Microsoft.DP-100.v2026-06-25.q212

Exam Code:	DP-100
Exam Name:	Designing and Implementing a Data Science Solution on Azure
Certification Provider:	Microsoft
Free Question Number:	212
Version:	v2026-06-25
# of views:	255
# of Questions views:	2120
https://www.freepdfdumps.com/Microsoft.DP-100.v2026-06-25.q212.html	

NEW QUESTION: 1
You need to resolve the local machine learning pipeline performance issue. What should you do?
A. Increase Graphic Processing Units (GPUs).
B. Increase the learning rate.
C. Increase the training iterations,
D. Increase Central Processing Units (CPUs).
Answer: A (LEAVE A REPLY)

NEW QUESTION: 2
You create an Azure Machine Learning workspace.
You need to detect data drift between a baseline dataset and a subsequent target dataset by using the DataDriftDetector class.
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 azureml.core import Workspace, Dataset
from datetime import datetime

ws = Workspace.from_config()
dset = Dataset.get_by_name(ws, 'target')
baseline = target.time_before(datetime(2021, 2, 1))
features = ['windAngle', 'windSpeed', 'temperature', 'stationName']

monitor = DataDriftDetector.                    (ws, 'drift-monitor', baseline,
```

- backfill
- create_from_datasets
- create_from_model

```
target, compute_target='cpu-cluster', frequency='Week', feature_list=None,
drift_threshold=.6, latency=24)

monitor = DataDriftDetector.get_by_name(ws, 'drift-monitor')
monitor = monitor.update(feature_list=features)
complete = monitor.                    (datetime(2021, 1, 1), datetime.today())
```

- backfill
- list
- update

Answer:

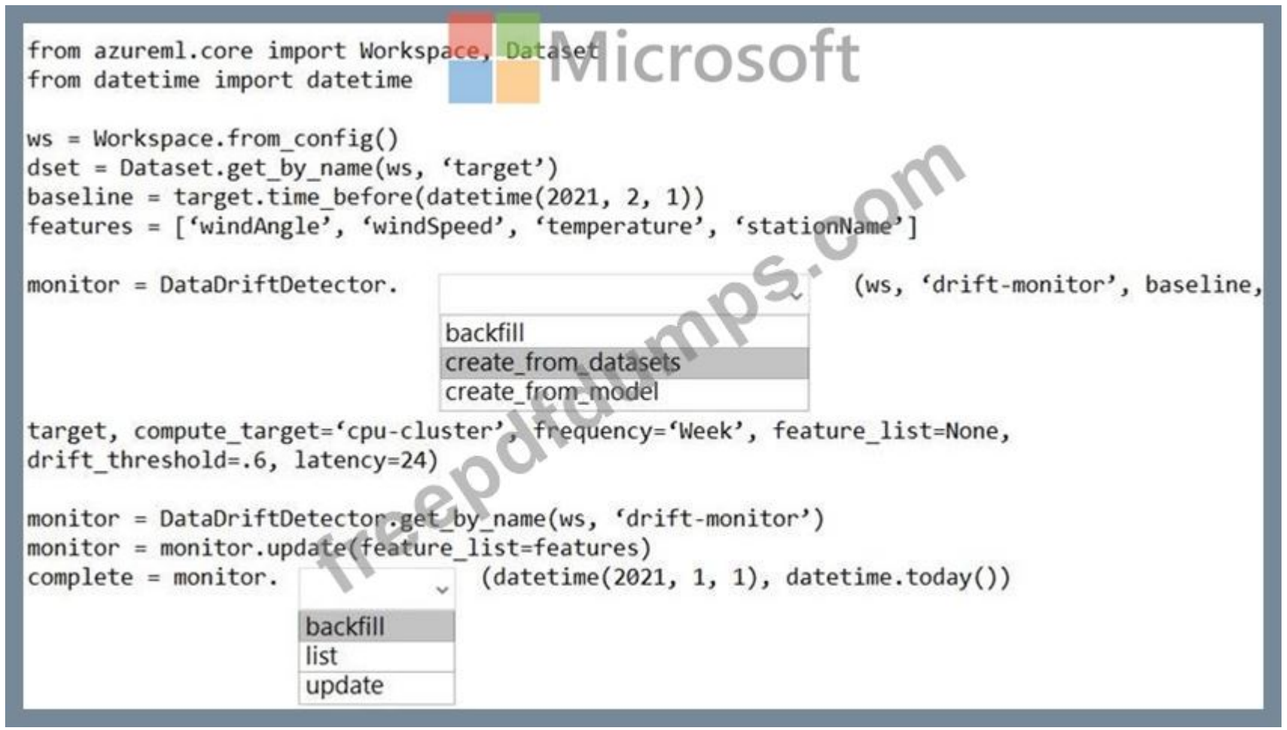
```
from azureml.core import Workspace, Dataset
from datetime import datetime

ws = Workspace.from_config()
dset = Dataset.get_by_name(ws, 'target')
baseline = target.time_before(datetime(2021, 2, 1))
features = ['windAngle', 'windSpeed', 'temperature', 'stationName']

monitor = DataDriftDetector.                    (ws, 'drift-monitor', baseline,
```

- backfill
- create_from_datasets
- create_from_model

```
target, compute_target='cpu-cluster', frequency='Week', feature_list=None,
drift_threshold=.6, latency=24)

monitor = DataDriftDetector.get_by_name(ws, 'drift-monitor')
monitor = monitor.update(feature_list=features)
complete = monitor.                    (datetime(2021, 1, 1), datetime.today())
```

- backfill
- list
- update

Explanation:



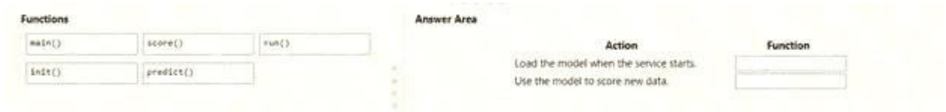
Box 1: create_from_datasets
The create_from_datasets method creates a new DataDriftDetector object from a baseline tabular dataset and a target time series dataset.
Box 2: backfill
The backfill method runs a backfill job over a given specified start and end date.
Syntax: backfill(start_date, end_date, compute_target=None, create_compute_target=False) Reference:
https://docs.microsoft.com/en-us/python/api/azureml-datadrift/azureml.datadrift.datadriftdetector(class)

NEW QUESTION: 3
You manage an Azure AI Foundry project in your subscription. You deploy a gpt-4o model. You must test the model before you use it in an existing front-end application. You need to adjust the parameters to get more creative responses.
Solution: Increase Temperature.
Does the solution meet the goal?
A. No
B. Yes
Answer: B (LEAVE A REPLY)

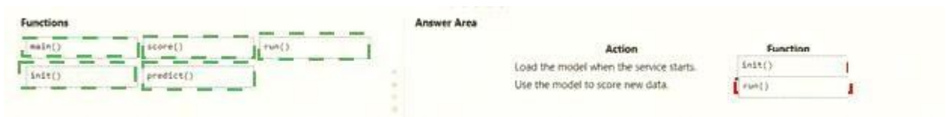
NEW QUESTION: 4
You have fine-tuned an Azure OpenAI Service model by using the Azure Ai Foundry portal. The fine-tuned model is overfitting.
You plan to correct overfitting by fine-tuning the model again
You need to modify the default value of a fine-tuning task parameter to minimize the possibility of overfitting. Which modification should you apply?
A. Decrease the batch_size.

B. Decrease the leaming_rate_multiplier.
C. increase the batch_size.
D. Increase the learning_rate_multiplier.
Answer: B (LEAVE A REPLY)

NEW QUESTION: 5
You use Azure Machine Learning to deploy a model as a real-time web service.
You need to create an entry script for the service that ensures that the model is loaded when the service starts and is used to score new data as it is received.
Which functions should you include in the script? To answer, drag the appropriate functions to the correct actions. Each function may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content NOTE: Each correct selection is worth one point.



Answer:



Explanation:

Box 1: init()
The entry script has only two required functions, init() and run(data). These functions are used to initialize the service at startup and run the model using request data passed in by a client. The rest of the script handles loading and running the model(s).
Box 2: run()
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-deploy-existing-model

NEW QUESTION: 6
You have the following Azure subscriptions and Azure Machine Learning service workspaces:

Subscription	Workspace	Comment
385bdfe5-4cef-4ad4-b977-3f86d92727c9	ml-default	This is the default subscription.
5a5891d1-557a-4234-9b83-2e90412b1068	ml-project	The information required to uniquely identify this workspace is stored in the file config.json in the same folder as the Python script.

You need to obtain a reference to the mi-protect workspace
Solution: Run the following Python code.

```
from azure.ai.ml import MLClient
ws = MLClient.workspaces.get("ml-project")
```

Does the solution meet the goal?
A. No
B. Yes
Answer: B (LEAVE A REPLY)

NEW QUESTION: 7
You create a binary classification model. You use the Fairlearn package to assess model fairness. You must eliminate the need to retrain the model. You need to implement the Fair learn package. Which algorithm should you use?
A. fair Icarn.postprocessing.ThresholdOplimizer
B. fatrlearn.reductions.GridSearch
C. fairlearn.reductions.ExponentiatedGradient
D. fairlearn.preprocessing.CorrelationRemover
Answer: D (LEAVE A REPLY)

NEW QUESTION: 8
You create a multi-class image classification deep learning model that uses the PyTorch deep learning framework.
You must configure Azure Machine Learning Hyperdrive to optimize the hyperparameters for the classification model.
You need to define a primary metric to determine the hyperparameter values that result in the model with the best accuracy score.
Which three actions must you perform? Each correct answer presents part of the solution.
NOTE: Each correct selection is worth one point.
A. Set the primary_metric_goal of the estimator used to run the bird_classifier_train.py script to maximize.
B. Add code to the bird_classifier_train.py script to calculate the validation loss of the model and log it as a float value with the key loss.
C. Set the primary_metric_goal of the estimator used to run the bird_classifier_train.py script to minimize.
D. Set the primary_metric_name of the estimator used to run the bird_classifier_train.py script to accuracy.
E. Set the primary_metric_name of the estimator used to run the bird_classifier_train.py script to loss.
F. Add code to the bird_classifier_train.py script to calculate the validation accuracy of the model and log it as a float value with the key accuracy.
Answer: A,D,F (LEAVE A REPLY)
AD:
primary_metric_name="accuracy",
primary_metric_goal=PrimaryMetricGoal.MAXIMIZE
Optimize the runs to maximize "accuracy". Make sure to log this value in your training script.
Note:
primary_metric_name: The name of the primary metric to optimize. The name of the primary metric needs to exactly match the name of the metric logged by the training script.
primary_metric_goal: It can be either PrimaryMetricGoal.MAXIMIZE or PrimaryMetricGoal.MINIMIZE and determines whether the primary metric will be maximized or minimized when evaluating the runs.
F: The training script calculates the val_accuracy and logs it as "accuracy", which is used as the primary metric.

NEW QUESTION: 9

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

run = Run.get_context()
data = pd.read_csv('./data.csv')
rows = (len(data))
# record row_count metric here
...
```


You need to record the row count as a metric named row_count that can be returned using the get_metrics method of the Run object after the experiment run completes. Which code should you use?

A. run.upload_file('row_count', './data.csv')

B. run.log('row_count', rows)

C. run.tag('row_count', rows)

D. run.log_table('row_count', rows)

E. run.log_row('row_count', rows)

Answer: B (LEAVE A REPLY)

Log a numerical or string value to the run with the given name using log(name, value, description= ' ').

Logging a metric to a run causes that metric to be stored in the run record in the experiment. You can log the same metric multiple times within a run, the result being considered a vector of that metric.

Example: run.log(" accuracy " , 0.95)

Reference:

https://docs.microsoft.com/en-us/python/api/azureml-core/azureml.core.run

NEW QUESTION: 10

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

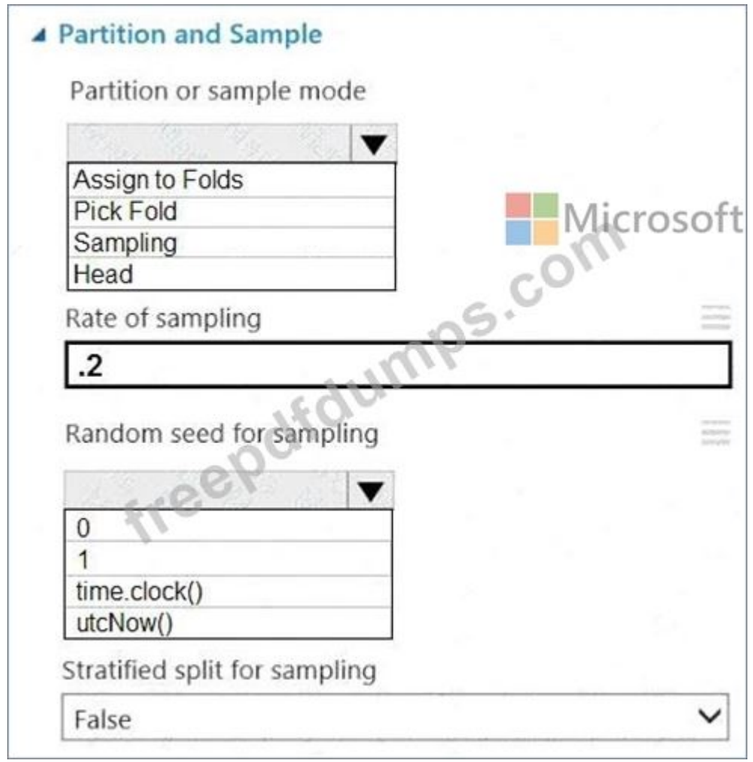
You must create a subset of the data for testing purposes using a random sampling seed based on the system clock.

You add the Partition and Sample module to your experiment.

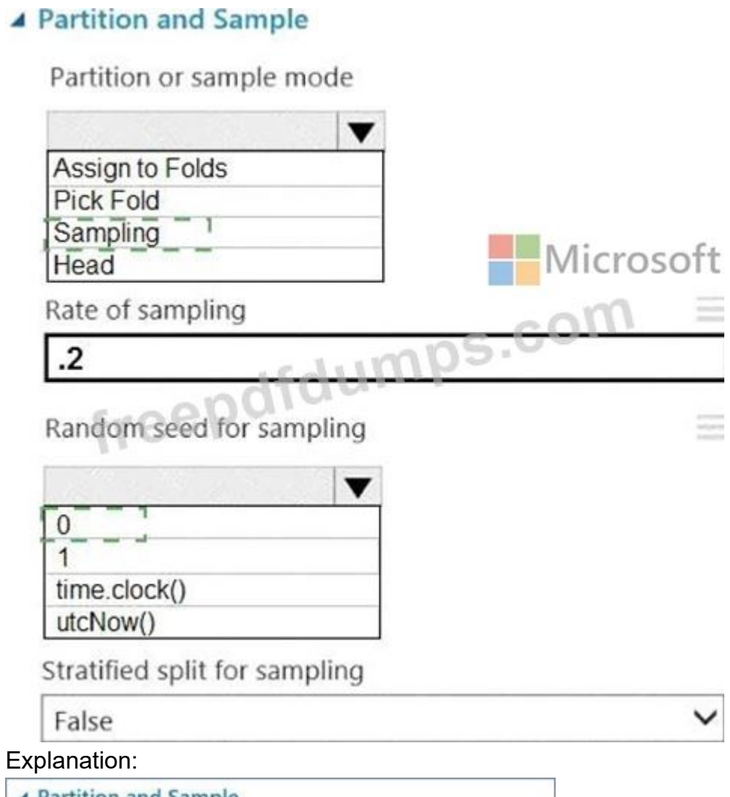
You need to select the properties for the module.

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

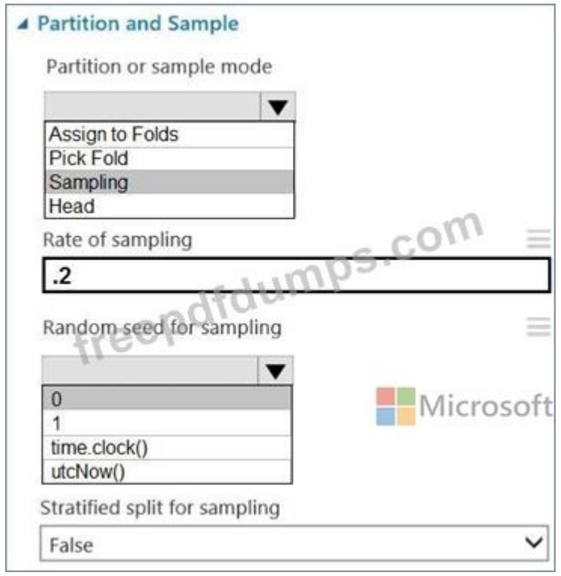
NOTE: Each correct selection is worth one point.


Partition and Sample
Partition or sample mode
Assign to Folds
Pick Fold
Sampling
Head
Rate of sampling
.2
Random seed for sampling
0
1
time.clock()
utcNow()
Stratified split for sampling
False

Answer:



Explanation:



Box 1: Sampling

Create a sample of data

This option supports simple random sampling or stratified random sampling. This is useful if you want to create a smaller representative sample dataset for testing.

