ways to achieve this goal? Each correct answer presents a complete solution.

A. Use the get_metrics() method of the run object to retrieve the experiment run logs.
B. Use the get_details_with_logs() method of the run object to display the experiment run logs.
C. View the log files for the experiment run in the experiment folder.
D. View the logs for the experiment run in Azure Machine Learning studio.
E. Use the get_output() method of the run object to retrieve the experiment run logs.

Answer: BD

Explanation:
Use get_details_with_logs() to fetch the run details and logs created by the run.
You can monitor Azure Machine Learning runs and view their logs with the Azure Machine Learning studio. Reference:
https://docs.microsoft.com/en-us/python/api/azureml-pipeline-core/azureml.pipeline.core.steprun https://docs.microsoft.com/en-us/azure/machine-learning/how-to-monitor-view-training-logs

NEW QUESTION 57
- (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.
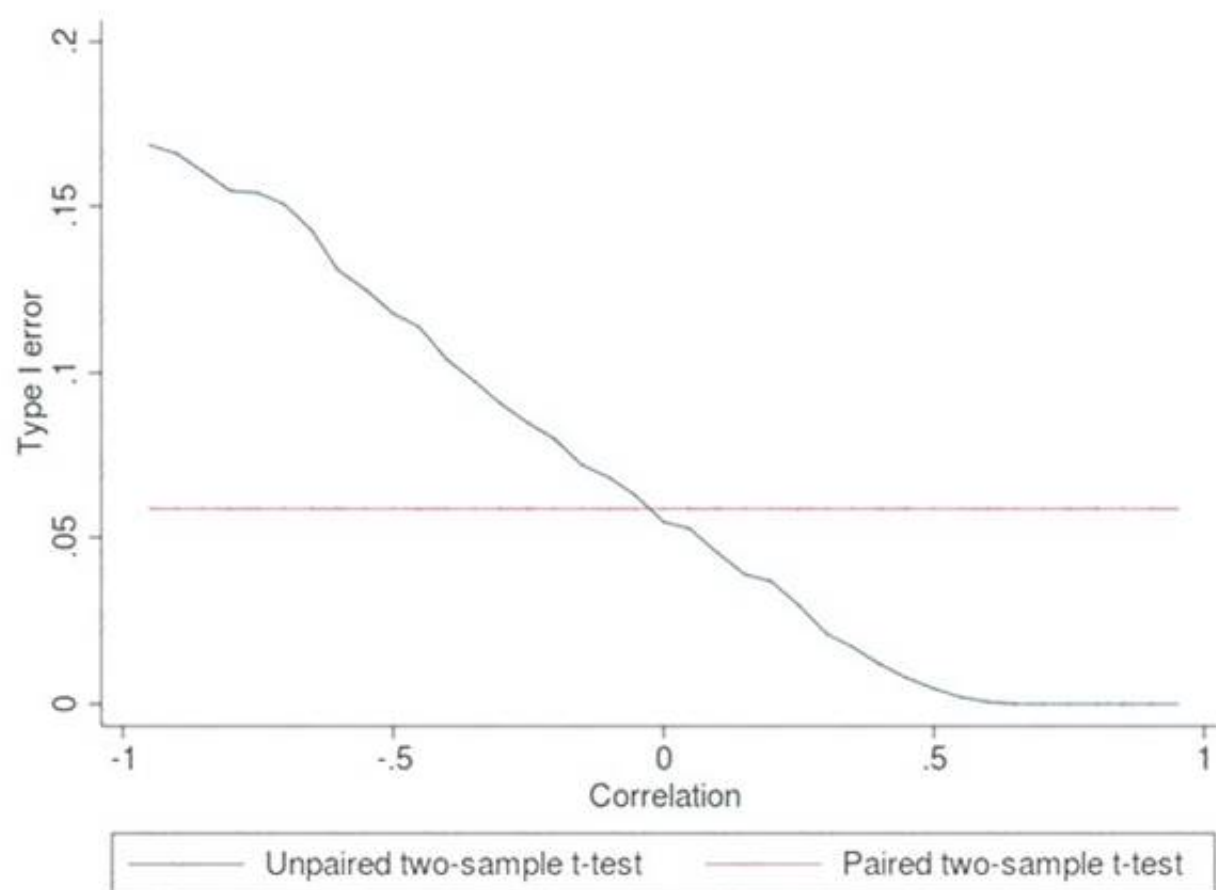Which type of distribution should you produce?

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 59
- (Exam Topic 3)
You create an Azure Databricks workspace and a linked Azure Machine Learning workspace. You have the following Python code segment in the Azure Machine
Learning workspace: import mlflow
import mlflow.azureml import azureml.mlflow import azureml.core
from azureml.core import Workspace subscription_id = 'subscription_id' resourse_group = 'resource_group_name' workspace_name = 'workspace_name'

ws = Workspace.get(name=workspace_name, subscription_id=subscription_id, resource_group=resource_group)
experimentName = "/Users/{user_name}/{experiment_folder}/{experiment_name}" mlflow.set_experiment(experimentName)
uri = ws.get_mlflow_tracking_uri() mlflow.set_tracking_uri(uri)
Instructions: 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
A resource group and Azure Machine Learning workspace will be created.	○	○
An Azure Databricks experiment will be tracked only in the Azure Machine Learning workspace.	○	○
The epoch loss metric is set to be tracked.	○	○

A. Mastered
B. Not Mastered

Answer: A

Explanation:
A screenshot of a computer Description automatically generated with medium confidence
Box 1: No
The Workspace.get method loads an existing workspace without using configuration files. ws = Workspace.get(name="myworkspace",
subscription_id='<azure-subscription-id>', resource_group='myresourcegroup')
Box 2: Yes
MLflow Tracking with Azure Machine Learning lets you store the logged metrics and artifacts from your local runs into your Azure Machine Learning workspace.
The get_mlflow_tracking_uri() method assigns a unique tracking URI address to the workspace, ws, and set_tracking_uri() points the MLflow tracking URI to that
address.
Box 3: Yes
Note: In Deep Learning, epoch means the total dataset is passed forward and backward in a neural network once.
Reference:
https://docs.microsoft.com/en-us/python/api/azureml-core/azureml.core.workspace.workspace https://docs.microsoft.com/en-us/azure/machine-learning/how-to-
use-mlflow

NEW QUESTION 62
- (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_input = PipelineData("raw_data", datastore=rawdatastore)
data_output = PipelineData("processed_data", datastore=datastore)
process_step = PythonScriptStep(script_name="process.py",
  arguments=["--data_for_train", data_input],
  outputs=[data_output], compute_target=aml_compute,
  source_directory=process_directory)
train_step = PythonScriptStep(script_name="train.py",
  arguments=["--data_for_train", data_input], inputs=[data_output],
  compute_target=aml_compute, source_directory=train_directory)
pipeline = Pipeline(workspace=ws, steps=[process_step, train_step])
```

Does the solution meet the goal?

A. Yes
B. No

Answer: B

Explanation:
Note: Data used in pipeline can be produced by one step and consumed in another step by providing a PipelineData object as an output of one step and an input
of one or more subsequent steps.
Compare with this example, the pipeline train step depends on the process_step_output output of the pipeline process step:
from azureml.pipeline.core import Pipeline, PipelineData from azureml.pipeline.steps import PythonScriptStep
datastore = ws.get_default_datastore()
process_step_output = PipelineData("processed_data", datastore=datastore) process_step = PythonScriptStep(script_name="process.py",
arguments=["--data_for_train", process_step_output], outputs=[process_step_output], compute_target=aml_compute, source_directory=process_directory)
train_step = PythonScriptStep(script_name="train.py", arguments=["--data_for_train", process_step_output], inputs=[process_step_output],

compute_target=aml_compute, source_directory=train_directory)
pipeline = Pipeline(workspace=ws, steps=[process_step, train_step]) Reference:
https://docs.microsoft.com/en-us/python/api/azureml-pipeline-core/azureml.pipeline.core.pipelinedata?view=azu

NEW QUESTION 64
- (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 66
- (Exam Topic 3)
You are developing a hands-on workshop to introduce Docker for Windows to attendees. You need to ensure that workshop attendees can install Docker on their devices.
Which two prerequisite components should attendees install on the devices? Each correct answer presents part of the solution.
NOTE: Each correct selection is worth one point.

A. Microsoft Hardware-Assisted Virtualization Detection Tool
B. Kitematic
C. BIOS-enabled virtualization
D. VirtualBox
E. Windows 10 64-bit Professional

Answer: CE

Explanation:
C: Make sure your Windows system supports Hardware Virtualization Technology and that virtualization is enabled.
Ensure that hardware virtualization support is turned on in the BIOS settings. For example:

E: To run Docker, your machine must have a 64-bit operating system running Windows 7 or higher. References:
https://docs.docker.com/toolbox/toolbox_install_windows/ https://blogs.technet.microsoft.com/canitpro/2015/09/08/step-by-step-enabling-hyper-v-for-use-on-windows-10/

NEW QUESTION 68
- (Exam Topic 3)
You have a Python script that executes a pipeline. The script includes the following code:
from azureml.core import Experiment
pipeline_run = Experiment(ws, 'pipeline_test').submit(pipeline) You want to test the pipeline before deploying the script.
You need to display the pipeline run details written to the STDOUT output when the pipeline completes. Which code segment should you add to the test script?

A. pipeline_run.get.metrics()
B. pipeline_run.wait_for_completion(show_output=True)
C. pipeline_param = PipelineParameter(name="stdout", default_value="console")
D. pipeline_run.get_status()

Answer: B

Explanation:
wait_for_completion: Wait for the completion of this run. Returns the status object after the wait. Syntax: wait_for_completion(show_output=False, wait_post_processing=False, raise_on_error=True) Parameter: show_output
Indicates whether to show the run output on sys.stdout.

NEW QUESTION 70
- (Exam Topic 3)
You are planning to register a trained model in an Azure Machine Learning workspace.
You must store additional metadata about the model in a key-value format. You must be able to add new metadata and modify or delete metadata after creation.
You need to register the model. Which parameter should you use?

A. description
B. model_framework
C. cags
D. properties

Answer: D

Explanation:
azureml.core.Model.properties:
Dictionary of key value properties for the Model. These properties cannot be changed after registration, however new key value pairs can be added.
Reference:
https://docs.microsoft.com/en-us/python/api/azureml-core/azureml.core.model.model

NEW QUESTION 75
- (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.

◢ SMOTE

Label column

Selected columns:
All labels

Launch column selector

SMOTE percentage ▼

| 0 |
| 300 |
| 3000 |
| 4000 |

Number of nearest neighbors ▼

| 0 |
| 1 |
| 5 |
| 4000 |

Random seed

0

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 77
- (Exam Topic 3)

A biomedical research company plans to enroll people in an experimental medical treatment trial.

You create and train a binary classification model to support selection and admission of patients to the trial. The model includes the following features: Age, Gender, and Ethnicity.

The model returns different performance metrics for people from different ethnic groups.

You need to use Fairlearn to mitigate and minimize disparities for each category in the Ethnicity feature. Which technique and constraint should you use? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.

Option	Value
Technique	⌄
	Grid search
	Outlier detection
	Dimensionality reduction
Constraint	⌄
	Demographic parity
	False-positive rate parity

A. Mastered
B. Not Mastered

Answer: A

Explanation:

Graphical user interface, text, application, chat or text message Description automatically generated

Box 1: Grid Search

Fairlearn open-source package provides postprocessing and reduction unfairness mitigation algorithms: ExponentiatedGradient, GridSearch, and ThresholdOptimizer.

Note: The Fairlearn open-source package provides postprocessing and reduction unfairness mitigation algorithms types:

≫ Reduction: These algorithms take a standard black-box machine learning estimator (e.g., a LightGBM model) and generate a set of retrained models using a sequence of re-weighted training datasets.

≫ Post-processing: These algorithms take an existing classifier and the sensitive feature as input.

Box 2: Demographic parity

The Fairlearn open-source package supports the following types of parity constraints: Demographic parity, Equalized odds, Equal opportunity, and Bounded group loss.

Reference:

https://docs.microsoft.com/en-us/azure/machine-learning/concept-fairness-ml

NEW QUESTION 80
- (Exam Topic 3)

You use the following code to define the steps for a pipeline: from azureml.core import Workspace, Experiment, Run from azureml.pipeline.core import Pipeline from azureml.pipeline.steps import PythonScriptStep ws = Workspace.from_config()

. . .

step1 = PythonScriptStep(name="step1", ...) step2 = PythonScriptsStep(name="step2", ...) pipeline_steps = [step1, step2]

You need to add code to run the steps.

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

NOTE: Each correct selection is worth one point.

A. experiment = Experiment(workspace=ws, name='pipeline-experiment')run = experiment.submit(config=pipeline_steps)
B. run = Run(pipeline_steps)
C. pipeline = Pipeline(workspace=ws, steps=pipeline_steps) experiment = Experiment(workspace=ws, name='pipeline-experiment')run = experiment.submit(pipeline)
D. pipeline = Pipeline(workspace=ws, steps=pipeline_steps)run = pipeline.submit(experiment_name='pipeline-experiment')

Answer: CD

Explanation:

After you define your steps, you build the pipeline by using some or all of those steps.

Build the pipeline. Example:

pipeline1 = Pipeline(workspace=ws, steps=[compare_models])

Submit the pipeline to be run

pipeline_run1 = Experiment(ws, 'Compare_Models_Exp').submit(pipeline1) Reference:

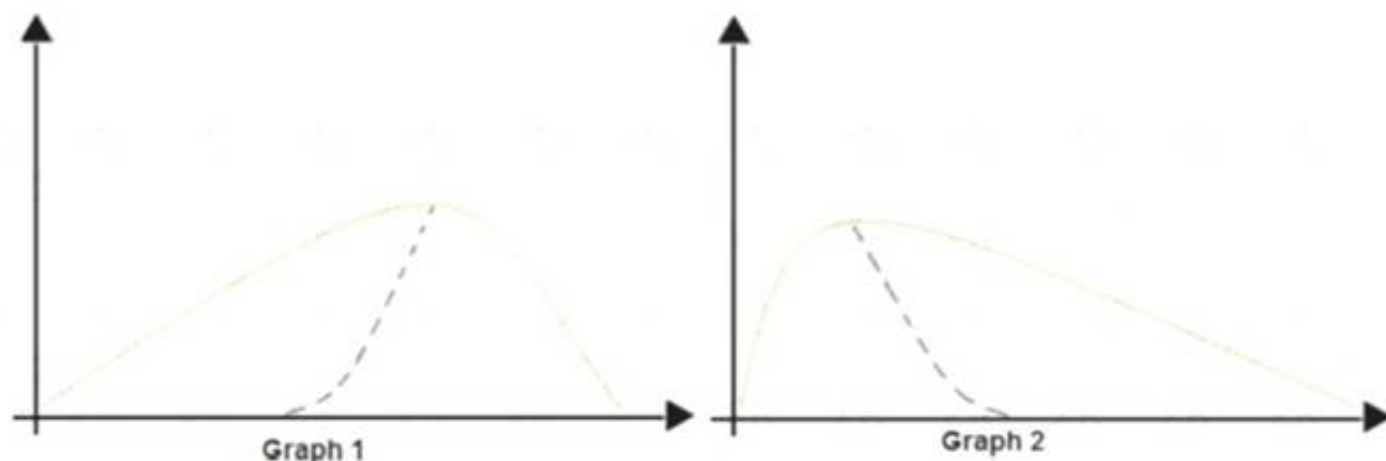https://docs.microsoft.com/en-us/azure/machine-learning/how-to-create-machine-learning-pipelines

NEW QUESTION 85
- (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.

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 87
- (Exam Topic 3)
You are building recurrent neural network to perform a binary classification.
The training loss, validation loss, training accuracy, and validation accuracy of each training epoch has been provided. You need to identify whether the
classification model is over fitted.
Which of the following is correct?

A. The training loss increases while the validation loss decreases when training the model.
B. The training loss decreases while the validation loss increases when training the model.
C. The training loss stays constant and the validation loss decreases when training the model.
D. 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.

Answer: B

Explanation:
An overfit model is one where performance on the train set is good and continues to improve, whereas performance on the validation set improves to a point and
then begins to degrade.
References:
https://machinelearningmastery.com/diagnose-overfitting-underfitting-lstm-models/

NEW QUESTION 92
- (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 95
- (Exam Topic 3)
You are using Azure Machine Learning to train machine learning models. You need a compute target on which to remotely run the training script. You run the following Python code:

```
from azureml.core.compute import ComputeTarget, AmlCompute
from azureml.core.compute_target import ComputeTargetException
the_cluster_name = "NewCompute"
config = AmlCompute.provisioning_configuration(vm_size='STANDARD_D2', max_nodes=3)
the_cluster = ComputeTarget.create(ws, the_cluster_name, config)
```

Answer Area

	Yes	No
The compute is created in the same region as the Machine Learning service workspace.	☑	○
The compute resource created by the code is displayed as a compute cluster in Azure Machine Learning studio.	○	○
The minimum number of nodes will be zero.	○	○

A. Mastered
B. Not Mastered

Answer: A

Explanation:
Box 1: Yes
The compute is created within your workspace region as a resource that can be shared with other users. Box 2: Yes
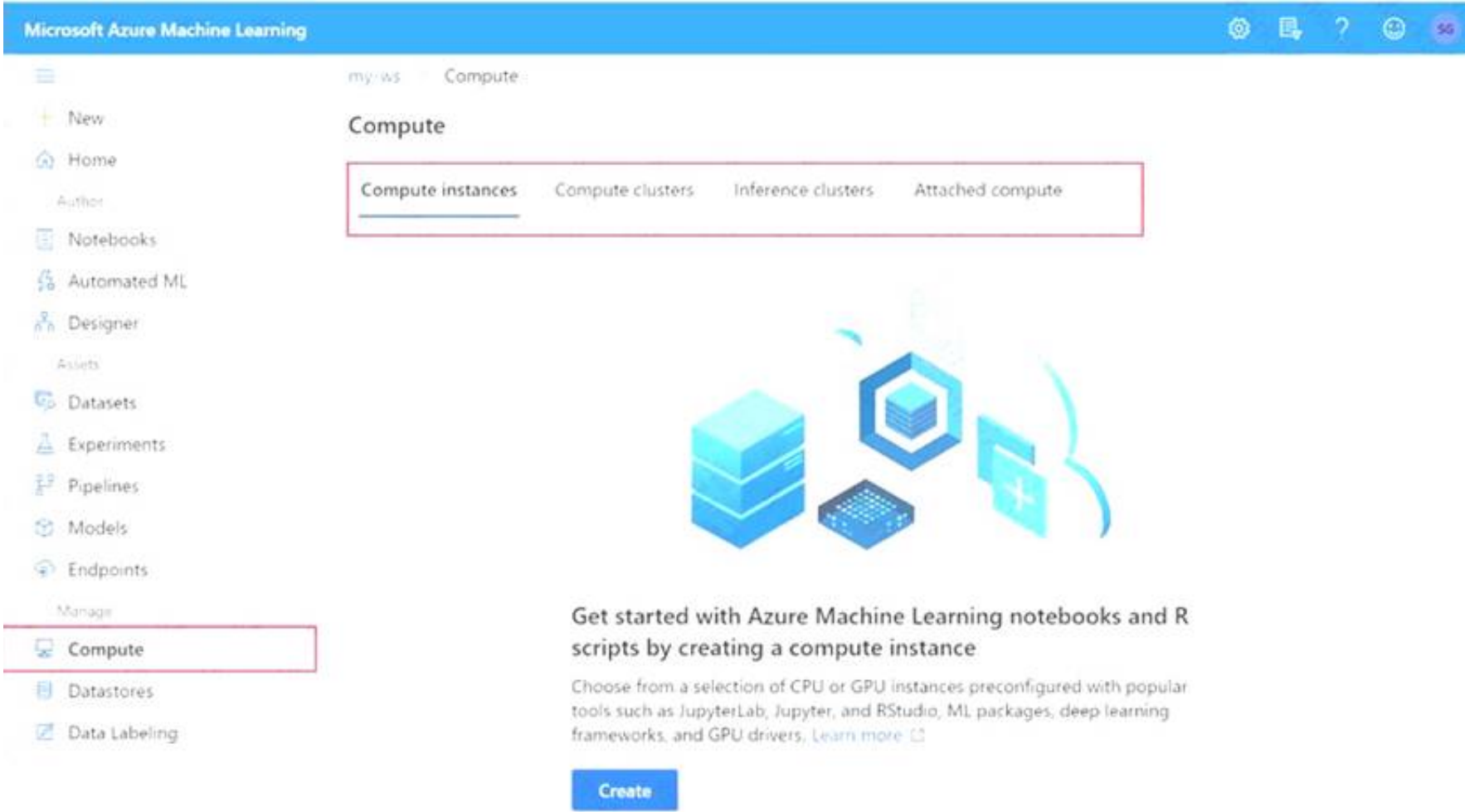It is displayed as a compute cluster. View compute targets
* 1. To see all compute targets for your workspace, use the following steps:
* 2. Navigate to Azure Machine Learning studio.
* 3. Under Manage, select Compute.
* 4. Select tabs at the top to show each type of compute target.

Box 3: Yes
min_nodes is not specified, so it defaults to 0. Reference:
https://docs.microsoft.com/en-us/python/api/azureml-core/azureml.core.compute.amlcompute.amlcomputeprovi https://docs.microsoft.com/en-us/azure/machine-
learning/how-to-create-attach-compute-studio

NEW QUESTION 100
- (Exam Topic 3)
You are building an experiment using the Azure Machine Learning designer.
You split a dataset into training and testing sets. You select the Two-Class Boosted Decision Tree as the algorithm.
You need to determine the Area Under the Curve (AUC) of the model.
Which three modules should you use in sequence? To answer, move the appropriate modules from the list of modules to the answer area and arrange them in the
correct order.

A. Mastered
B. Not Mastered

Answer: A

Explanation:
Step 1: Train Model
Two-Class Boosted Decision Tree
First, set up the boosted decision tree model.
* 1. Find the Two-Class Boosted Decision Tree module in the module palette and drag it onto the canvas.
* 2. Find the Train Model module, drag it onto the canvas, and then connect the output of the Two-Class Boosted Decision Tree module to the left input port of the
Train Model module.
The Two-Class Boosted Decision Tree module initializes the generic model, and Train Model uses training data to train the model.
* 3. Connect the left output of the left Execute R Script module to the right input port of the Train Model
module (in this tutorial you used the data coming from the left side of the Split Data module for training). This portion of the experiment now looks something like
this:

Step 2: Score Model
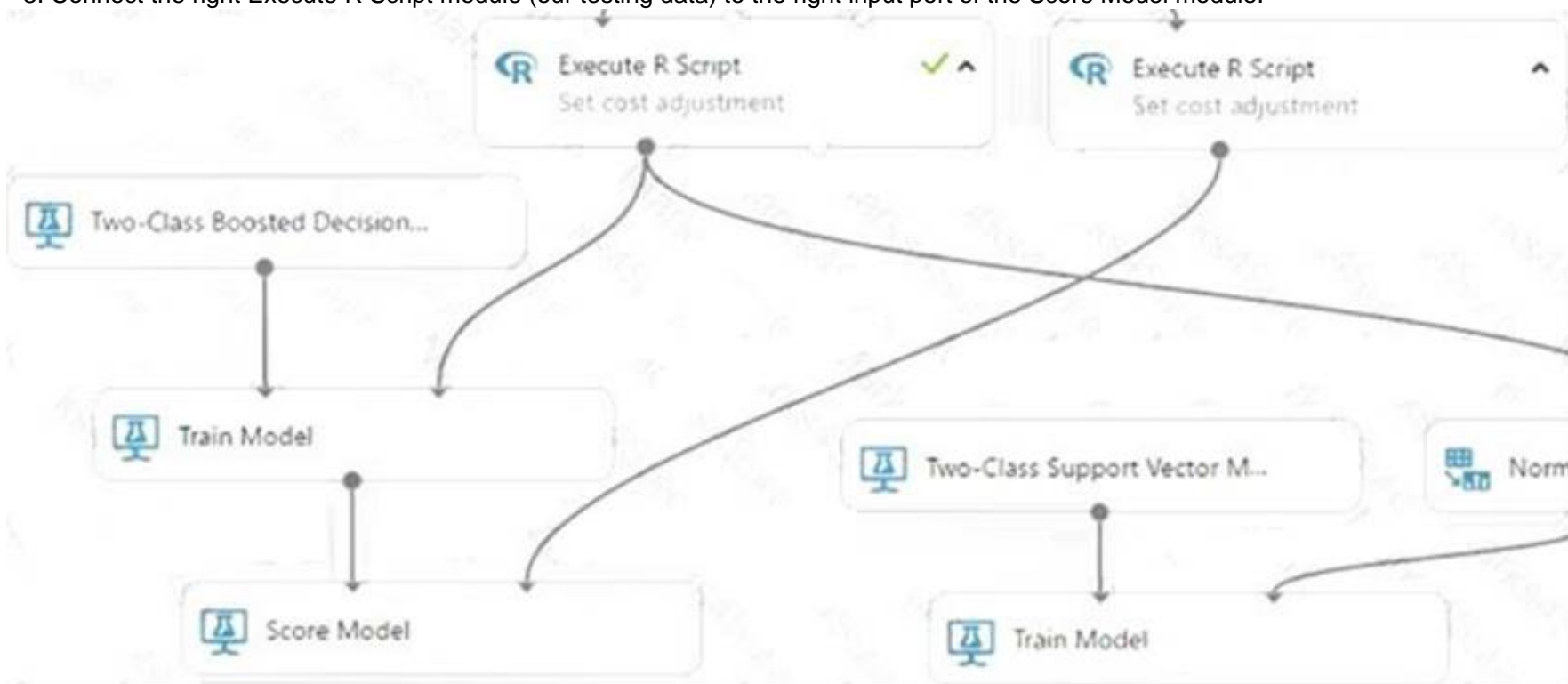Score and evaluate the models
You use the testing data that was separated out by the Split Data module to score our trained models. You can then compare the results of the two models to see which generated better results.
Add the Score Model modules
* 1. Find the Score Model module and drag it onto the canvas.
* 2. Connect the Train Model module that's connected to the Two-Class Boosted Decision Tree module to the left input port of the Score Model module.
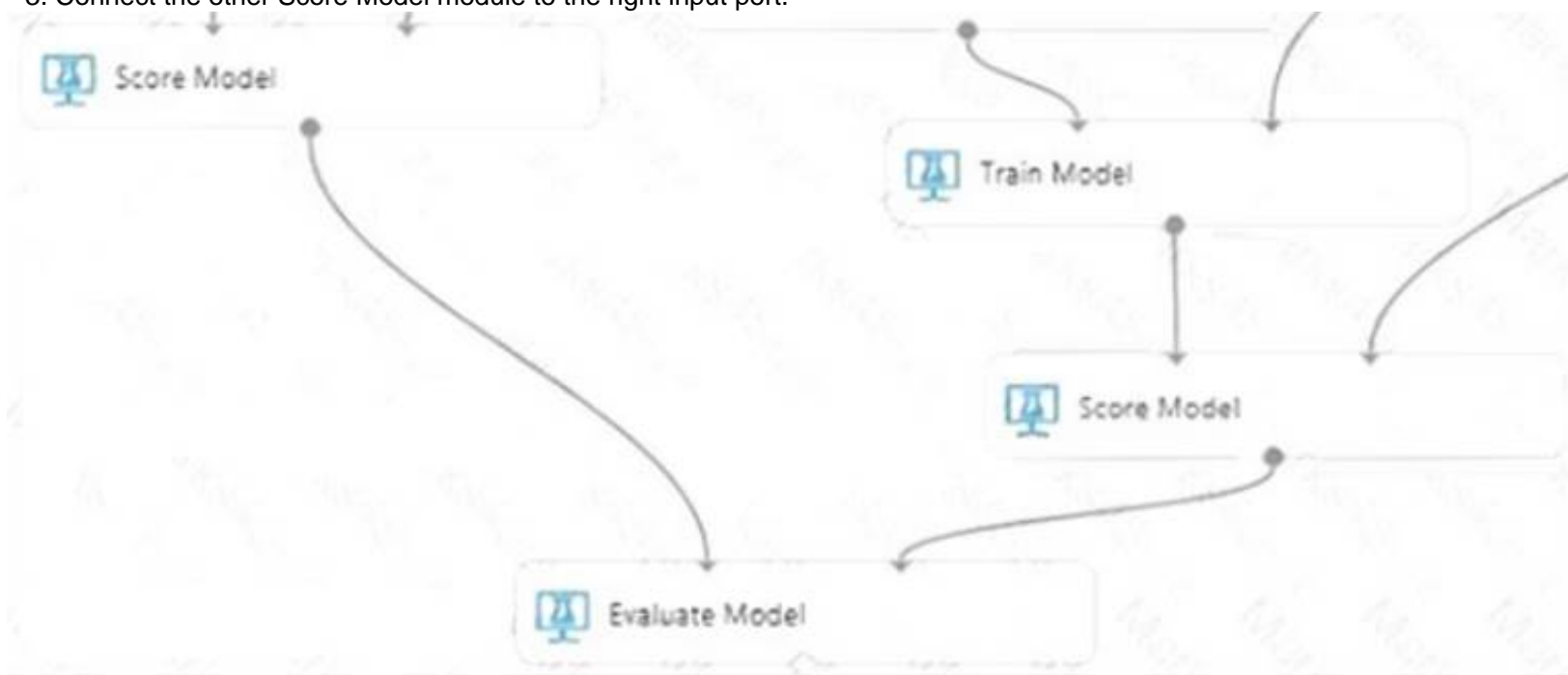* 3. Connect the right Execute R Script module (our testing data) to the right input port of the Score Model module.

Step 3: Evaluate Model
To evaluate the two scoring results and compare them, you use an Evaluate Model module.
* 1. Find the Evaluate Model module and drag it onto the canvas.
* 2. Connect the output port of the Score Model module associated with the boosted decision tree model to the left input port of the Evaluate Model module.
* 3. Connect the other Score Model module to the right input port.

NEW QUESTION 102
- (Exam Topic 3)
You plan to run a Python script as an Azure Machine Learning experiment. The script contains the following code:

import os, argparse, glob
from azureml.core import Run parser = argparse.ArgumentParser() parser.add_argument('--input-data',
type=str, dest='data_folder') args = parser.parse_args() data_path = args.data_folder
file_paths = glob.glob(data_path + "/*.jpg")
You must specify a file dataset as an input to the script. The dataset consists of multiple large image files and must be streamed directly from its source.
You need to write code to define a ScriptRunConfig object for the experiment and pass the ds dataset as an argument.
Which code segment should you use?

A. arguments = ['--input-data', ds.to_pandas_dataframe()]
B. arguments = ['--input-data', ds.as_mount()]
C. arguments = ['--data-data', ds]
D. arguments = ['--input-data', ds.as_download()]

Answer: A

Explanation:
If you have structured data not yet registered as a dataset, create a TabularDataset and use it directly in your training script for your local or remote experiment.
To load the TabularDataset to pandas DataFrame df = dataset.to_pandas_dataframe()
Note: TabularDataset represents data in a tabular format created by parsing the provided file or list of files. Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-train-with-datasets

NEW QUESTION 103
- (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 106
- (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 train and register a machine learning model.
You plan to deploy the model as a real-time web service. Applications must use key-based authentication to use the model.