1. Add the Partition and Sample module to your experiment in Studio, and connect the dataset.

2. Partition or sample mode: Set this to Sampling.
3. Rate of sampling. See box 2 below.
Box 2: 0
3. Rate of sampling. Random seed for sampling: Optionally, type an integer to use as a seed value.
This option is important if you want the rows to be divided the same way every time. The default value is 0, meaning that a starting seed is generated based on the system clock. This can lead to slightly different results each time you run the experiment.
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/partition-and-sample

NEW QUESTION: 11
You are developing a machine learning model.
You must inference the machine learning model for testing.
You need to use a minimal cost compute target
Which two compute targets should you use? Each correct answer presents a complete solution.
NOTE: Each correct selection is worth one point
A. Remote VM
B. Azure Machine Learning Kubernetes
C. Local web service
D. Azure Databricks
E. Azure Container Instances
Answer: C,E (LEAVE A REPLY)

NEW QUESTION: 12
You have a deployment of an Azure OpenAI Service base model.
You plan to fine-tune the model.
You need to prepare a file that contains training data.
Which file format should you use?
A. TSV
B. CSV
C. JSON
D. JSONL
Answer: (SHOW ANSWER)

NEW QUESTION: 13
You create a training pipeline by using the Azure Machine Learning designer. You need to load data into a machine learning pipeline by using the Import Data component. Which two data sources could you use? Each correct answer presents a complete solution.
NOTE: Each correct selection is worth one point
A. URL via HTTP
B. Azure Data Lake Storage Gen2
C. Azure Blob storage container through a registered datastore
D. Azure SQL Database

E. Registered dataset
Answer: A,C (LEAVE A REPLY)

NEW QUESTION: 14
You have an Azure AI Foundry project.
You plan to use the Azure AI Foundry portal to fine-tune a base Azure OpenAI Service model that can accept both text and images as input You need to choose the suitable model. Which model should you choose?
A. gpt-35-turbo
B. davinci-002
C. gpt-4o
D. gpt-4
Answer: (SHOW ANSWER)

NEW QUESTION: 15
You are a data scientist building a deep convolutional neural network (CNN) for image classification.
The CNN model you built shows signs of overfitting.
You need to reduce overfitting and converge the model to an optimal fit.
Which two actions should you perform? Each correct answer presents a complete solution.
NOTE: Each correct selection is worth one point.
A. Reduce the amount of training data.
B. Add an additional dense layer with 64 input units
C. Add L1/L2 regularization.
D. Use training data augmentation
E. Add an additional dense layer with 512 input units.
Answer: A,C (LEAVE A REPLY)
References:
https://machinelearningmastery.com/how-to-reduce-overfitting-in-deep-learning-with-weight-regularization/
https://en.wikipedia.org/wiki/Convolutional_neural_network

NEW QUESTION: 16
You develop and train a machine learning model to predict fraudulent transactions for a hotel booking website.
Traffic to the site varies considerably. The site experiences heavy traffic on Monday and Friday and much lower traffic on other days. Holidays are also high web traffic days. You need to deploy the model as an Azure Machine Learning real-time web service endpoint on compute that can dynamically scale up and down to support demand. Which deployment compute option should you use?
A. attached Azure Databricks cluster
B. Azure Container Instance (ACI)
C. Azure Kubernetes Service (AKS) inference cluster
D. Azure Machine Learning Compute Instance
E. attached virtual machine in a different region
Answer: E (LEAVE A REPLY)
Azure Machine Learning compute cluster is a managed-compute infrastructure that allows you to easily create a single or multi-node compute. The compute is created within your workspace region as a resource that can be shared with other users in your workspace. The compute scales up automatically when a job is submitted, and can be put in an Azure Virtual Network.
Reference:

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

NEW QUESTION: 17
You have a Jupyter Notebook that contains Python code that is used to train a model.
You must create a Python script for the production deployment. The solution must minimize code maintenance.
Which two actions should you perform? Each correct answer presents part of the solution.
NOTE: Each correct selection is worth one point.
A. Refactor the Jupyter Notebook code into functions
B. Save each function to a separate Python file
C. Define a main() function in the Python script
D. Remove all comments and functions from the Python script
Answer: (SHOW ANSWER)
Reference:
https://www.guru99.com/learn-python-main-function-with-examples-understand-main.html
https://towardsdatascience.com/from-jupyter-notebook-to-deployment-a-straightforward-example-1838c203a437

NEW QUESTION: 18
You plan to use automated machine learning by using Azure Machine Learning Python SDK v2 to train a regression model. You have data that has features with missing values, and categorical features with few distinct values.
You need to control whether automated machine learning automatically imputes missing values and encode categorical features as part of the training task. Which enemy of the autumn package should you use?
A. ForecastHorizonMode
B. FeaturizationMode
C. RegressionPrimaryMetrics
D. RegressionModels
Answer: B (LEAVE A REPLY)

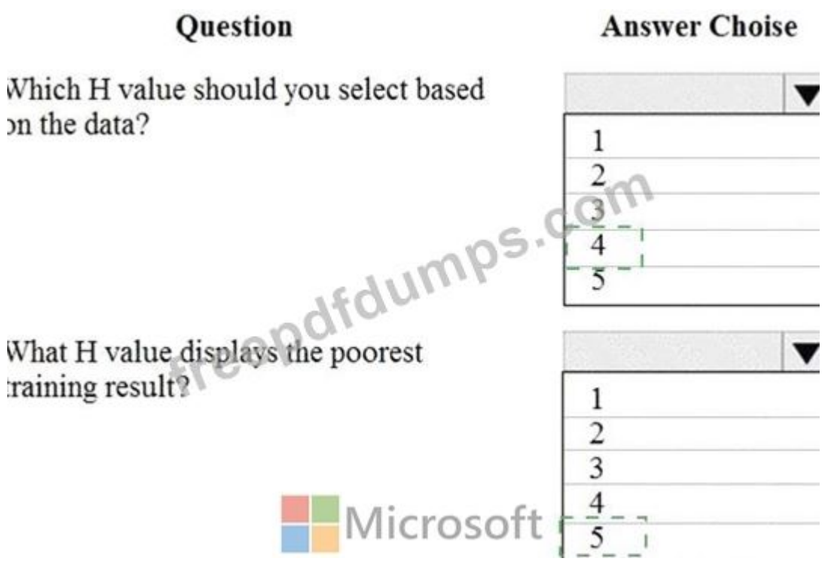
NEW QUESTION: 19
You are tuning a hyperparameter for an algorithm. The following table shows a data set with different hyperparameter, training error, and validation errors.

Hyperparameter (H)	Training error (TE)	Validation error (VE)
1	105	95
2	200	85
3	250	100
4	105	100
5	400	50

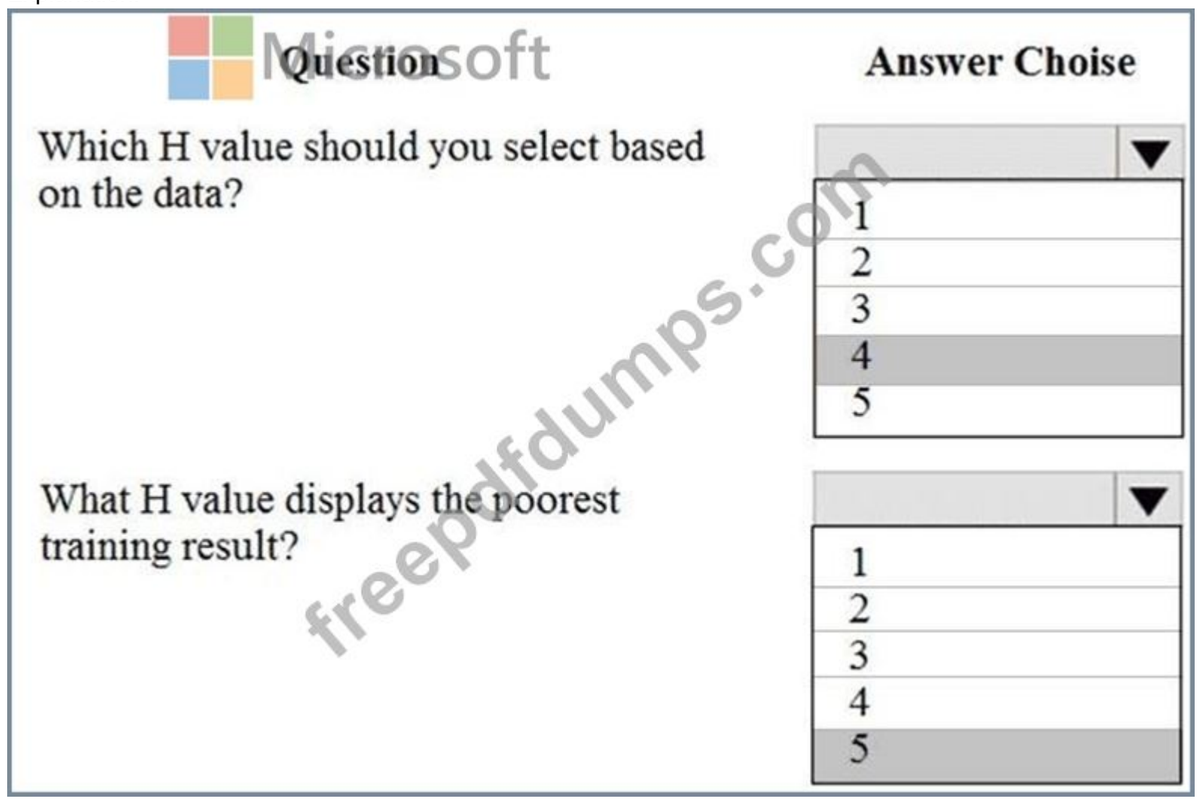
Use the drop-down menus to select the answer choice that answers each question based on the information presented in the graphic.

Question	Answer Choise
Which H value should you select based on the data?	1 2 3 4 5
What H value displays the poorest training result?	1 2 3 4 5

Answer:



Explanation:



Box 1: 4

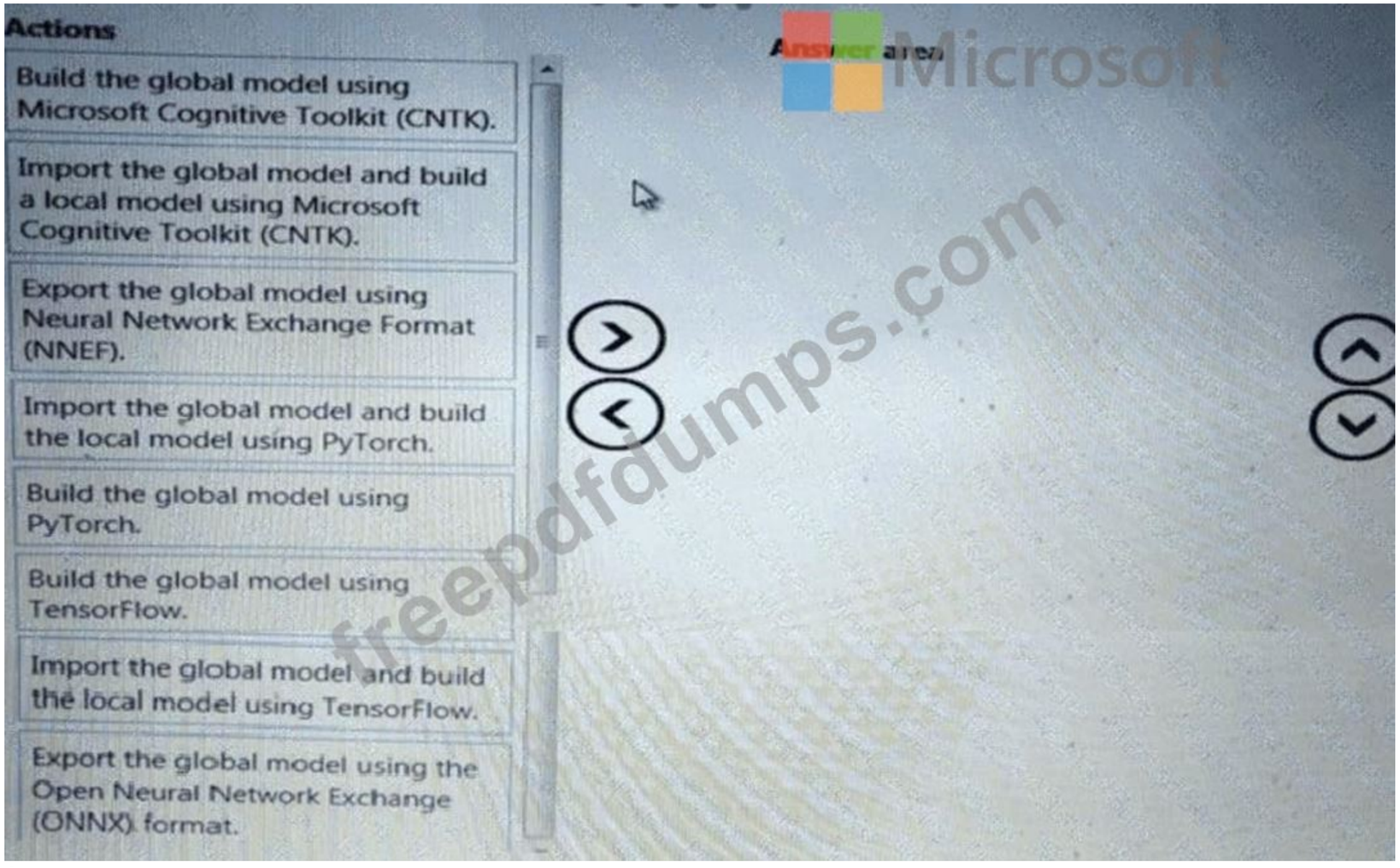
Choose the one which has lower training and validation error and also the closest match.

Minimize variance (difference between validation error and train error).

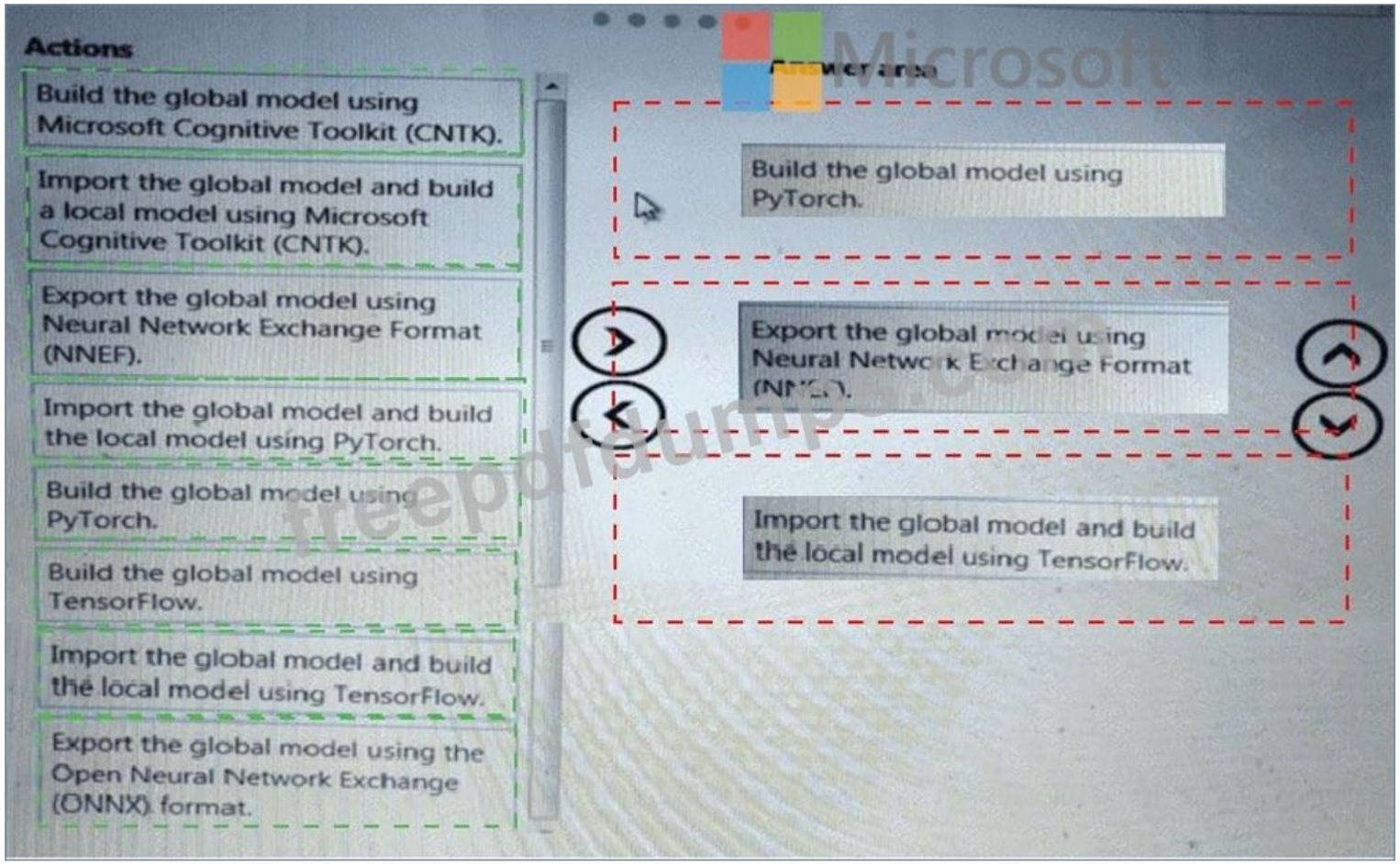
Box 2: 5
Minimize variance (difference between validation error and train error).
Reference:
https://medium.com/comet-ml/organizing-machine-learning-projects-project-management-guidelines-2d2b85651bbd

NEW QUESTION: 20
You use Azure Machine Learning studio to analyze an mltable data asset containing a decimal column named column1. You need to verify that the column1 values are normally distributed.
Which statistic should you use?
A. Mean
B. Max
C. Profile
D. Type
Answer: (SHOW ANSWER)

NEW QUESTION: 21
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.