You need to deploy the web service.
Solution:
Create an AksWebservice instance.
Set the value of the auth_enabled property to False.
Set the value of the token_auth_enabled property to True.
Deploy the model to the service. Does the solution meet the goal?

A. Yes
B. No

Answer: B

Explanation:
Instead use only auth_enabled = TRUE Note: Key-based authentication.
Web services deployed on AKS have key-based auth enabled by default. ACI-deployed services have
key-based auth disabled by default, but you can enable it by setting auth_enabled = TRUE when creating the ACI web service. The following is an example of creating an ACI deployment configuration with key-based auth enabled.
deployment_config <- aci_webservice_deployment_config(cpu_cores = 1, memory_gb = 1,
auth_enabled = TRUE) Reference:
https://azure.github.io/azureml-sdk-for-r/articles/deploying-models.html

NEW QUESTION 111
- (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 114
- (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.

Answer Area

	Yes	No
The estimator will look for the files it needs to run an experiment in the training-experiment directory of the local compute environment.	☑	○
The estimator will mount the local data-folder folder and make it available to the script through a parameter.		
The train.py script file will be created if it does not exist.	○	○

A. Mastered
B. Not Mastered

Answer: A

Explanation:

Answer Area

	Yes	No
The estimator will look for the files it needs to run an experiment in the training-experiment directory of the local compute environment.	⊡	○
The estimator will mount the local data-folder folder and make it available to the script through a parameter.	◉	○
The train.py script file will be created if it does not exist.	○	⊡

NEW QUESTION 115
- (Exam Topic 3)
You create a new Azure Databricks workspace.
You configure a new cluster for long-running tasks with mixed loads on the compute cluster as shown in the image below.

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

Code for each user runs as a separate process	▼
	Yes
	No

The number of workers is fixed for the entire duration of the job	▼
	Yes
	No

A. Mastered
B. Not Mastered

Answer: A

Explanation:
 Box 1: No
Running user code in separate processes is not possible in Scala. Box 2: No
Autoscaling is enabled. Minimum 2 workers, Maximum 8 workers. Reference:
https://docs.databricks.com/clusters/configure.html

NEW QUESTION 119
- (Exam Topic 3)
You plan to use a Data Science Virtual Machine (DSVM) with the open source deep learning frameworks Caffe2 and Theano. You need to select a pre configured
DSVM to support the framework.
What should you create?

A. Data Science Virtual Machine for Linux (CentOS)
B. Data Science Virtual Machine for Windows 2012
C. Data Science Virtual Machine for Windows 2016
D. Geo AI Data Science Virtual Machine with ArcGIS
E. Data Science Virtual Machine for Linux (Ubuntu)

Answer: E

NEW QUESTION 121
- (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 train and register a machine learning model.
You plan to deploy the model as a real-time web service. Applications must use key-based authentication to use the model.
You need to deploy the web service. Solution:
Create an AciWebservice instance.
Set the value of the ssl_enabled property to True.
Deploy the model to the service. Does the solution meet the goal?

A. Yes
B. No

Answer: B

Explanation:
Instead use only auth_enabled = TRUE Note: Key-based authentication.
Web services deployed on AKS have key-based auth enabled by default. ACI-deployed services have
key-based auth disabled by default, but you can enable it by setting auth_enabled = TRUE when creating the ACI web service. The following is an example of
creating an ACI deployment configuration with key-based auth enabled.
deployment_config <- aci_webservice_deployment_config(cpu_cores = 1,
memory_gb = 1, auth_enabled = TRUE) Reference:
https://azure.github.io/azureml-sdk-for-r/articles/deploying-models.html

NEW QUESTION 123
- (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 127
- (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 130
- (Exam Topic 3)
You are creating a machine learning model. You have a dataset that contains null rows.
You need to use the Clean Missing Data module in Azure Machine Learning Studio to identify and resolve the null and missing data in the dataset.
Which parameter should you use?

A. Replace with mean
B. Remove entire column
C. Remove entire row
D. Hot Deck

Answer: B

Explanation:
Remove entire row: Completely removes any row in the dataset that has one or more missing values. This is useful if the missing value can be considered randomly missing.
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/clean-missing-data

NEW QUESTION 134
- (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 train a classification model by using a logistic regression algorithm.
You must be able to explain the model's predictions by calculating the importance of each feature, both as an overall global relative importance value and as a measure of local importance for a specific set of predictions.
You need to create an explainer that you can use to retrieve the required global and local feature importance values.
Solution: Create a PFIExplainer. Does the solution meet the goal?

A. Yes
B. No

Answer: A

Explanation:
Permutation Feature Importance Explainer (PFI): Permutation Feature Importance is a technique used to explain classification and regression models. At a high level, the way it works is by randomly shuffling data one feature at a time for the entire dataset and calculating how much the performance metric of interest changes. The larger the change, the more important that feature is. PFI can explain the overall behavior of any underlying model but does not explain individual predictions.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-machine-learning-interpretability

NEW QUESTION 139
- (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 140
- (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

Which algorithm minimizes differences between actual and predicted values?

Answer choice
- Bayesian Linear Regression
- Neural Network Regression
- Boosted Decision Tree Regression
- Linear Regression
- 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.
- Set the Decrease learning rate option to True.
- Set the Create trainer mode option to Parameter Range.
- Increase the number of epochs.
- 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 142
- (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 143
- (Exam Topic 3)
You are creating a new Azure Machine Learning pipeline using the designer.
The pipeline must train a model using data in a comma-separated values (CSV) file that is published on a website. You have not created a dataset for this file.
You need to ingest the data from the CSV file into the designer pipeline using the minimal administrative effort.
Which module should you add to the pipeline in Designer?

A. Convert to CSV
B. Enter Data Manually D
C. Import Data
D. Dataset

Answer: D

Explanation:
The preferred way to provide data to a pipeline is a Dataset object. The Dataset object points to data that lives in or is accessible from a datastore or at a Web URL. The Dataset class is abstract, so you will create an instance of either a FileDataset (referring to one or more files) or a TabularDataset that's created by from one or more files with delimited columns of data.
Example:
from azureml.core import Dataset
iris_tabular_dataset = Dataset.Tabular.from_delimited_files([(def_blob_store, 'train-dataset/iris.csv')]) Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-create-your-first-pipeline

NEW QUESTION 148
- (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.

Category | Variables

Categorical variables ▼

Gender, IsPlaySoccer
Gender, PrevExamMarks, Height, Weight
PrevExamMarks, Height, Weight
IsPlaySoccer

Continuous variables ▼

Gender, IsPlaySoccer
Gender, PrevExamMarks, Height, Weight
PrevExamMarks, Height, Weight
IsPlaySoccer

A. Mastered
B. Not Mastered

Answer: A

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

NEW QUESTION 153
- (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 158
- (Exam Topic 3)
You are a lead data scientist for a project that tracks the health and migration of birds. You create a multi-class image classification deep learning model that uses a set of labeled bird photographs collected by experts.
You have 100,000 photographs of birds. All photographs use the JPG format and are stored in an Azure blob container in an Azure subscription.
You need to access the bird photograph files in the Azure blob container from the Azure Machine Learning service workspace that will be used for deep learning model training. You must minimize data movement.
What should you do?

A. Create an Azure Data Lake store and move the bird photographs to the store.
B. Create an Azure Cosmos DB database and attach the Azure Blob containing bird photographs storage to the database.
C. Create and register a dataset by using TabularDataset class that references the Azure blob storage containing bird photographs.
D. Register the Azure blob storage containing the bird photographs as a datastore in Azure Machine Learning service.
E. Copy the bird photographs to the blob datastore that was created with your Azure Machine Learning service workspace.

Answer: D

Explanation:
We recommend creating a datastore for an Azure Blob container. When you create a workspace, an Azure blob container and an Azure file share are automatically registered to the workspace.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-access-data

NEW QUESTION 163
- (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 168
- (Exam Topic 3)
You create a binary classification model. The model is registered in an Azure Machine Learning workspace. You use the Azure Machine Learning Fairness SDK to assess the model fairness.
You develop a training script for the model on a local machine.
You need to load the model fairness metrics into Azure Machine Learning studio. What should you do?

A. Implement the download_dashboard_by_upload_id function
B. Implement the creace_group_metric_sec function
C. Implement the upload_dashboard_dictionary function
D. Upload the training script

Answer: C

Explanation:
import azureml.contrib.fairness package to perform the upload:
from azureml.contrib.fairness import upload_dashboard_dictionary, download_dashboard_by_upload_id Reference:

https://docs.microsoft.com/en-us/azure/machine-learning/how-to-machine-learning-fairness-aml

NEW QUESTION 169
- (Exam Topic 3)