Answer:



NEW QUESTION: 22

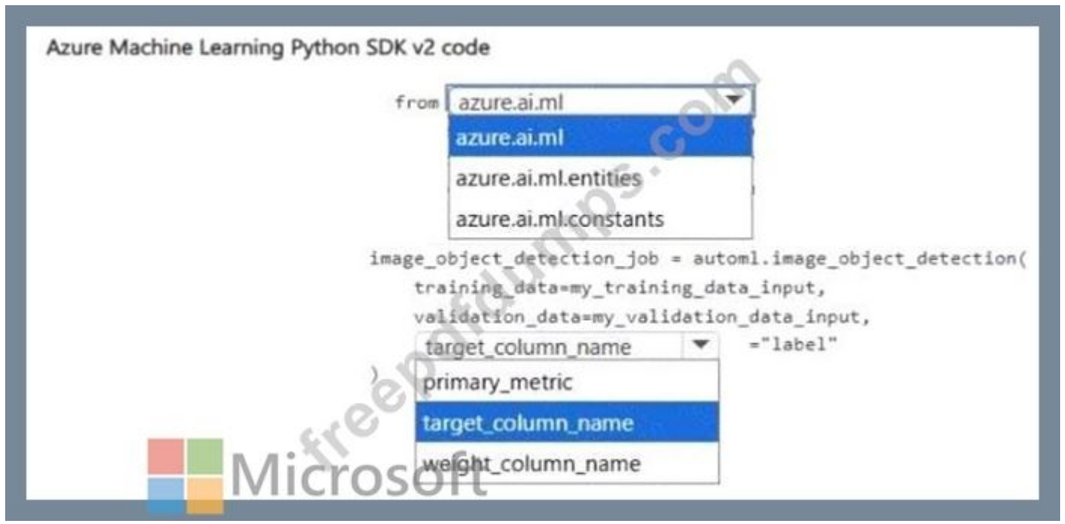
You have an Azure Machine Learning workspace named Workspaces

You plan to train an image object detection model by using Automated ML in Workspace1.

You need to complete the provided Azure Machine Learning Python SDK v2 code to start an image object detection job.

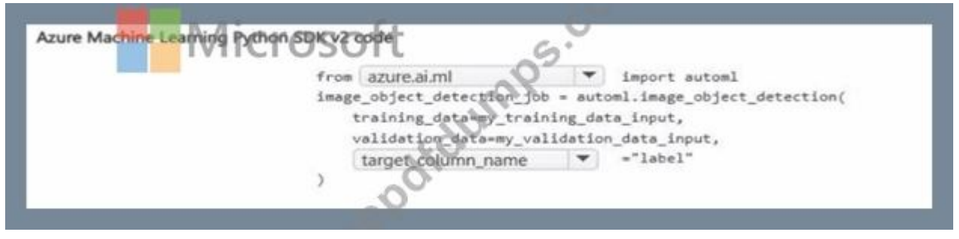
How should you complete the code? To answer, select the appropriate options in the answer area.

NOTE Each correct selection is worth one point.



Answer:

Explanation:



NEW QUESTION: 23

You manage an Azure At Foundry project.

You are implementing a RAG solution. The documents contain tables and images that must be broken into semantically relevant chunks.

You need to generate textual representations of images and tables to be used as chunks.

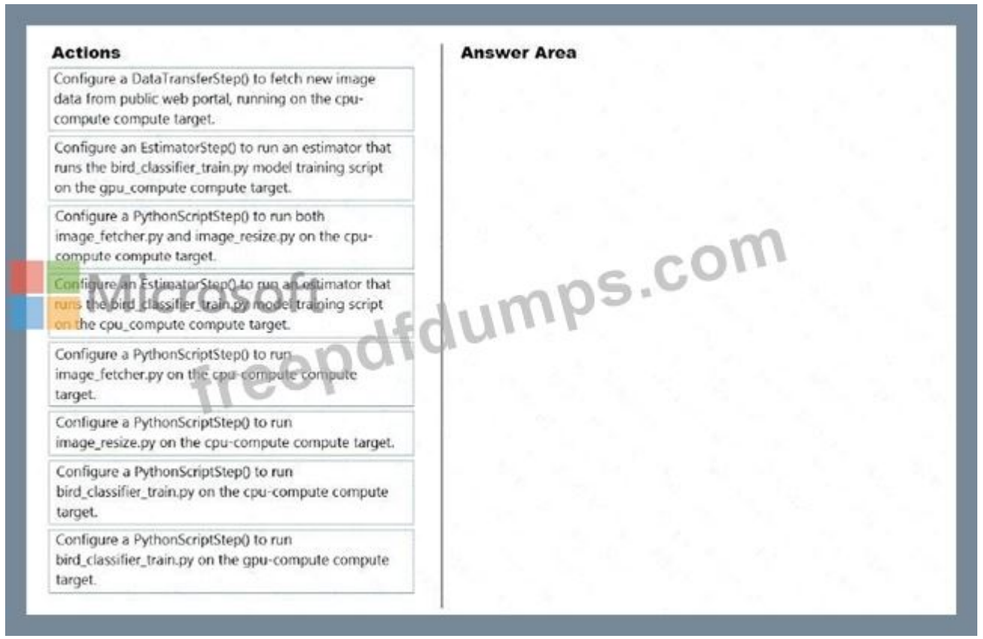
Which two chunking approaches should you use? Each correct answer presents a complete solution. Choose two.

NOTE: Each correct selection is worth one point.
A. Sentence-based parsing
B. Fixed-size parsing
C. Large language model augmentation
D. Document layout analysis
Answer: C,D (LEAVE A REPLY)

NEW QUESTION: 24
You create an Azure Machine Learning workspace named workspaces. You create a Python SDK v2 notebook to perform custom model training in workspace1. You need to run the notebook from Azure Machine Learning Studio in workspace1. What should you provision first?
A. real-time endpoint
B. Azure Machine Learning compute cluster
C. Azure Machine Learning compute instance
D. default storage account
Answer: C (LEAVE A REPLY)

NEW QUESTION: 25
You manage an Azure Machine Learning workspace.
An MLflow model is already registered. You plan to customize how the deployment does inference. You need to deploy the MLflow model to a batch endpoint for batch inferencing. What should you create first?
A. deployment definition
B. environment
C. deployment
D. scoring script
Answer: D (LEAVE A REPLY)

NEW QUESTION: 26
You create a multi-class image classification deep learning experiment by using the PyTorch framework. You plan to run the experiment on an Azure Compute cluster that has nodes with GPU's.
You need to define an Azure Machine Learning service pipeline to perform the monthly retraining of the image classification model. The pipeline must run with minimal cost and minimize the time required to train the model.
Which three pipeline steps should you run in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.



Answer:

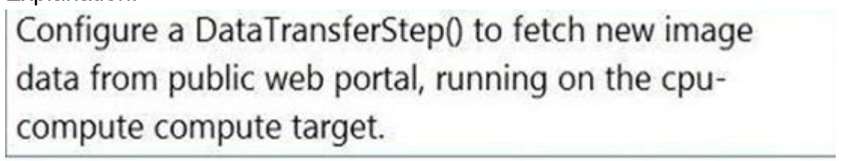
Actions

- Configure a DataTransferStep() to fetch new image data from public web portal, running on the cpu-compute compute target.
- Configure an EstimatorStep() to run an estimator that runs the bird_classifier_train.py model training script on the gpu_compute compute target.
- Configure a PythonScriptStep() to run both image_fetcher.py and image_resize.py on the cpu-compute compute target.
- Configure an EstimatorStep() to run an estimator that runs the bird_classifier_train.py model training script on the cpu_compute compute target.
- Configure a PythonScriptStep() to run image_fetcher.py on the cpu-compute compute target.
- Configure a PythonScriptStep() to run image_resize.py on the cpu-compute compute target.
- Configure a PythonScriptStep() to run bird_classifier_train.py on the cpu-compute compute target.
- Configure a PythonScriptStep() to run bird_classifier_train.py on the gpu-compute compute target.

Answer Area

- Configure a DataTransferStep() to fetch new image data from public web portal, running on the cpu-compute compute target.
- Configure a PythonScriptStep() to run image_resize.py on the cpu-compute compute target.
- Configure an EstimatorStep() to run an estimator that runs the bird_classifier_train.py model training script on the gpu_compute compute target.

Explanation:



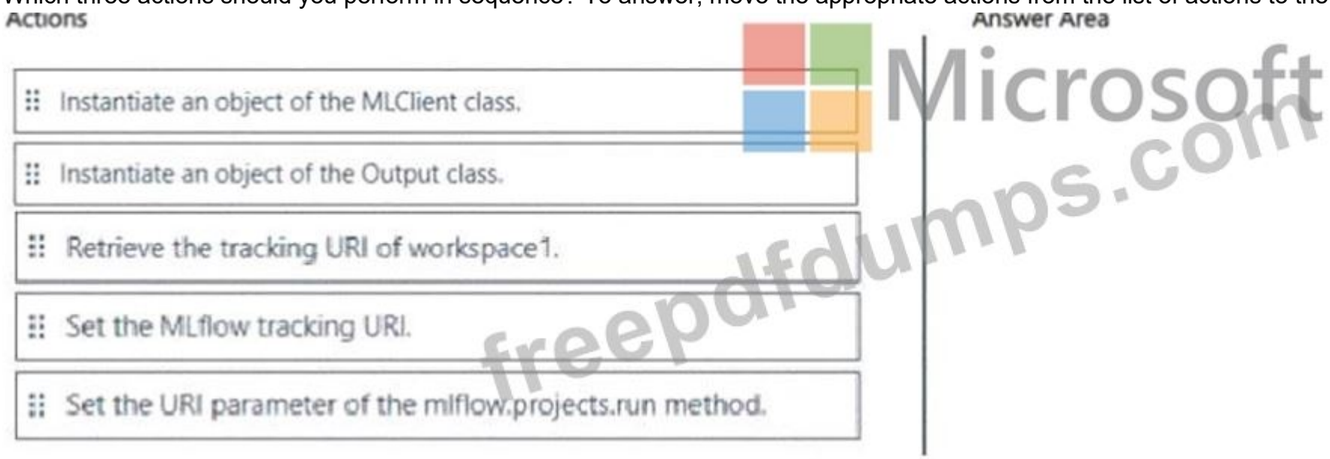
Step 1: Configure a DataTransferStep() to fetch new image data...

Step 2: Configure a PythonScriptStep() to run image_resize.y on the cpu-compute compute target.

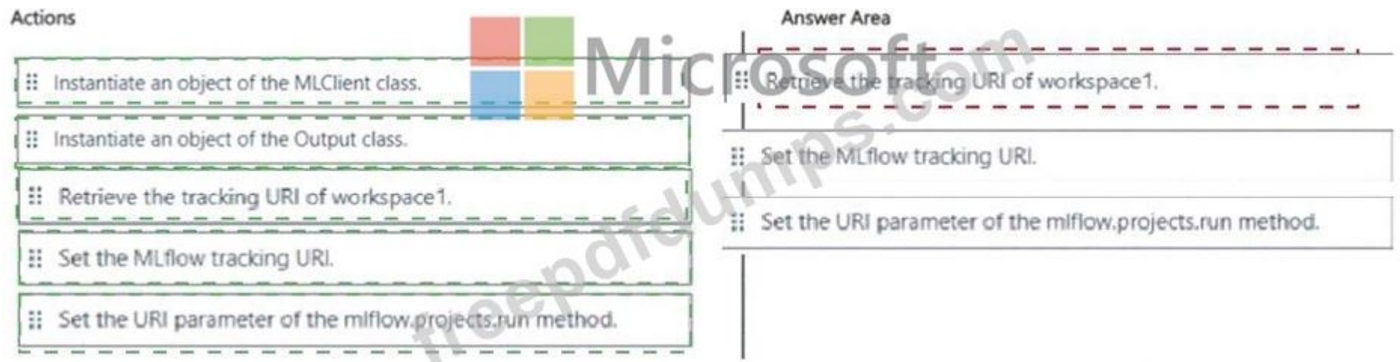
Step 3: Configure the EstimatorStep() to run training script on the gpu_compute computer target.

The PyTorch estimator provides a simple way of launching a PyTorch training job on a compute target.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-train-pytorch

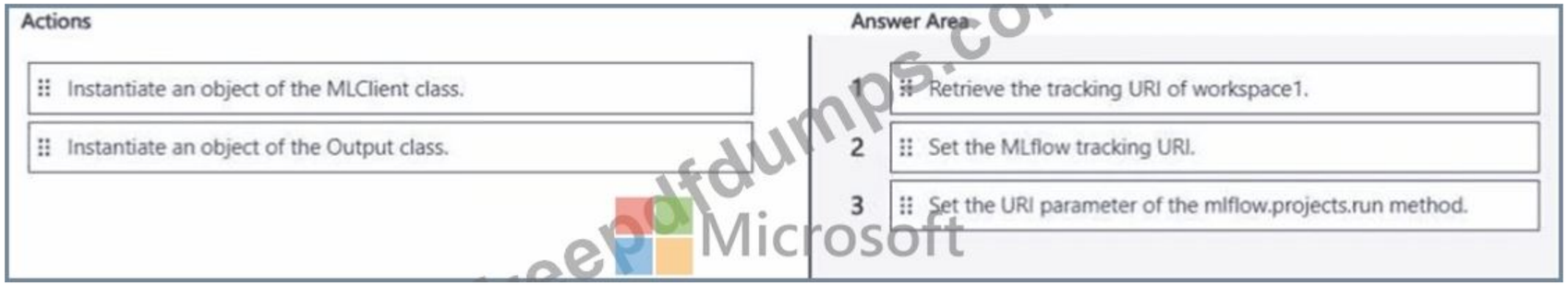
NEW QUESTION: 27
You manage an Azure Machine Learning workspace named workspace 1 and a Data Science Virtual Machine (DSVM) named DSMV1.
You must run an experiment on DSMV1 by using a Jupyter notebook and Python SDK v2 code. You must store metrics and artifacts in workspace1. You start by creating Python SDK v2 code to import all required packages.
You need to implement the Python SDK v2 code to store metrics and artifacts in workspace1.
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.



Answer:



Explanation:

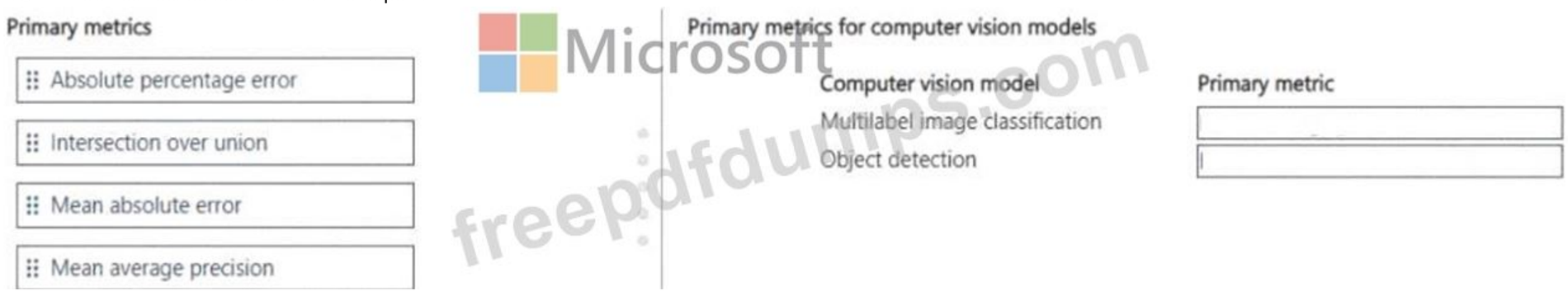


NEW QUESTION: 28
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.
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 (LEAVE A REPLY)
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

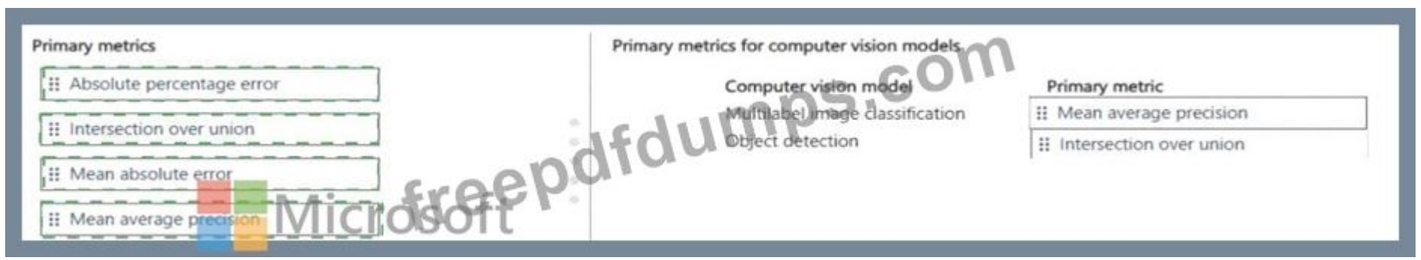
NEW QUESTION: 29
You manage an Azure At Foundry project. You deploy an Azure OpenAI Service chat model. You configure the development environment with the necessary packages to build a Prompt flow in Visual Studio Code IDE.
You create a Python file called chat.py. You must configure the large language model with 0.2 as the temperature.
You need to develop the chat application.
Which two actions should you perform? Each correct answer presents part of the solution. Choose two.
NOTE: Each correct selection is worth one point.

A. Use chat.complete0 to configure the temperature of the deployed model.
B. Use the render0 function to reference the deployed Azure OpenAI chat model.
C. Configure the temperature in the prompt template.
D. Configure the prompt template to reference the deployed Azure OpenAI chat model.
E. Configure the project client to reference the deployed Azure OpenAI chat model.
Answer: (SHOW ANSWER)

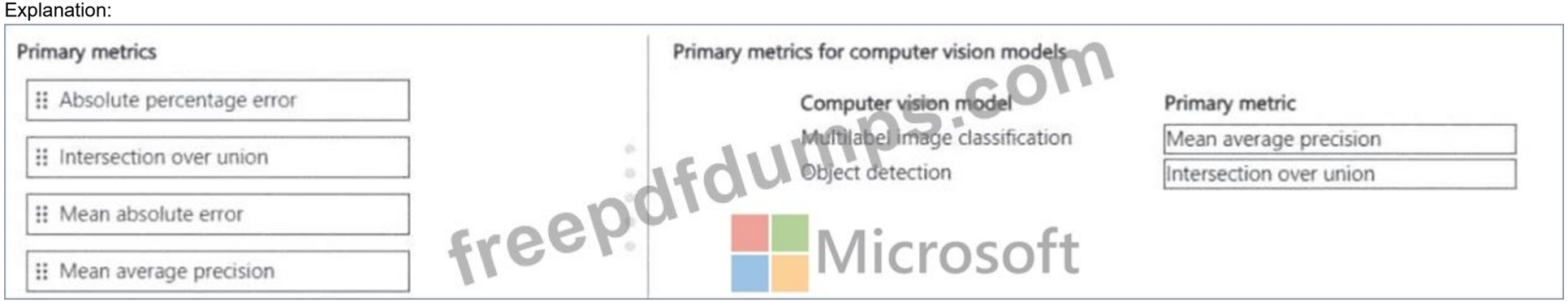
NEW QUESTION: 30
You manage an Azure Machine Learning workspace. You build automated machine learning training experiments for computer vision models.
You need to use a primary metric for model optimization and hyperparameter tuning for each model.
Which primary metrics should you use for the models? To answer, move the appropriate primary metrics to the correct computer vision models. You may use each primary metric once, more than once, or not at all.
You may need to move the split bar between panes or scroll to view content.
NOTE: Each correct selection is worth one point.