```
You create the following config.json file.

{
    "workspace_name" : "ml-workspace"
}
```

You must use the Azure Machine Learning SDK to interact with data and experiments in the workspace. You need to configure the config.json file to connect to the workspace from the Python environment. Which two additional parameters must you add to the config.json file in order to connect to the workspace?
Each correct answer presents part of the solution.
NOTE: Each correct selection is worth one point.

A. subscription_Id
B. Key
C. resource_group
D. region
E. Login

Answer: AC

Explanation:
To use the same workspace in multiple environments, create a JSON configuration file. The configuration file saves your subscription (subscription_id), resource (resource_group), and workspace name so that it can be easily loaded.
The following sample shows how to create a workspace. from azureml.core import Workspace
ws = Workspace.create(name='myworkspace', subscription_id='<azure-subscription-id>', resource_group='myresourcegroup', create_resource_group=True,
location='eastus2'
)
Reference:
https://docs.microsoft.com/en-us/python/api/azureml-core/azureml.core.workspace.workspace

NEW QUESTION 173
- (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 176
- (Exam Topic 2)
You need to set up the Permutation Feature Importance module according to the model training requirements. Which properties should you select? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.

▲ Tune Model Hyperparameters

Specify parameter sweeping mode

| Random sweep | ▼ |

Maximum number of runs on random sweep

| 5 |

Random seed

| 0 |

Label column

Selected columns:
Column names: MedianValue

Launch column selector

Metric for measuring performance for classification

| ▼ |

F-score
Precision
Recall
Accuracy

Metric for measuring performance for regression

| ▼ |

Root of mean squared error
R-squared
Mean zero one error
Mean absolute error

A. Mastered
B. Not Mastered

Answer: A

Explanation:
Box 1: Accuracy
Scenario: You want to configure hyperparameters in the model learning process to speed the learning phase by using hyperparameters. In addition, this configuration should cancel the lowest performing runs at each evaluation interval, thereby directing effort and resources towards models that are more likely to be successful.
Box 2: R-Squared

NEW QUESTION 177
- (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.

Code segments