Answer:

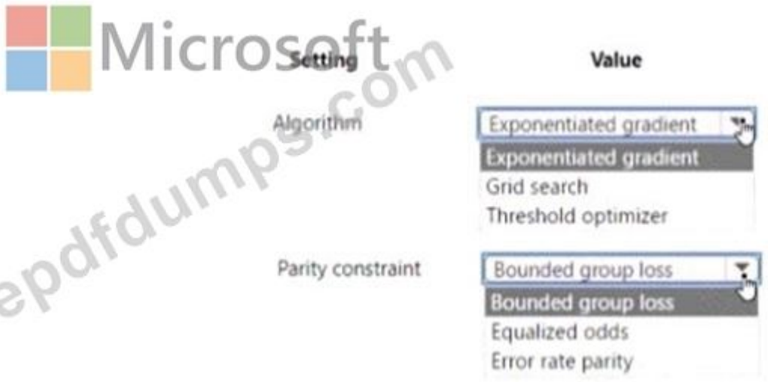


Explanation:

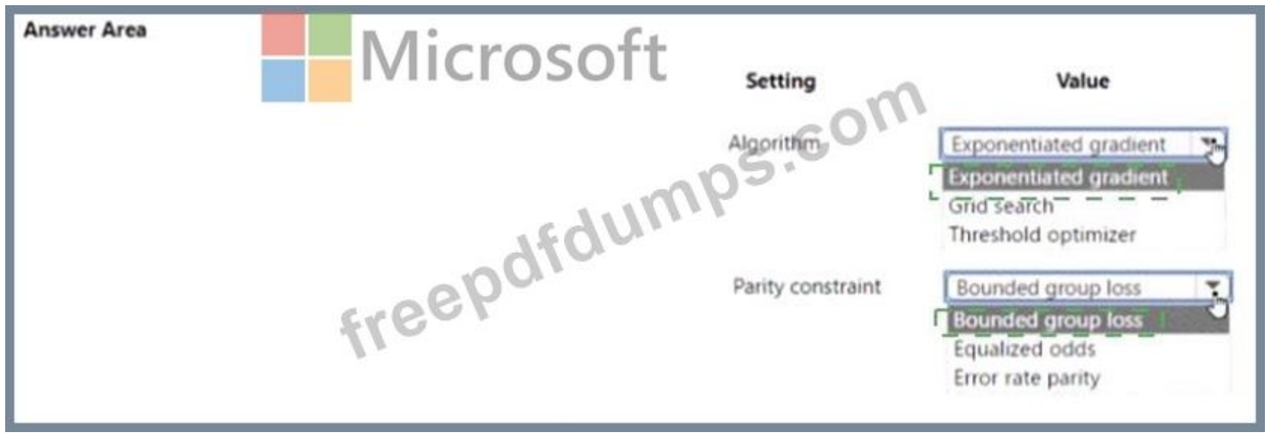


NEW QUESTION: 31
You are profiling mltabte data assets by using Azure Machine Learning studio. You need to detect columns with odd or missing values. Which statistic should you analyze?
A. Error count
B. Profile
C. Std deviation
D. Type
Answer: A (LEAVE A REPLY)

NEW QUESTION: 32
You have a binary classifier that predicts positive cases of diabetes within two separate age groups.
The classifier exhibits a high degree of disparity between the age groups.
You need to modify the output of the classifier to maximize its degree of fairness across the age groups and meet the following requirements:
* Eliminate the need to retrain the model on which the classifier is based.
* Minimize the disparity between true positive rates and false positive rates across age groups.
Which algorithm and panty constraint should you use? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.
Answer Area



Answer:



Explanation:

NEW QUESTION: 33

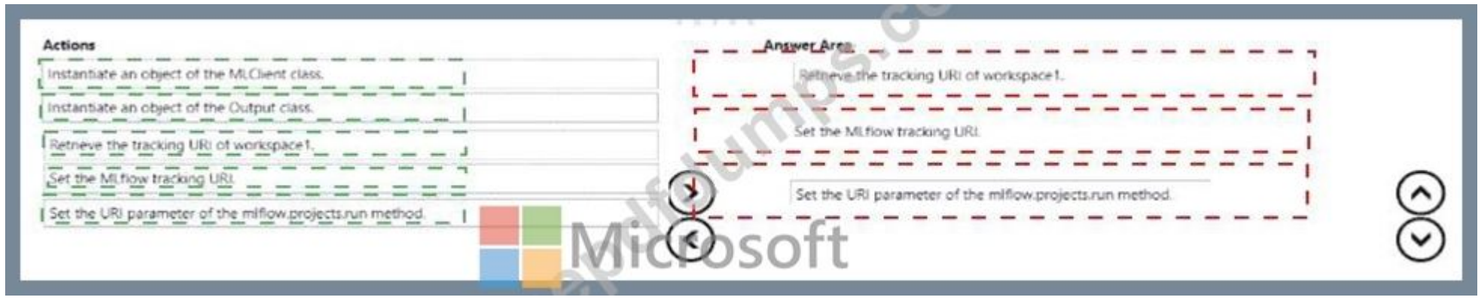
You manage an Azure Machine Learning workspace named workspace1 and a Data Science Virtual Machine (DSVM) named DSMV1.

You must an experiment in DSMV1 by using a Jupiter notebook and Python SDK v2 code. You must store metrics and artifacts in workspace 1 You start by creating Python SCK v2 code to import ail required packages.

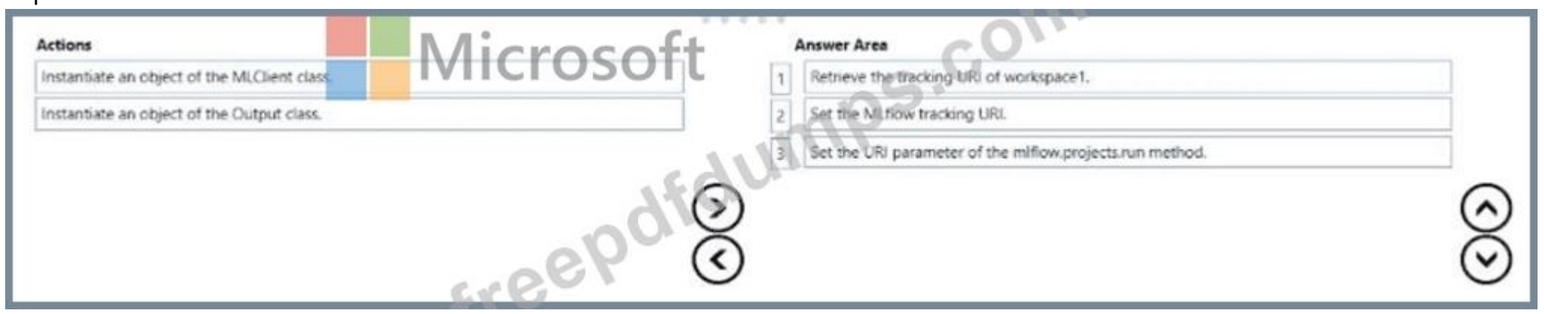
You need to implement the Python SOK v2 code to store metrics and article in workspace1.

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 the correctly order.

Answer:



Explanation:



NEW QUESTION: 34

You create a multi-class image classification model with automated machine learning in Azure Machine Learning.

You need to prepare labeled image data as input for model training in the form of an Azure Machine Learning tabular dataset.

Which data format should you use?

A. COCO

B. Pascal VOC

C. JSONL

D. JSON

Answer: D (LEAVE A REPLY)

NEW QUESTION: 35

You are creating a new experiment in Azure Machine Learning Studio. You have a small dataset that has missing values in many columns. The data does not require the application of predictors for each column. You plan to use the Clean Missing Data module to handle the missing data.

You need to select a data cleaning method.

Which method should you use?

A. Synthetic Minority Oversampling Technique (SMOTE)

B. Replace using MICE

C. Replace using; Probabilistic PCA

D. Normalization

Answer: (SHOW ANSWER)

Replace using Probabilistic PCA: Compared to other options, such as Multiple Imputation using Chained Equations (MICE), this option has the advantage of not requiring the application of predictors for each column. Instead, it approximates the covariance for the full dataset. Therefore, it might offer better performance for datasets that have missing values in many columns.

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

NEW QUESTION: 36
You have an Azure Machine Learning workspace.
You plan to tune a model hyperparameter when you train the model.
You need to define a search space that returns a normally distributed value.
Which parameter should you use?
A. Uniform
B. QLogNormal
C. LogUniform
D. QUniform
Answer: D (LEAVE A REPLY)

NEW QUESTION: 37
You are implementing a machine learning model to predict stock prices.
The model uses a PostgreSQL database and requires GPU processing.
You need to create a virtual machine that is pre-configured with the required tools.
What should you do?
A. Create a Deep Learning Virtual Machine (DLVM) Windows edition.
B. Create a Data Science Virtual Machine (DSVM) Windows edition.
C. Create a Data Science Virtual Machine (DSVM) Linux edition.
D. Create a Geo AI Data Science Virtual Machine (Geo-DSVM) Windows edition.
E. Create a Deep Learning Virtual Machine (DLVM) Linux edition.
Answer: (SHOW ANSWER)

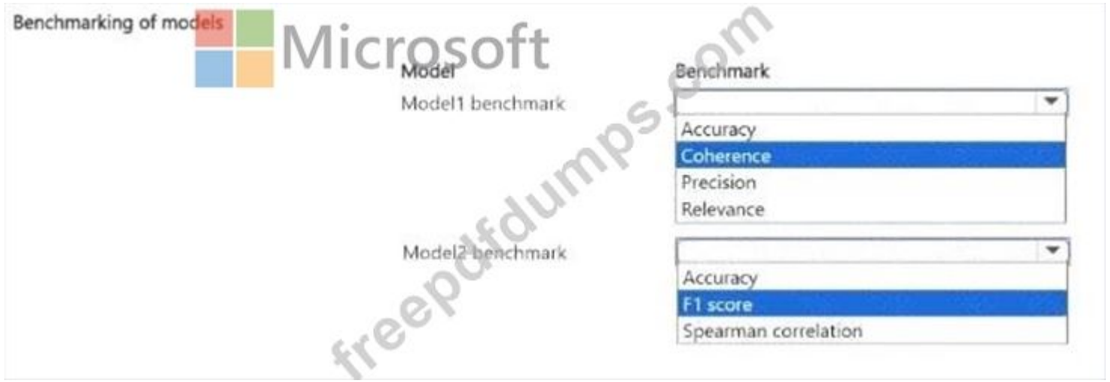
NEW QUESTION: 38
You have an Azure Machine Learning workspace that includes an AmICompute cluster and a batch endpoint.
You clone a repository that contains an MLflow model to your local computer. You need to ensure that you can deploy the model to the batch endpoint.
Solution: Add a compute resource to the workspace.
Does the solution meet the goal?
A. Yes
B. No
Answer: (SHOW ANSWER)

NEW QUESTION: 39
A set of CSV files contains sales records. All the CSV files have the same data schema.
Each CSV file contains the sales record for a particular month and has the filename sales.csv. Each file in stored in a folder that indicates the month and year when the data was recorded. The folders are in an Azure blob container for which a datastore has been defined in an Azure Machine Learning workspace. The folders are organized in a parent folder named sales to create the following hierarchical structure:

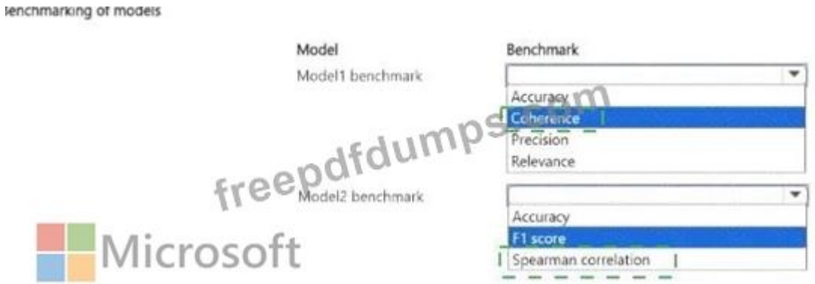
At the end of each month, a new folder with that month's sales file is added to the sales folder.
You plan to use the sales data to train a machine learning model based on the following requirements:
You must define a dataset that loads all of the sales data to date into a structure that can be easily converted to a dataframe.
You must be able to create experiments that use only data that was created before a specific previous month, ignoring any data that was added after that month.
You must register the minimum number of datasets possible.
You need to register the sales data as a dataset in Azure Machine Learning service workspace.
What should you do?
A. Create a tabular dataset that references the datastore and explicitly specifies each ' sales/mm-yyyy/sales.
csv ' file every month. Register the dataset with the name sales_dataset each month, replacing theexisting dataset and specifying a tag named month indicating the month and year it was registered.
Usethis dataset for all experiments.
B. Create a tabular dataset that references the datastore and specifies the path ' sales/*/sales.csv ' , register the dataset with the name sales_dataset and a tag named month indicating the month and year it was registered, and use this dataset for all experiments.
C. Create a new tabular dataset that references the datastore and explicitly specifies each ' sales/mm-yyyy/ sales.csv ' file every month. Register the dataset with the name sales_dataset_MM-YYYY each month with appropriate MM and YYYY values for the month and year. Use the appropriate month-specific dataset for experiments.
D. Create a tabular dataset that references the datastore and explicitly specifies each ' sales/mm-yyyy/sales.csv ' file. Register the dataset with the name sales_dataset each month as a new version and with a tag named month indicating the month and year it was registered. Use this dataset for all experiments, identifying the version to be used based on the month tag as necessary.
Answer: B (LEAVE A REPLY)
Specify the path.
Example:
The following code gets the workspace existing workspace and the desired datastore by name. And then passes the datastore and file locations to the path parameter to create a new TabularDataset, weather_ds.
from azureml.core import Workspace, Datastore, Dataset
datastore_name = ' your datastore name '
# get existing workspace
workspace = Workspace.from_config()
# retrieve an existing datastore in the workspace by name
datastore = Datastore.get(workspace, datastore_name)
# create a TabularDataset from 3 file paths in datastore
datastore_paths = [(datastore, ' weather/2018/11.csv '),
(datastore, ' weather/2018/12.csv '),
(datastore, ' weather/2019/*.csv ')]
weather_ds = Dataset.Tabular.from_delimited_files(path=datastore_paths)

NEW QUESTION: 40

You manage an Azure AI Foundry project.
You plan 10 build a RAG solution. The solution must include two models:
* One for text output, named Model1. This model must resemble human language and read naturally.
* One for creating embeddings, named Model2. This model must maximize the retrieval of relevant results (high recall) You need to compare different models by using benchmarking metrics to select the appropriate models for Model1 and Model?

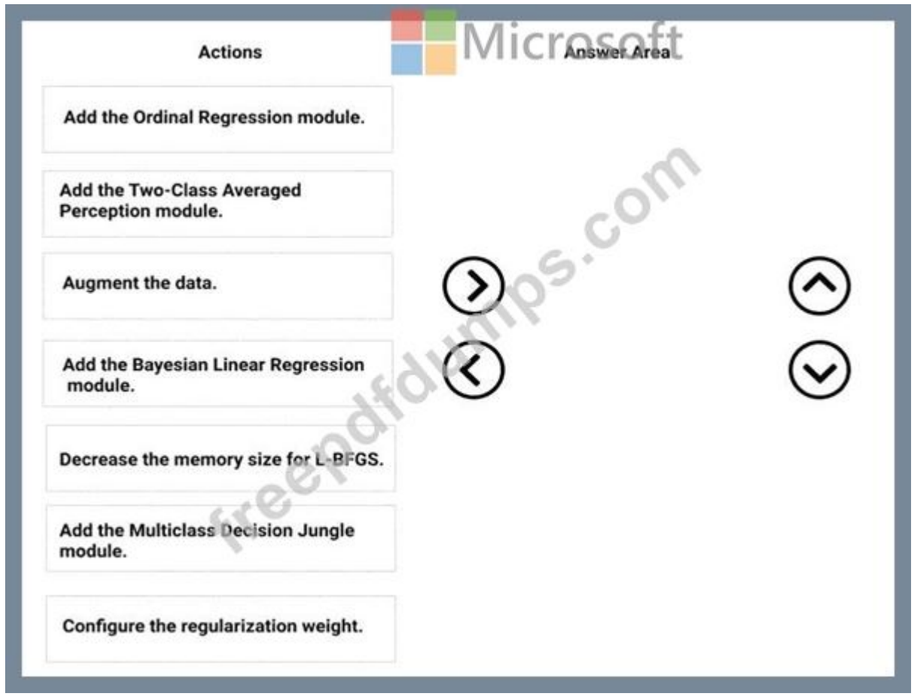


Answer:

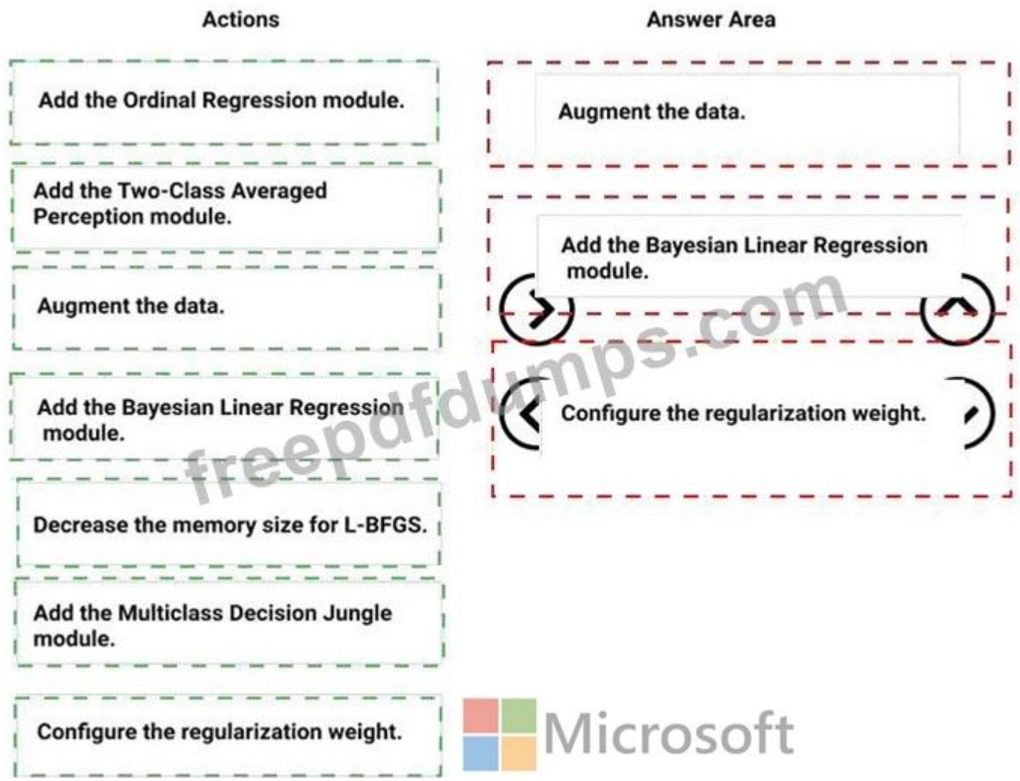


Explanation:

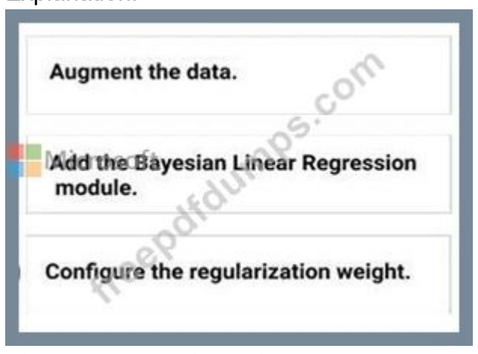
NEW QUESTION: 41
You need to correct the model fit issue.
Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.


Actions
Answer Area
Add the Ordinal Regression module.
Add the Two-Class Averaged Perception module.
Augment the data.
Add the Bayesian Linear Regression module.
Decrease the memory size for L-BFGS.
Add the Multiclass Decision Jungle module.
Configure the regularization weight.

Answer:



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: 42

You perform hyper parameter tuning with Azure Machine Learning.

You create the following Python code:

```
from azure.ai.ml.sweep import Normal, Uniform

command_job_for_sweep = command_job(
    learning_rate=Normal(mu=10, sigma=3),
    keep_probability=Uniform(min_value=0.05, max_value=0.1),
)
```

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

Answer Area

Statements	Yes	No
The code defines a search space by using the `learning_rate` and `keep_probability` parameters.	○	○
The logarithm of the `learning_rate` parameter has a normal distribution.	○	○
The `keep_probability` parameter has a uniform distribution with a minimum value of 0.05 and a maximum value of 0.1.	○	○

Answer:

Answer Area

Statements	Yes	No
The code defines a search space by using the `learning_rate` and `keep_probability` parameters.	◉	○
The logarithm of the `learning_rate` parameter has a normal distribution.	◉	○
The `keep_probability` parameter has a uniform distribution with a minimum value of 0.05 and a maximum value of 0.1.	◉	○

Explanation:

Answer Area

Statements	Yes	No
The code defines a search space by using the `learning_rate` and `keep_probability` parameters.	◉	○
The logarithm of the `learning_rate` parameter has a normal distribution.	◉	○
The `keep_probability` parameter has a uniform distribution with a minimum value of 0.05 and a maximum value of 0.1.	◉	○

NEW QUESTION: 43

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.estimator import Estimator
sk_est = Estimator(source_directory='./scripts',
 compute_target=aml-compute,
 entry_script='train.py',
 conda_packages=['scikit-learn'])
```

Does the solution meet the goal?
A. Yes
B. No
Answer: B (LEAVE A REPLY)
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: 44
You create and register a model in an Azure Machine Learning workspace.
You must use the Azure Machine Learning SDK to implement a batch inference pipeline that uses a ParallelRunStep to score input data using the model. You must specify a value for the ParallelRunConfig compute_target setting of the pipeline step.
You need to create the compute target.
Which class should you use?
A. BatchCompute
B. AdlaCompute
C. AmlCompute
D. Aks Compute
Answer: C (LEAVE A REPLY)
Compute target to use for ParallelRunStep. This parameter may be specified as a compute target object or the string name of a compute target in the workspace.
The compute_target target is of AmlCompute or string.

Note: An Azure Machine Learning Compute (AmlCompute) is a managed-compute infrastructure that allows you to easily create a single or multi-node compute. The compute is created within your workspace region as a resource that can be shared with other users Reference:
https://docs.microsoft.com/en-us/python/api/azureml-contrib-pipeline-steps/azureml.contrib.pipeline.steps.
parallelrunconfig
https://docs.microsoft.com/en-us/python/api/azureml-core/azureml.core.compute.amlcompute(class)

NEW QUESTION: 45
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)
```


Answer:

```
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)
```


Explanation:

```
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)
```


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: 46

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 use Azure Machine Learning designer to load the following datasets into an experiment:

Dataset1

Age	Length	Width
3	22	13
7	11	96
18	32	85

Dataset2

Age	Length	Width
11	101	65
6	98	23
33	22	54
17	52	12

You need to create a dataset that has the same columns and header row as the input datasets and contains all rows from both input datasets.
Solution: Use the Execute Python Script module.
Does the solution meet the goal?
A. No
B. Yes
Answer: (SHOW ANSWER)

NEW QUESTION: 47
You manage an Azure Machine Learning workspace. The development environment tor managing the workspace is configured to use Python SDK v2 in Azure Machine Learning Notebooks A Synapse Spark Compute is currently attached and uses system-assigned identity You need to use Python code to update the Synapse Spark Compute 10 use a user-assigned identity.
Solution: Configure the IdentityConfiguration class with the appropriate identity type.
Does the solution meet the goal?
A. Yes
B. No
Answer: (SHOW ANSWER)

NEW QUESTION: 48
You use differential privacy to ensure your reports are private. The calculated value of the epsilon for your data is 1.8. You need to modify your data to ensure your reports are private. Which epsilon value should you accept for your data?
A. between 3 and 10
B. between 2 and 3
C. between 0 and 1

D. more than 10
Answer: B (LEAVE A REPLY)

NEW QUESTION: 49
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.
After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.
You are using Azure Machine Learning to run an experiment that trains a classification model.
You want to use Hyperdrive to find parameters that optimize the AUC metric for the model. You configure a HyperDriveConfig for the experiment by running the following code:

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

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

```
import numpy as np
from sklearn.metrics import roc_auc_score
from azureml.core.run import Run
run = Run.get_context()
# code to train model omitted
auc = roc_auc_score(y_test, y_predicted)
run.log("AUC", np.float(auc))
```

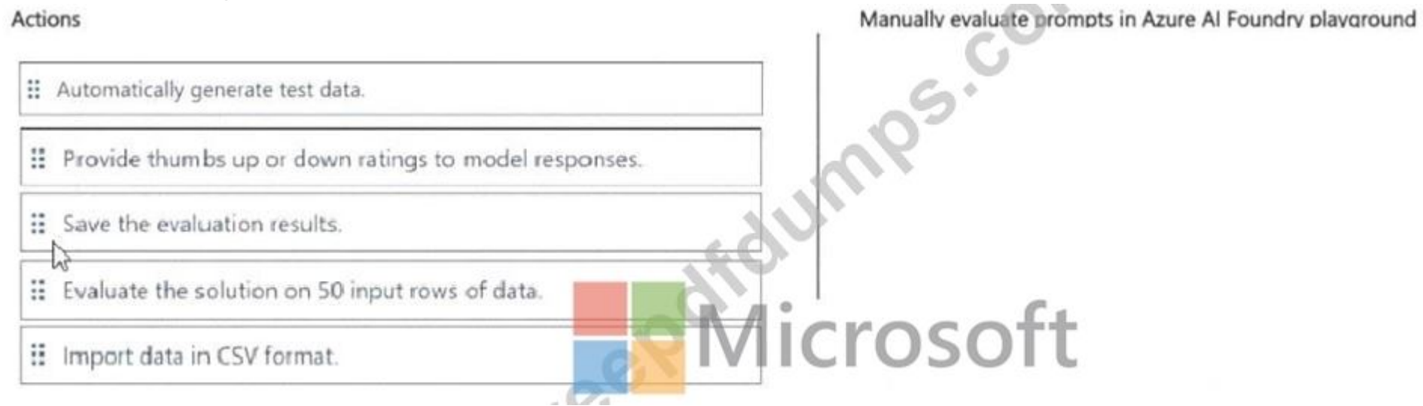
Does the solution meet the goal?
A. No
B. Yes
Answer: (SHOW ANSWER)

NEW QUESTION: 50
You have a dataset that is stored m an Azure Machine Learning workspace.
You must perform a data analysis for differentiate privacy by using the SmartNoise SDK.
You need to measure the distribution of reports for repeated queries to ensure that they are balanced Which type of test should you perform?
A. Utility
B. Accuracy
C. Bias
D. Privacy
Answer: B (LEAVE A REPLY)

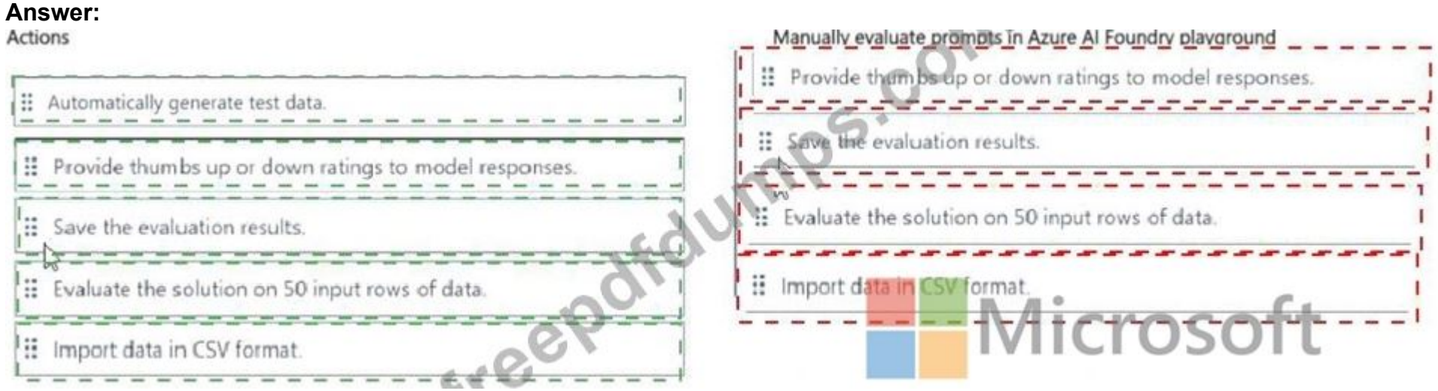
NEW QUESTION: 51
You create an Azure Machine learning workspace.
You are use the Azure Machine -learning Python SDK v2 to define the search space for concrete hyperparafneters. The hyper parameters must consist of a list of predetermined, comma-separated.

You need to import the class from the azure ai ml. sweep package used to create the list of values.
Which class should you import?
A. Uniform
B. Choice
C. Normal
D. Randint
Answer: (SHOW ANSWER)

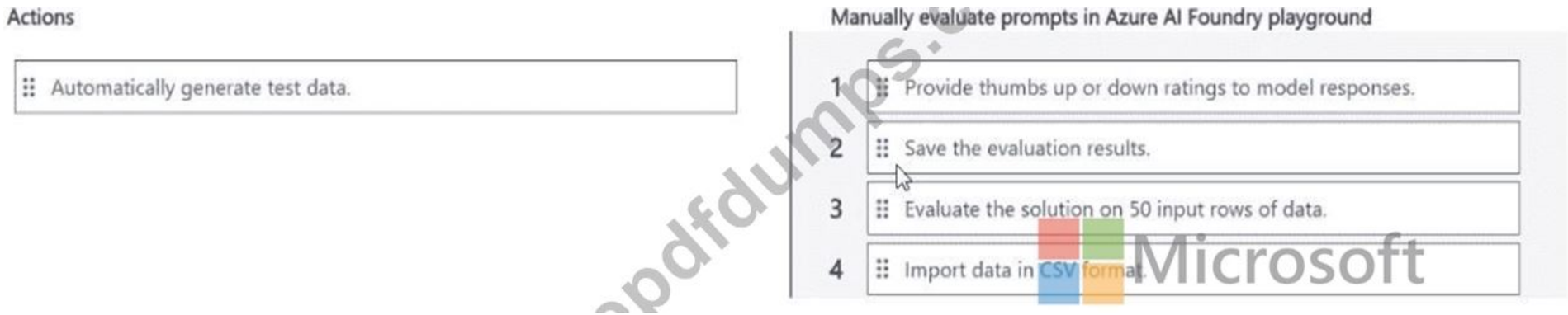
NEW QUESTION: 52
You manage an Azure AI Foundry project.
You deploy a large language model from the model catalog.
You need to manually evaluate the model, collect the statistics, and be able to review the results later.



Answer:



Explanation:



NEW QUESTION: 53
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 (LEAVE A REPLY)
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: 54
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 use Azure Machine Learning designer to load the following datasets into an experiment:

Dataset1

Age	Length	Width
3	22	13
7	11	96
18	32	85

Dataset2

Age	Length	Width
11	101	65
6	98	23
33	22	54
17	52	12

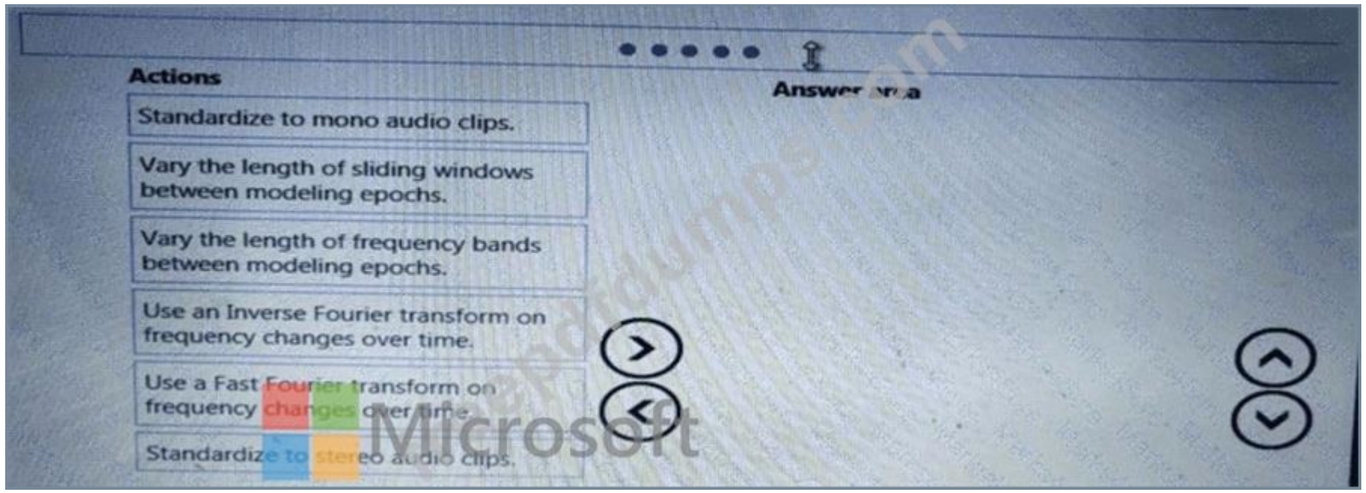
You need to create a dataset that has the same columns and header row as the input datasets and contains all rows from both input datasets.
Solution: Use the Add Rows module.
Does the solution meet the goal?
A. No
B. Yes
Answer: A (LEAVE A REPLY)

NEW QUESTION: 55
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.
After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.
An IT department creates the following Azure resource groups and resources:

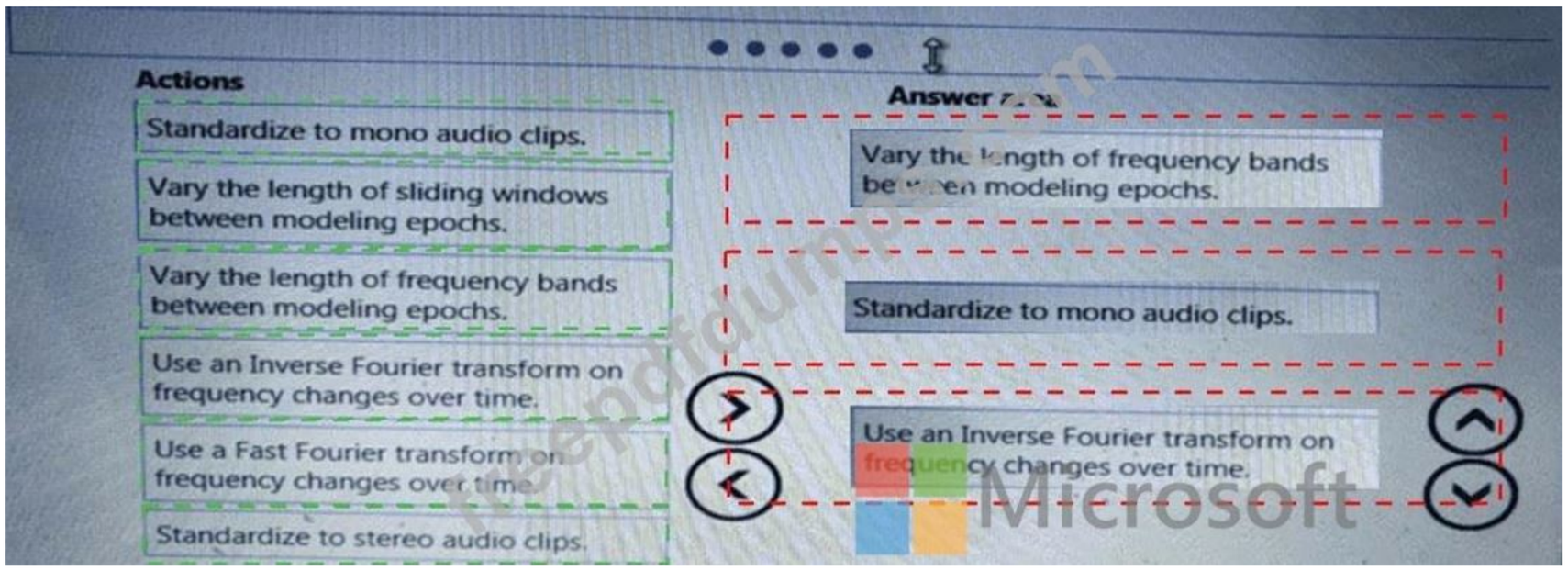
Resource group	Resources
ml_resources	• an Azure Machine Learning workspace named amlworkspace • an Azure Storage account named amlworkspace12345 • an Application Insights instance named amlworkspace54321 • an Azure Key Vault named amlworkspace67890 • an Azure Container Registry named amlworkspace09876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

The IT department creates an Azure Kubernetes Service (AKS)-based inference compute target named aks- cluster in the Azure Machine Learning workspace.
You have a Microsoft Surface Book computer with a GPU. Python 3.6 and Visual Studio Code are installed.
You need to run a script that trains a deep neural network (DNN) model and logs the loss and accuracy metrics.
Solution: Install the Azure ML SDK on the Surface Book. Run Python code to connect to the workspace and then run the training script as an experiment on local compute.
A. Yes
B. No
Answer: B (LEAVE A REPLY)
Need to attach the mlvm virtual machine as a compute target in the Azure Machine Learning workspace.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/concept-compute-target

NEW QUESTION: 56
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.