```
early_termination_policy =
TruncationSelectionPolicy(evaluation_interval=1,
truncation_percentage=20, delay_evaluation=5)
```

```
import TruncationSelectionPolicy
```

```
from azureml.train.hyperdrive
```

```
import BanditPolicy
```

```
early_termination_policy = BanditPolicy
(slack_factor = 0.1, evaluation_interval=1,
delay_evaluation=5)
```

Answer Area

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 181
- (Exam Topic 1)
You need to implement a model development strategy to determine a user's tendency to respond to an ad. Which technique should you use?

A. Use a Relative Expression Split module to partition the data based on centroid distance.
B. Use a Relative Expression Split module to partition the data based on distance travelled to the event.
C. Use a Split Rows module to partition the data based on distance travelled to the event.
D. Use a Split Rows module to partition the data based on centroid distance.

Answer: A

Explanation:
Split Data partitions the rows of a dataset into two distinct sets.
The Relative Expression Split option in the Split Data module of Azure Machine Learning Studio is helpful
when you need to divide a dataset into training and testing datasets using a numerical expression.
Relative Expression Split: Use this option whenever you want to apply a condition to a number column. The number could be a date/time field, a column containing age or dollar amounts, or even a percentage. For example, you might want to divide your data set depending on the cost of the items, group people by age ranges, or separate data by a calendar date.
Scenario:
Local market segmentation models will be applied before determining a user's propensity to respond to an advertisement.
The distribution of features across training and production data are not consistent References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/split-data

NEW QUESTION 185
- (Exam Topic 1)
You need to implement a scaling strategy for the local penalty detection data.
Which normalization type should you use?

A. Streaming
B. Weight
C. Batch
D. Cosine

Answer: C

Explanation:
Post batch normalization statistics (PBN) is the Microsoft Cognitive Toolkit (CNTK) version of how to evaluate the population mean and variance of Batch Normalization which could be used in inference Original Paper.