Answer:



NEW QUESTION: 57

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: (SHOW ANSWER)

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: 58

You manage an Azure Machine Learning workspace named workspace1by using the Python SDK v2.

You must register datastores in workspace 1 for Azure Blot storage and Azure Fetes storage to meet the following requirements.

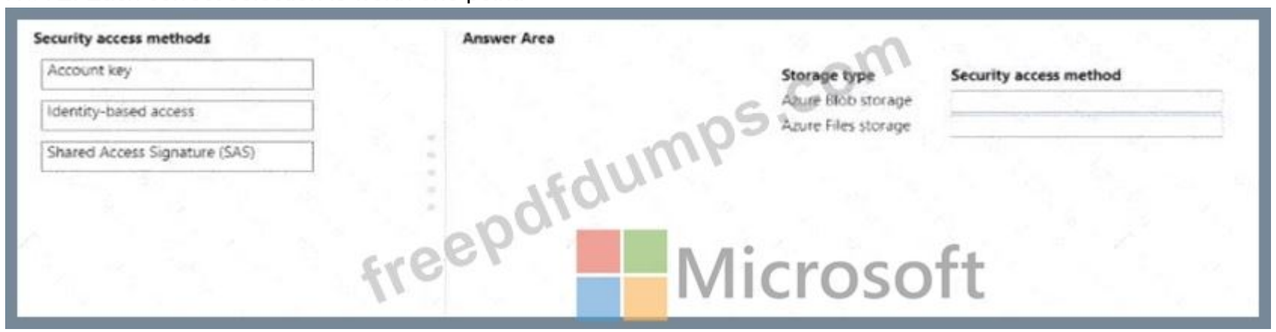
* Azure Active Directory (Azure AD) authentication must be used for access to storage when possible.

* Credentials and secrets steed in workspace1 must be valid lot a specified time period when accessing Azure Files storage.

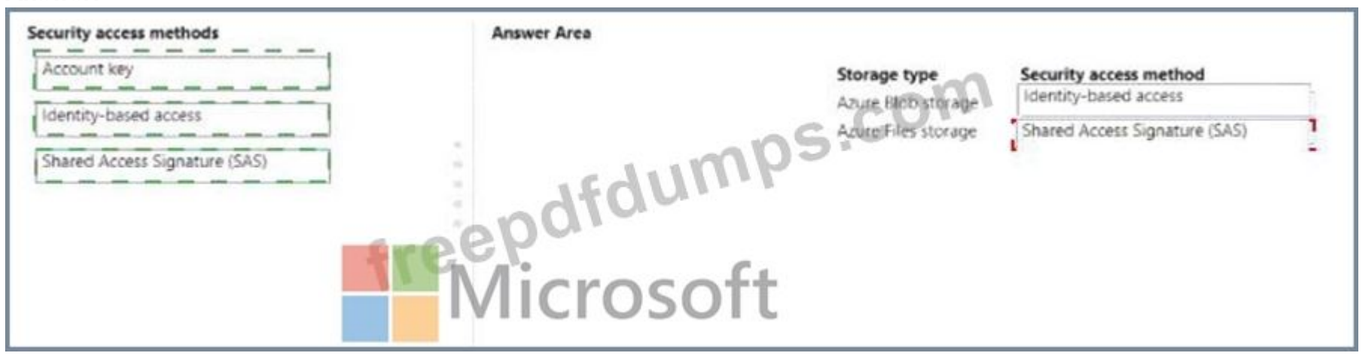
You need to configure a security access method used to register the Azure Blob and azure files storage in workspace1.

Which security access method should you configure? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.



Answer:



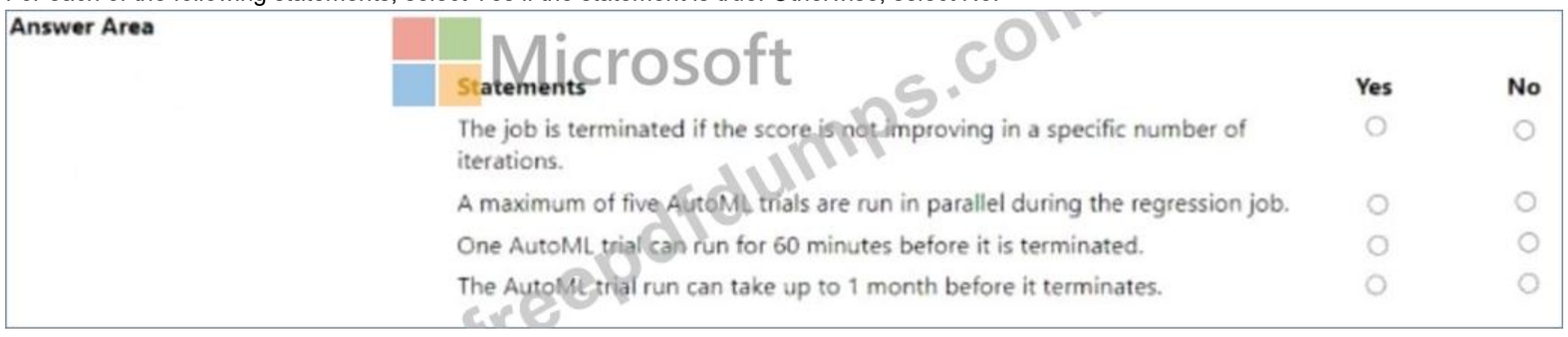
Explanation:

NEW QUESTION: 59

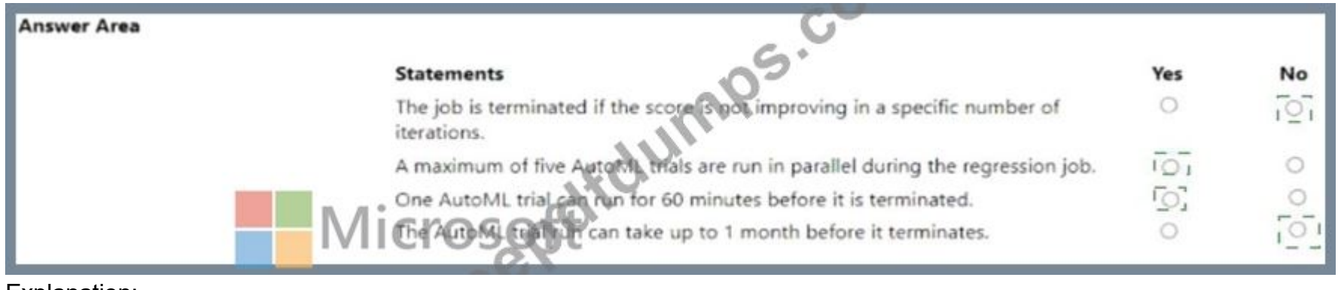
You manage an Azure Machine Learning workspace. You configure an automated machine learning regression training job by using the Azure Machine Learning Python SDK v2. You configure the regression job by using the following script:

```
regression_job.set_limits(
    timeout_minutes = 60,
    max_concurrent_trials = 5,
    enable_early_termination = True
)
```

For each of the following statements, select Yes if the statement is true. Otherwise, select No.



Answer:



Explanation:

Answer Area

Statements	Yes	No
The job is terminated if the score is not improving in a specific number of iterations.	○	●
A maximum of five AutoML trials are run in parallel during the regression job.	●	○
One AutoML trial can run for 60 minutes before it is terminated.	●	○
The AutoML trial run can take up to 1 month before it terminates.	○	●

NEW QUESTION: 60

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: A,C (LEAVE A REPLY)
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: 61
You manage an Azure Machine Learning workspace. You use Azure Machine Learning Python SDK v2 to configure a trigger to schedule a pipeline job. You need to create a time-based schedule with recurrence pattern.

Which two properties must you use to successfully configure the trigger? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

A. time_zone

B. interval

C. schedule

D. frequency

E. start.time

Answer: B,D (LEAVE A REPLY)

NEW QUESTION: 62

You have an Azure subscription that contains a resource group named rg-ml.

You plan to create an Azure Machine Learning workspace named workspacel in rg-ml by using Azure Machine Learning Python SDK v2.

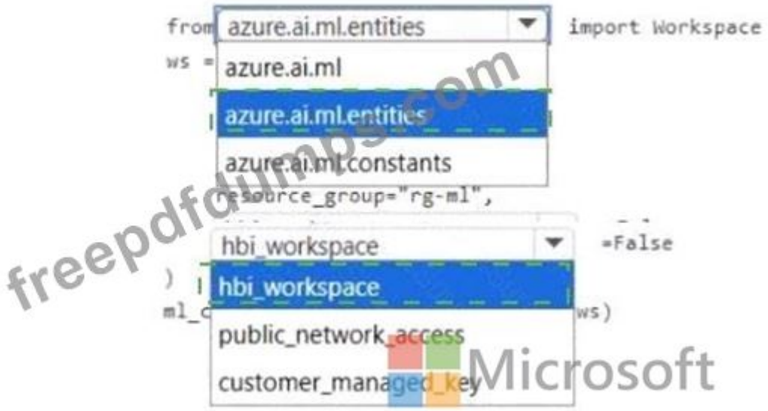
You need to ensure workspacel is configured to prevent the collection of potentially sensitive data by Microsoft diagnostics.

How should you complete the provided code? To answer, select the appropnate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer:

Azure Machine Learning Python SDK v2 code



Explanation:

NEW QUESTION: 63

You manage an Azure Machine learning workspace. The workspace includes an Azure Machine Learning kubernetes compute target configured as an Azure Kubemetes Service (AKS) cluster named AKS1 AKS1 is configured to enable the targeting of different nodes to train workloads.

You must run a command job on AK51 by using the Azure ML Python SDK v2? The command job must select different types of compute nodes. The compare node types must be specified by using a command parameter.

You need to configure the command parameter.

Which parameter should you use?

A. compute

B. environment

C. instance_type

D. limits

Answer: C (LEAVE A REPLY)

from azure.ai.ml import command

define the command

command_job = command(

command="python -c "print('Hello world!')"",

environment="AzureML-lightgbm-3.2-ubuntu18.04-py37-cpu@latest",

compute="<Kubernetes-compute_target_name>",
instance_type="<instance_type_name>"

NEW QUESTION: 64

You have an Azure Machine Learning workspace.

You plan to use Azure Machine Learning Python SDK v2 to define a pipeline component that trains an image classification model. The execution logic of the component is contained in the train() function in the file named modeljrain.py.

You write code to import all required libraries and store it as train_component.py in the same folder that contains model_train.py.

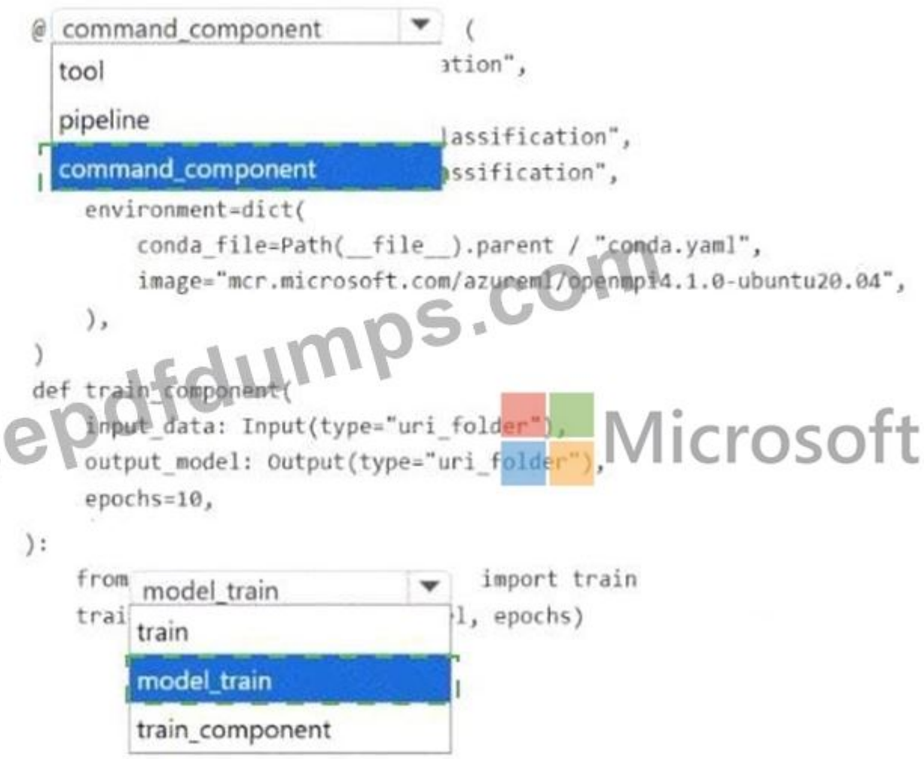
You need to complete the remaining code in train_component.py.

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

NOTE: Each correct selection is worth one point.

Answer:

Azure Machine Learning Python SDK v2 code



Explanation:

NEW QUESTION: 65

You use the Two-Class Neural Network module in Azure Machine Learning Studio to build a binary classification model. You use the Tune Model Hyperparameters module to tune accuracy for the model.

You need to select the hyperparameters that should be tuned using the Tune Model Hyperparameters module.

Which two hyperparameters should you use? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

A. Number of hidden nodes

B. Learning Rate

C. The type of the normalizer

D. Number of learning iterations

E. Hidden layer specification

Answer: D,E (LEAVE A REPLY)

D: For Number of learning iterations, specify the maximum number of times the algorithm should process the training cases.

E: For Hidden layer specification, select the type of network architecture to create.

Between the input and output layers you can insert multiple hidden layers. Most predictive tasks can be accomplished easily with only one or a few hidden layers.

References:

https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/two-class-neural-network

NEW QUESTION: 66

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 (LEAVE A REPLY)

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: 67

You are using Azure Machine Learning to monitor a trained and deployed model. You implement Event Grid to respond to Azure Machine Learning events.

Model performance has degraded due to model input data changes.

You need to trigger a remediation ML pipeline based on an Azure Machine Learning event.

Which event should you use?

A. RunStatusChanged

B. ModelDeployed

C. RunCompleted

D. DatasetDriftDetected

Answer: D (LEAVE A REPLY)

NEW QUESTION: 68

You have a Python data frame named salesData in the following format:

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

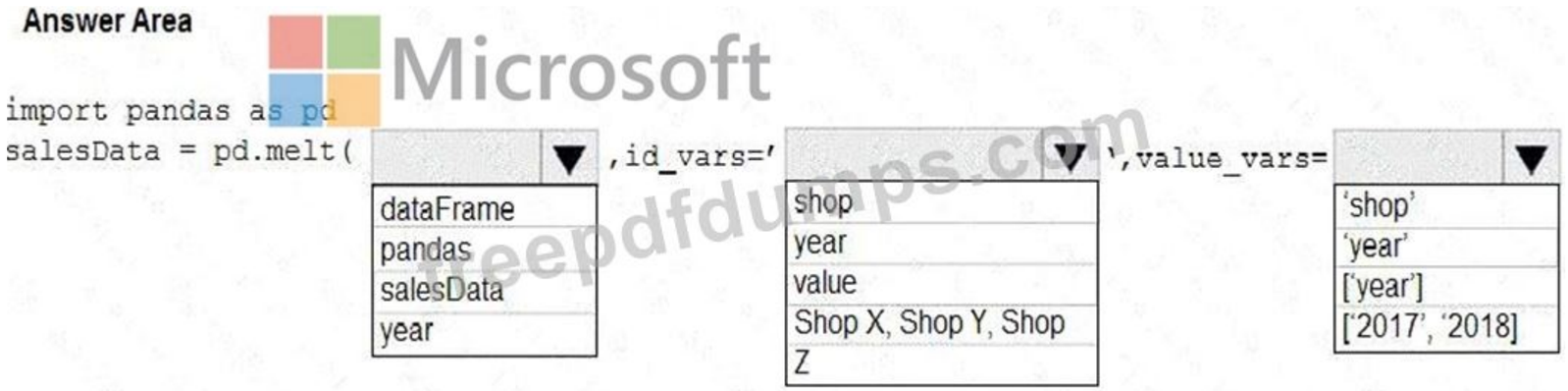
The data frame must be unpivoted to a long data format as follows:

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

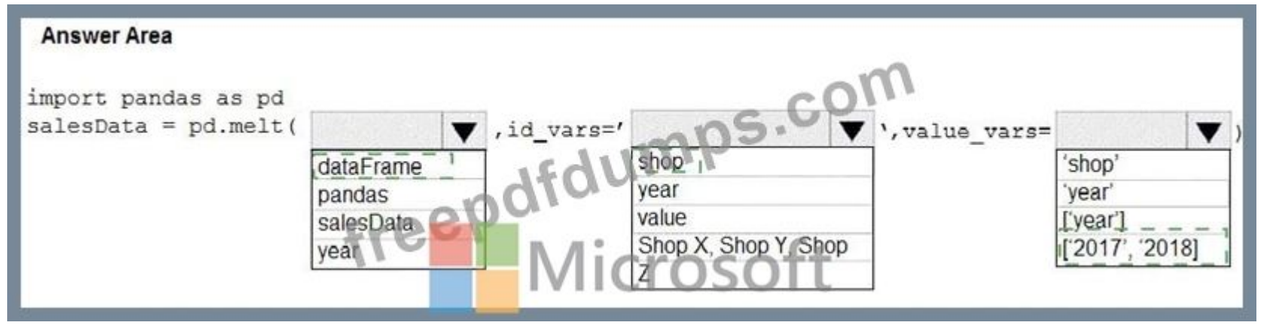
You need to use the pandas.melt() function in Python to perform the transformation.

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

NOTE: Each correct selection is worth one point.



Answer:



Explanation:

Box 1: dataFrame

Syntax: pandas.melt(frame, id_vars=None, value_vars=None, var_name=None, value_name= ' value ' , col_level=None)[source] Where frame is a DataFrame Box 2: shop Paramter id_vars id_vars : tuple, list, or ndarray, optional Column(s) to use as identifier variables.

Box 3: [' 2017 ' , ' 2018 ']

value_vars : tuple, list, or ndarray, optional

Column(s) to unpivot. If not specified, uses all columns that are not set as id_vars.

Example:

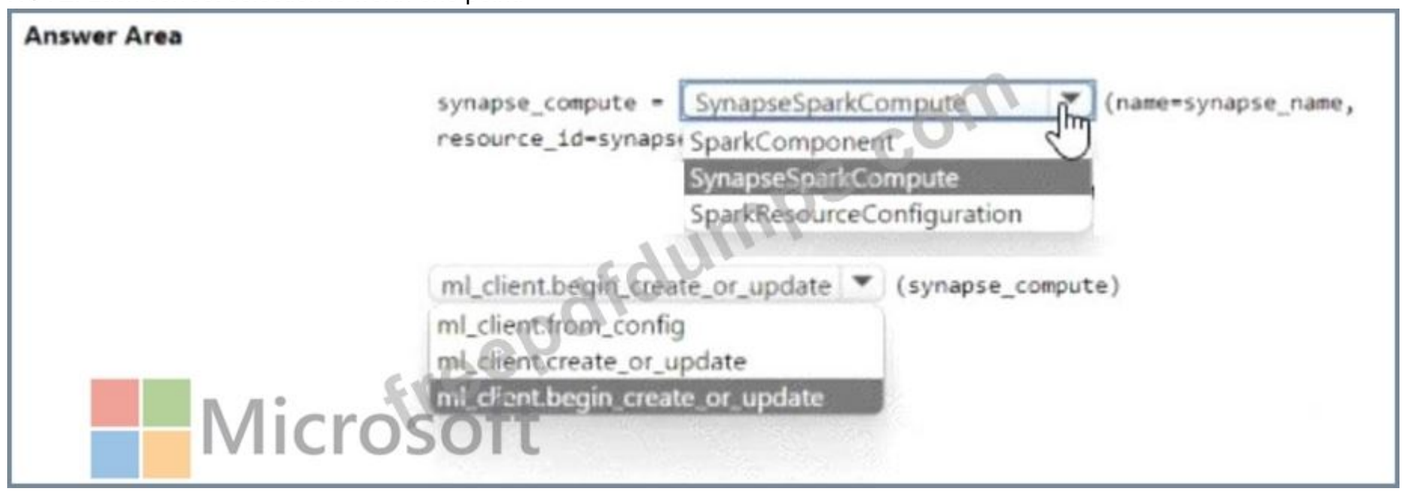
df = pd.DataFrame({ ' A ' : {0: ' a ' , 1: ' b ' , 2: ' c ' },

' B ' : {0: 1, 1: 3, 2: 5},

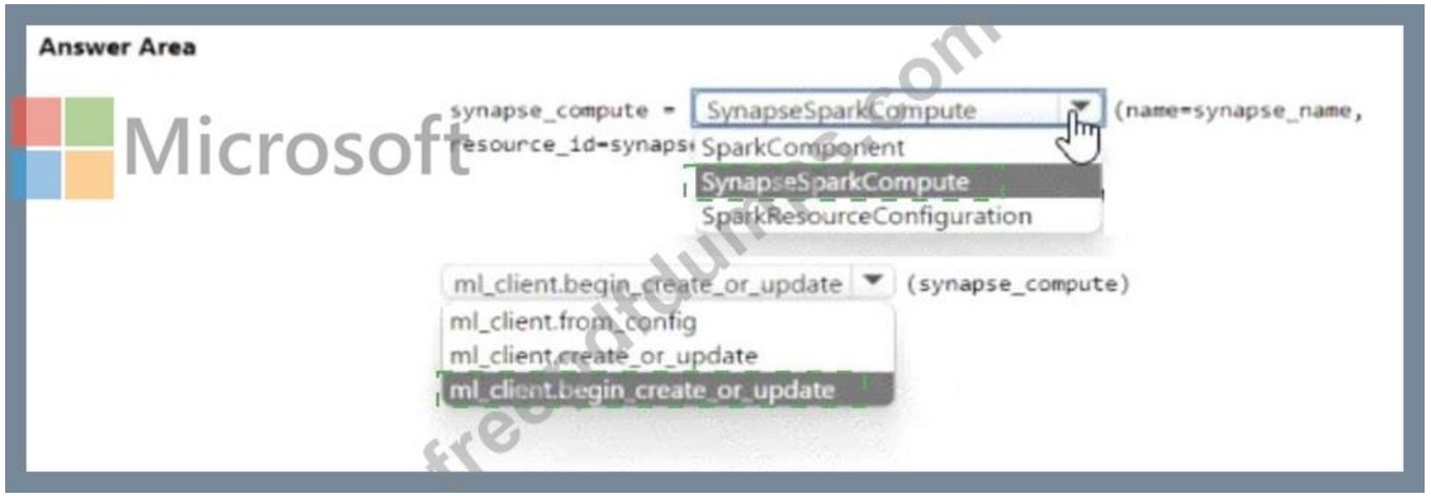
' C ' : {0: 2, 1: 4, 2: 6}})

pd.melt(df, id_vars=[' A '], value_vars=[' B ' , ' C '])
A variable value
0 a B 1
1 b B 3
2 c B 5
3 a C 2
4 b C 4
5 c C 6
References:
https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.melt.html

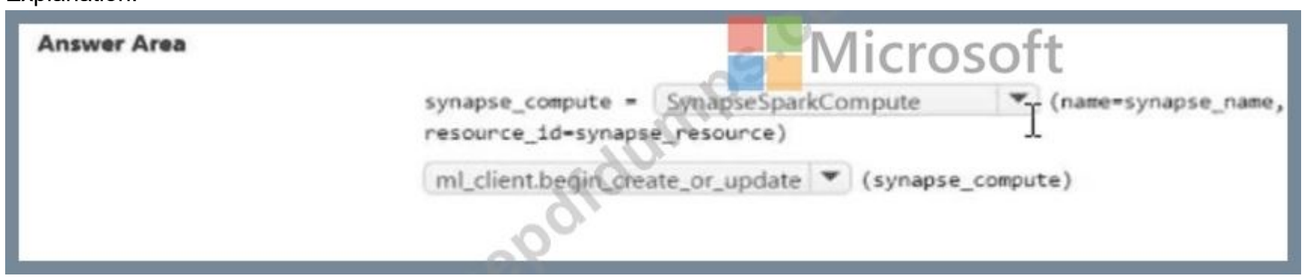
NEW QUESTION: 69
You are creating data wrangling and model training solutions in an Azure Machine Learning workspace.
You must use the same Python notebook to perform both data wrangling and model training.
You need to use the Azure Machine Learning Python SDK v2 to define and configure the Synapse Spark pool asynchronously in the workspace as dedicated compute How should you complete the rode segment? To answer, select the appropriate options in the answer area.
NOTE: Lach correct selection is worth one point.



Answer:



Explanation:



NEW QUESTION: 70
You manage an Azure Machine Learning workspace.
You choose the urijolder data type as an output of a pipeline component.
You need to define the data access mode that is supported by your configuration.
Which mode should you define?
A. direct
B. download
C. eva|-download
D. rw_mount
Answer: D (LEAVE A REPLY)

NEW QUESTION: 71
You use the Azure Machine Learning service to create a tabular dataset named training.data. You plan to use this dataset in a training script.
You create a variable that references the dataset using the following code:
training_ds = workspace.datasets.get(" training_data ")
You define an estimator to run the script.
You need to set the correct property of the estimator to ensure that your script can access the training.data dataset Which property should you set?

A.
```
inputs = [training_ds.as_named_input('training_ds')]
```
B.
```
script_params = {"--training_ds":training_ds}
```
C.
```
environment_definition = {"training_data":training_ds}
```
D.
```
source_directory = training_ds
```

Answer: (SHOW ANSWER)

Example:

Get the training dataset

diabetes_ds = ws.datasets.get(" Diabetes Dataset ")

Create an estimator that uses the remote compute

hyper_estimator = SKLearn(source_directory=experiment_folder,

inputs=[diabetes_ds.as_named_input(' diabetes ')], # Pass the dataset as an input compute_target = cpu_cluster, conda_packages=[' pandas ' , ' ipykernel ' , ' matplotlib '], pip_packages=[' azureml-sdk ' , ' argparse ' , ' pyarrow '], entry_script= ' diabetes_training.py ') Reference:

https://notebooks.azure.com/GraemeMalcolm/projects/azureml-primers/html/04%20-%20Optimizing%20Model%20Training.ipynb

NEW QUESTION: 72

You use Azure Machine Learning Designer lo load the following datasets into an experiment:

Dataset1:

Age	Length	Width
3	22	13
7	11	96
18	32	85

Dataset2:

Age	Length	Width
11	101	65
6	98	23
33	22	54
17	52	12

You need to create a dataset that has the same columns and header row as the input datasets and contains all rows from both input datasets.

Solution: Use the Add Rows component.

Does the solution meet the goal?

A. Yes

B. No

Answer: B (LEAVE A REPLY)

NEW QUESTION: 73

You manage an Azure Machine Learning workspace named workspace!.

You plan to author custom pipeline components by using Azure Machine Learning Python SDK v2.

You must transform the Python code into a YAML specification that can be processed by the pipeline service.

You need to import the Python library that provides the transformation functionality.

Which Python library should you import?
A. azure.ai.ml.entities
B. sklearn
C. azure.ai ml.automl
D. mldesigner
Answer: D (LEAVE A REPLY)

NEW QUESTION: 74
You train and register a model in your Azure Machine Learning workspace.
You must publish a pipeline that enables client applications to use the model for batch inferencing. You must use a pipeline with a single ParallelRunStep step that runs a Python inferencing script to get predictions from the input data.
You need to create the inferencing script for the ParallelRunStep pipeline step.
Which two functions should you include? Each correct answer presents part of the solution.
NOTE: Each correct selection is worth one point.
A. run(mini_batch)D
B. main()
C. batch()
D. init()
E. score(mini_batch)
Answer: A,D (LEAVE A REPLY)
Reference:
https://github.com/Azure/MachineLearningNotebooks/tree/master/how-to-use-azureml/machine- learningpipelines/parallel-run

NEW QUESTION: 75
You create an Azure Machine Learning workspace and a dataset. The dataset includes age values for a large group of diabetes patients. You use the dp.mean function from the SmartNoise library to calculate the mean of the age value. You store the value in a variable named age.mean.
You must output the value of the interval range of released mean values that will be returned 95 percent of the time.
You need to complete the code.
Which code values should you use? To answer, select the appropriate options in the answer area NOTE: Each correct selection is worth one point.

Answer:

Answer Area

Explanation:

To complete the code to output the value of the interval range of released mean values that will be returned 95 percent of the time using the SmartNoise library, you would need to use the privacy_usage_to_accuracy method with a value of 0.95 for the interval percentage. The dp.mean function from the SmartNoise library can return this interval when the required methods and parameters are used correctly.

print(age_mean.privacy_usage_to_accuracy(0.95))

age_mean is the variable that holds the differentially private mean value calculated using the dp.mean function.

privacy_usage_to_accuracy(0.95) method is called on age_mean to get the interval range for 95% confidence.

Therefore, the selections should be:

privacy_usage_to_accuracy

0.95

This will give you the interval range of released mean values that will be returned 95 percent of the time.

NEW QUESTION: 76

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 (LEAVE A REPLY)

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%.

Topic 1, Case Study 1

Overview

You are a data scientist in a company that provides data science for professional sporting events. Models will be global and local market data to meet the following business goals:

*Understand sentiment of mobile device users at sporting events based on audio from crowd reactions.

*Access a user ' s tendency to respond to an advertisement.

*Customize styles of ads served on mobile devices.

*Use video to detect penalty events.

Current environment

Requirements

* Media used for penalty event detection will be provided by consumer devices. Media may include images and videos captured during the sporting event and snared using social media. The images and videos will have varying sizes and formats.

* The data available for model building comprises of seven years of sporting event media. The sporting event media includes: recorded videos, transcripts of radio commentary, and logs from related social media feeds feeds captured during the sporting events.

*Crowd sentiment will include audio recordings submitted by event attendees in both mono and stereo Formats.

Advertisements

* Ad response models must be trained at the beginning of each event and applied during the sporting event.

* Market segmentation nxxlels must optimize for similar ad resporr.r history.

* Sampling must guarantee mutual and collective exclusivity local and global segmentation models that share the same features.

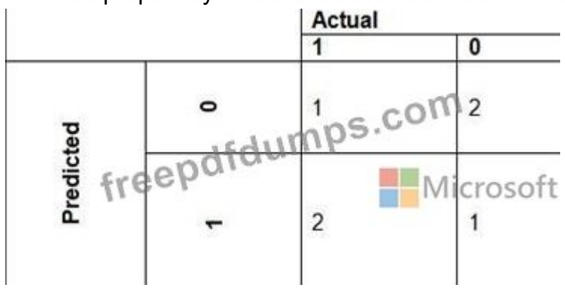
* Local market segmentation models will be applied before determining a user's propensity to respond to an advertisement.

* Data scientists must be able to detect model degradation and decay.

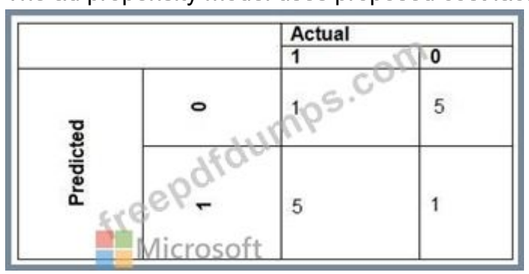
* Ad response models must support non linear boundaries features.

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

* The ad propensity model uses cost factors shown in the following diagram:



The ad propensity model uses proposed cost factors shown in the following diagram:



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

Penalty detection and sentiment

Findings

*Data scientists must build an intelligent solution by using multiple machine learning models for penalty event detection.

*Data scientists must build notebooks in a local environment using automatic feature engineering and model building in machine learning pipelines.

*Notebooks must be deployed to retrain by using Spark instances with dynamic worker allocation

*Notebooks must execute with the same code on new Spark instances to recode only the source of the data.

*Global penalty detection models must be trained by using dynamic runtime graph computation during training.

*Local penalty detection models must be written by using BrainScript.

* Experiments for local crowd sentiment models must combine local penalty detection data.

* Crowd sentiment models must identify known sounds such as cheers and known catch phrases. Individual crowd sentiment models will detect similar sounds.

* All shared features for local models are continuous variables.

* Shared features must use double precision. Subsequent layers must have aggregate running mean and standard deviation metrics Available.

segments

During the initial weeks in production, the following was observed:

*Ad response rates declined.

*Drops were not consistent across ad styles.

*The distribution of features across training and production data are not consistent.

Analysis shows that of the 100 numeric features on user location and behavior, the 47 features that come from location sources are being used as raw features. A suggested experiment to remedy the bias and variance issue is to engineer 10 linearly uncorrected features.

Penalty detection and sentiment

*Initial data discovery shows a wide range of densities of target states in training data used for crowd sentiment models.

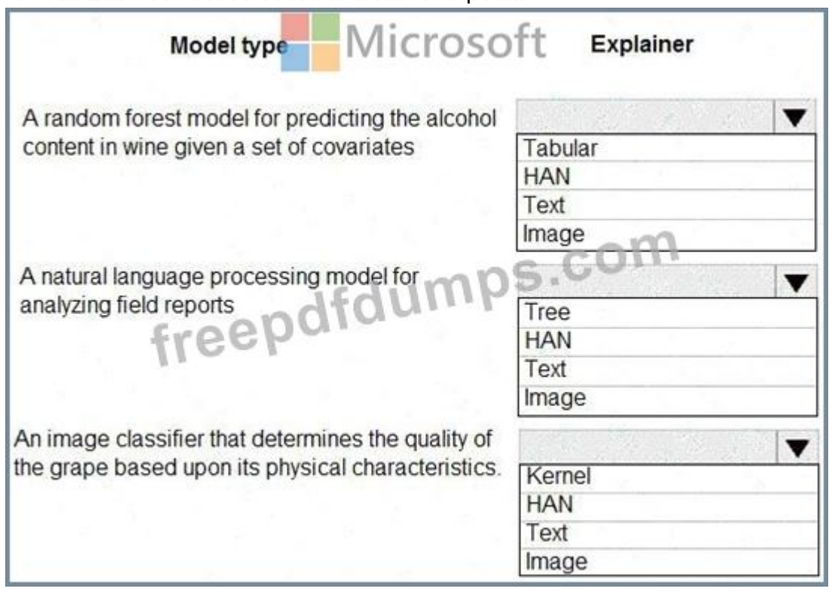
*All penalty detection models show inference phases using a Stochastic Gradient Descent (SGD) are running too stow.