In CNTK, custom networks are defined using the BrainScriptNetworkBuilder and described in the CNTK network description language "BrainScript."
Scenario:
Local penalty detection models must be written by using BrainScript. References:

https://docs.microsoft.com/en-us/cognitive-toolkit/post-batch-normalization-statistics

NEW QUESTION 190
- (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.

Actions

| Define a cross-entropy function activation. |
| Add cost functions for each target state. |
| Evaluate the classification error metric. |
| Evaluate the distance error metric. |
| Add cost functions for each component metric. |
| Define a sigmoid loss function activation. |

Answer Area

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 195
- (Exam Topic 1)
You need to implement a new cost factor scenario for the ad response models as illustrated in the performance curve exhibit.
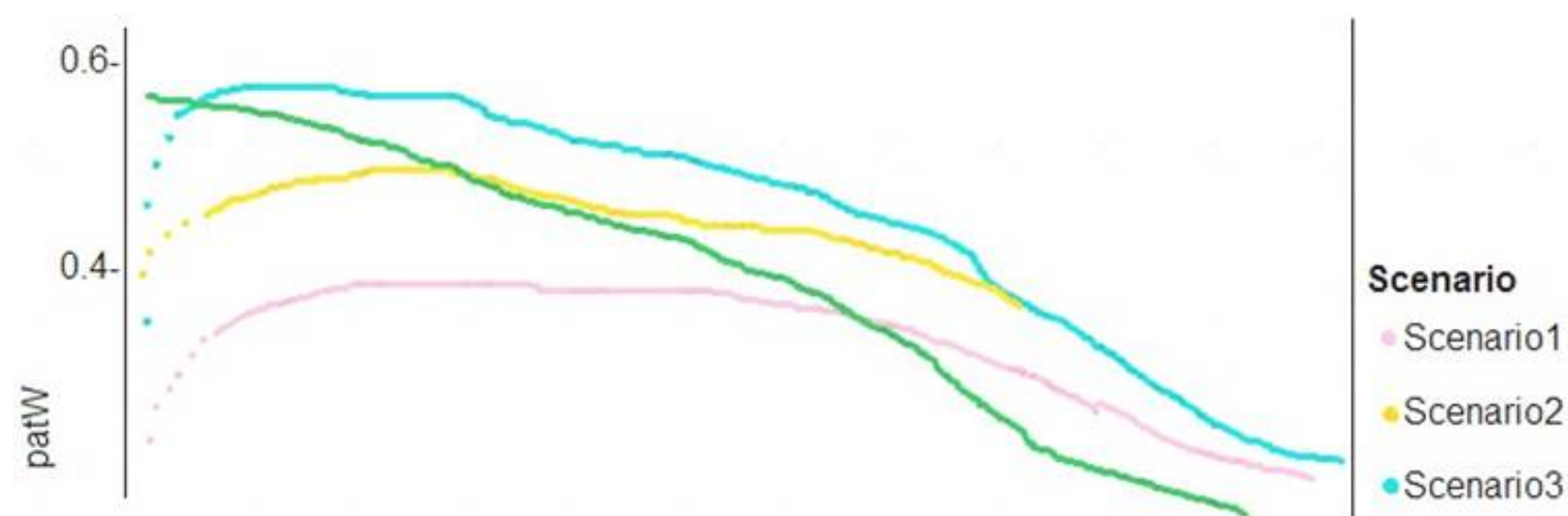Which technique should you use?

A. Set the threshold to 0.5 and retrain if weighted Kappa deviates +/- 5% from 0.45.
B. Set the threshold to 0.05 and retrain if weighted Kappa deviates +/- 5% from 0.5.
C. Set the threshold to 0.2 and retrain if weighted Kappa deviates +/- 5% from 0.6.
D. Set the threshold to 0.75 and retrain if weighted Kappa deviates +/- 5% from 0.15.

Answer: A

Explanation:
 Scenario:
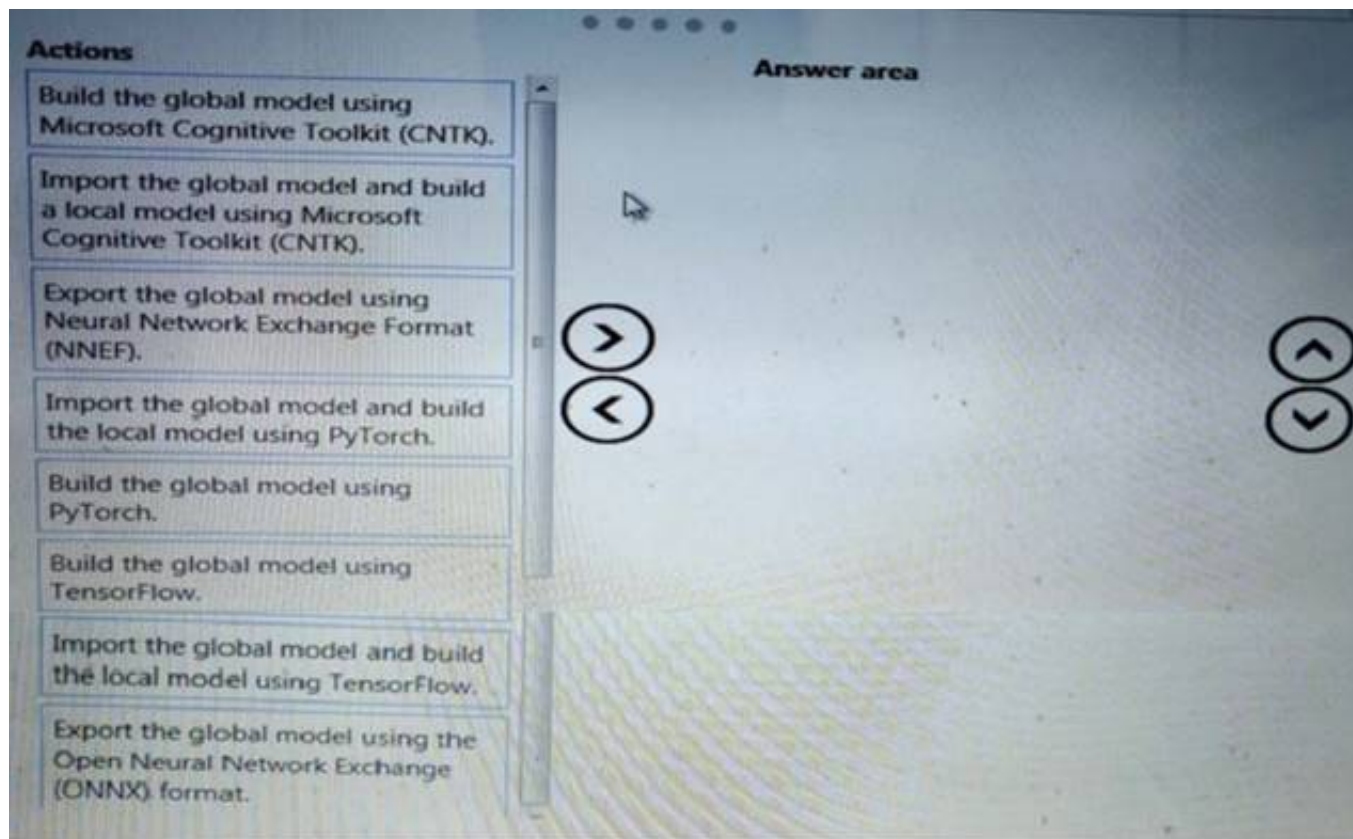Performance curves of current and proposed cost factor scenarios are shown in the following diagram:

The ad propensity model uses a cut threshold is 0.45 and retrains occur if weighted Kappa deviated from 0.1 +/- 5%.

NEW QUESTION 197
- (Exam Topic 1)
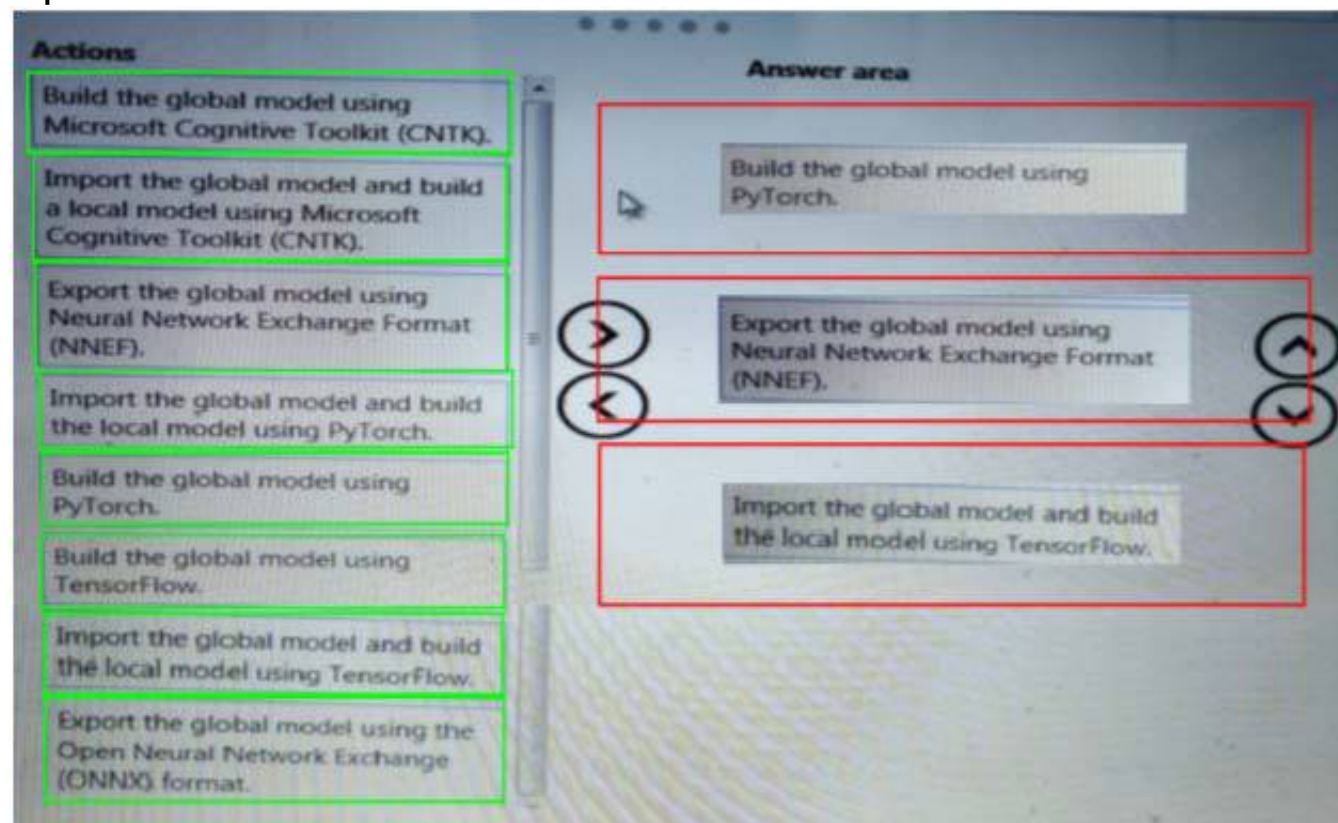You need to define a process for penalty event detection.
Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

A. Mastered
B. Not Mastered

Answer: A

Explanation:

NEW QUESTION 201
......

Thank You for Trying Our Product

* 100% Pass or Money Back

    All our products come with a 90-day Money Back Guarantee.

* One year free update

    You can enjoy free update one year. 24x7 online support.

* Trusted by Millions

    We currently serve more than 30,000,000 customers.

* Shop Securely

    All transactions are protected by VeriSign!

100% Pass Your DP-100 Exam with Our Prep Materials Via below:

https://www.certleader.com/DP-100-dumps.html