*Audio samples show that the length of a catch phrase varies between 25%-47%, depending on region.

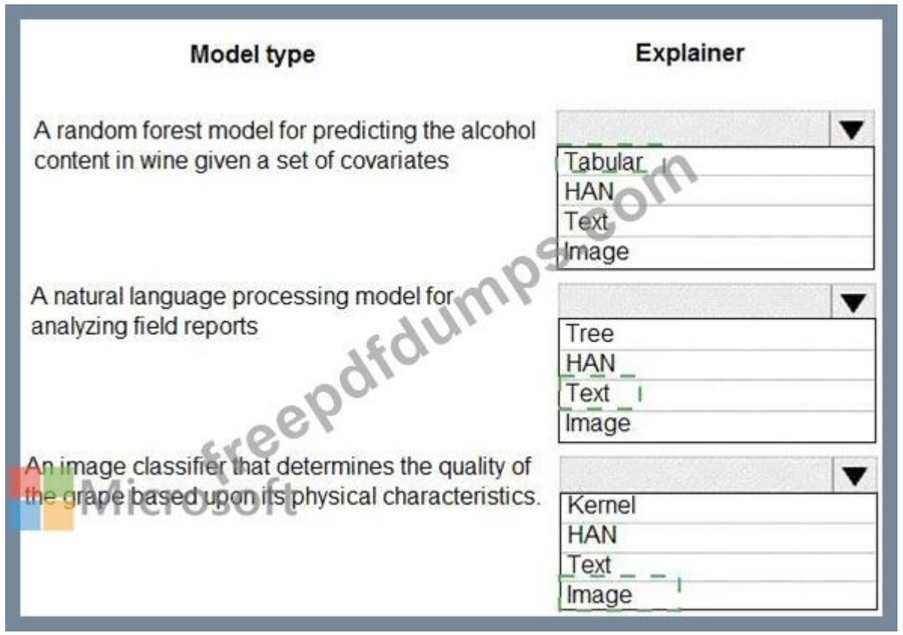
*The performance of the global penalty detection models show lower variance but higher bias when comparing training and validation sets. Before implementing any feature changes, you must confirm the bias and variance using all training and validation cases.

NEW QUESTION: 77
You plan to run a script as an experiment using a Script Run Configuration. The script uses modules from the scipy library as well as several Python packages that are not typically installed in a default conda environment.
You plan to run the experiment on your local workstation for small datasets and scale out the experiment by running it on more powerful remote compute clusters for larger datasets.
You need to ensure that the experiment runs successfully on local and remote compute with the least administrative effort.
What should you do?
A. Create and register an Environment that includes the required packages. Use this Environment for all experiment runs.
B. Always run the experiment with an Estimator by using the default packages.
C. Do not specify an environment in the run configuration for the experiment. Run the experiment by using the default environment.
D. Create a config. yaml file defining the conda packages that are required and save the file in the experiment folder.
E. Create a virtual machine (VM) with the required Python configuration and attach the VM as a compute target. Use this compute target for all experiment runs.
Answer: A (LEAVE A REPLY)
If you have an existing Conda environment on your local computer, then you can use the service to create an environment object. By using this strategy, you can reuse your local interactive environment on remote runs.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-use-environments

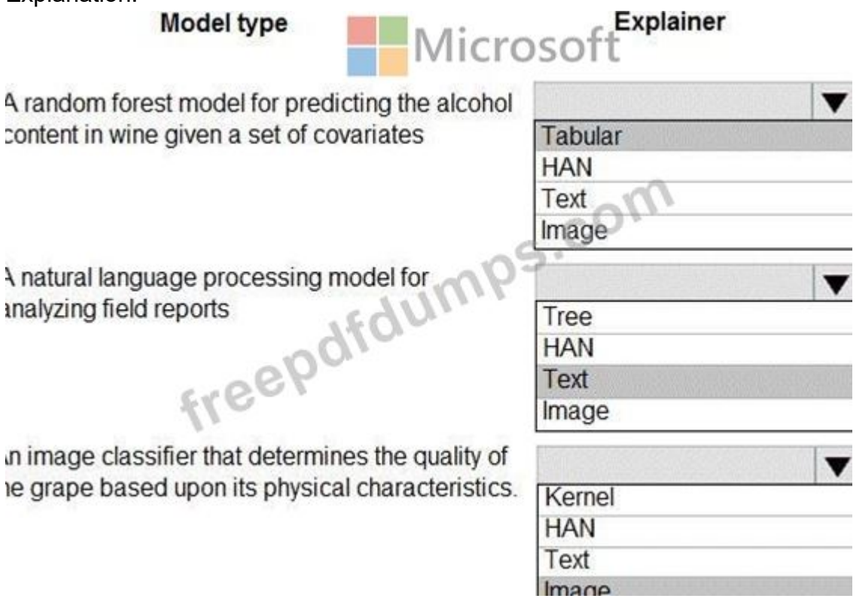
NEW QUESTION: 78
You are hired as a data scientist at a winery. The previous data scientist used Azure Machine Learning.
You need to review the models and explain how each model makes decisions.
Which explainer modules should you use? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.



Answer:



Explanation:



Meta explainers automatically select a suitable direct explainer and generate the best explanation info based on the given model and data sets. The meta explainers leverage all the libraries (SHAP, LIME, Mimic, etc.) that we have integrated or developed. The following are the meta explainers available in the SDK:

Tabular Explainer: Used with tabular datasets.

Text Explainer: Used with text datasets.

Image Explainer: Used with image datasets.

Box 1: Tabular
Box 2: Text
Box 3: Image
Reference:
https://medium.com/microsoftazure/automated-and-interpretable-machine-learning-d07975741298

NEW QUESTION: 79
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_table('Label Values', label_vals)
```

Does the solution meet the goal?
A. Yes
B. No
Answer: (SHOW ANSWER)
Instead use the run_log function to log the contents in label_vals:

```
for label_val in label_vals:
run.log('Label Values', label_val)
```

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

NEW QUESTION: 80
You create a binary classification model by using Azure Machine Learning Studio.
You must tune hyperparameters by performing a parameter sweep of the model. The parameter sweep must meet the following requirements:
* iterate all possible combinations of hyperparameters
* minimize computing resources required to perform the sweep
* You need to perform a parameter sweep of the model.
Which parameter sweep mode should you use?
A. Random sweep

B. Sweep clustering
C. Entire grid
D. Random grid
E. Random seed
Answer: D (LEAVE A REPLY)
Maximum number of runs on random grid: This option also controls the number of iterations over a random sampling of parameter values, but the values are not generated randomly from the specified range; instead, a matrix is created of all possible combinations of parameter values and a random sampling is taken over the matrix. This method is more efficient and less prone to regional oversampling or undersampling.
If you are training a model that supports an integrated parameter sweep, you can also set a range of seed values to use and iterate over the random seeds as well. This is optional, but can be useful for avoiding bias introduced by seed selection.

NEW QUESTION: 81
You create a binary classification model.
You need to evaluate the model performance.
Which two metrics can you use? Each correct answer presents a complete solution.
NOTE: Each correct selection is worth one point.
A. relative absolute error
B. precision
C. accuracy
D. mean absolute error
E. coefficient of determination
Answer: B,C (LEAVE A REPLY)
The evaluation metrics available for binary classification models are: Accuracy, Precision, Recall, F1 Score, and AUC.
Note: A very natural question is: 'Out of the individuals whom the model, how many were classified correctly (TP)?' This question can be answered by looking at the Precision of the model, which is the proportion of positives that are classified correctly.
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio/evaluate-model-performance

NEW QUESTION: 82
You register a file dataset named csvjolder that references a folder. The folder includes multiple com ma- separated values (CSV) files in an Azure storage blob container. You plan to use the following code to run a script that loads data from the file dataset. You create and instantiate the following variables:

Variable	Description
remote_cluster	References the Azure Machine Learning compute cluster
ws	References the Azure Machine Learning workspace

You have the following code:

```
from azureml.train.estimator import Estimator
file_dataset = ws.datasets.get('csv_folder')
estimator = Estimator(source_directory=script_folder,
```

```
compute_target = remote_cluster,
entry_script='script.py')
run = experiment.submit(config=estimator)
run.wait_for_completion(show_output=True)
```

You need to pass the dataset to ensure that the script can read the files it references. Which code segment should you insert to replace the code comment?

A.
```
inputs=[file_dataset.as_named_input('training_files').to_pandas_dataframe()],
```

B.
```
inputs=[file_dataset.as_named_input('training_files').as_mount()],
```

C.
```
script_params={'--training_files': file_dataset},
```

D.
```
inputs=[file_dataset.as_named_input('training_files')],
```

Answer: D (LEAVE A REPLY)

Example:

from azureml.train.estimator import Estimator

script_params = {

to mount files referenced by mnist dataset

' --data-folder ' : mnist_file_dataset.as_named_input(' mnist_opendataset ').as_mount(),

' --regularization ' : 0.5

}

est = Estimator(source_directory=script_folder,

script_params=script_params,

compute_target=compute_target,

environment_definition=env,

entry_script= ' train.py ')

Reference:

https://docs.microsoft.com/en-us/azure/machine-learning/tutorial-train-models-with-aml

NEW QUESTION: 83

You have an Azure Machine Learning workspace.

You plan to set up logging and tracking experiments by using MLflow Tracking.

You need to log the accuracy as a numerical value and the training loss as a plot.

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

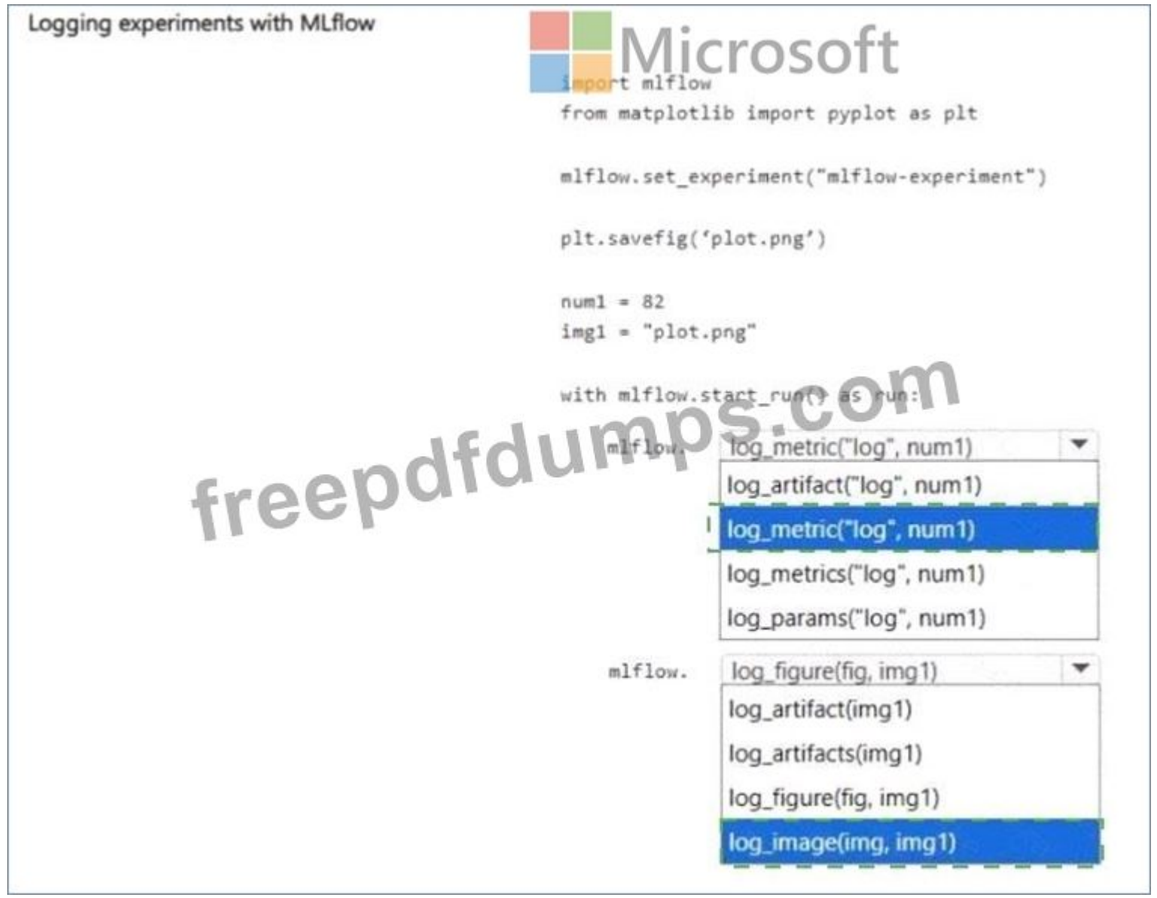
NOTE: Each correct selection is worth one point.

Logging experiments with MLflow

```
import mlflow
from matplotlib import pyplot as plt

mlflow.set_experiment("mlflow-experiment")

plt.savefig('plot.png')

num1 = 82
img1 = "plot.png"

with mlflow.start_run() as run:
    mlflow.[log_metric("log", num1)]
    mlflow.[log_figure(fig, img1)]
```

- log_metric("log", num1)
  - log_artifact("log", num1)
  - log_metric("log", num1)
  - log_metrics("log", num1)
  - log_params("log", num1)
- log_figure(fig, img1)
  - log_artifact(img1)
  - log_artifacts(img1)
  - log_figure(fig, img1)
  - log_image(img, img1)

Answer:



Explanation:

NEW QUESTION: 84

You create a script for training a machine learning model in Azure Machine Learning service.

You create an estimator by running the following code:

```
from azureml.core import Workspace, Datastore
from azureml.core.compute import ComputeTarget
from azureml.train.estimator import Estimator
work_space = Workspace.from_config()
data_source = work_space.get_default_datastore()
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.

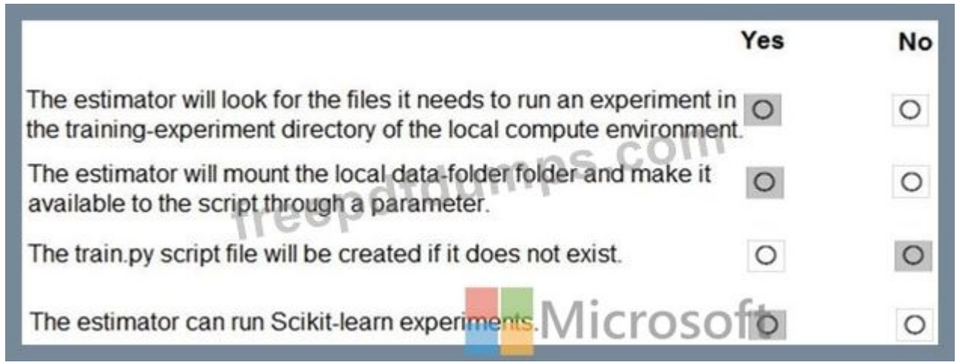
NOTE: Each correct selection is worth one point.

	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.	○	○
The estimator can run Scikit-learn experiments.	○	○

Answer:

	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.	○	○
The estimator can run Scikit-learn experiments.	○	○

Explanation:



Box 1: Yes

Parameter source_directory is a local directory containing experiment configuration and code files needed for a training job.

Box 2: Yes

script_params is a dictionary of command-line arguments to pass to the training script specified in entry_script.

Box 3: No

Box 4: Yes

The conda_packages parameter is a list of strings representing conda packages to be added to the Python environment for the experiment.

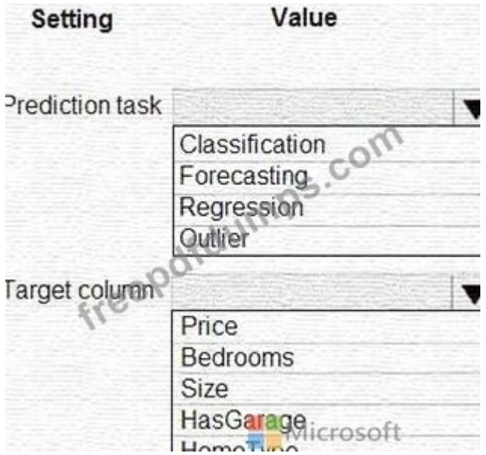
NEW QUESTION: 85

You have a dataset that includes home sales data for a city. The dataset includes the following columns.

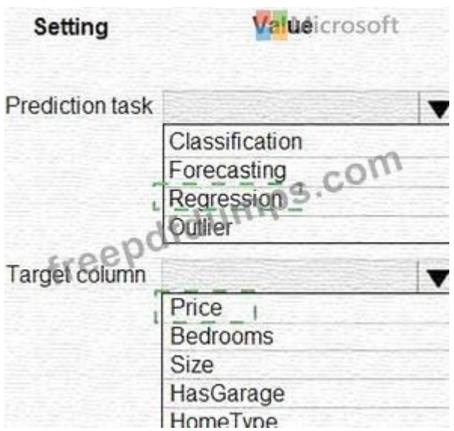
Name	Description
Price	The sales price for the house.
Bedrooms	The number of bedrooms in the house.
Size	The size of the house in square feet.
HasGarage	A binary value indicating whether or not the house has a garage.
HomeType	The category of home, for example, apartment, townhouse, single-family home.

Each row in the dataset corresponds to an individual home sales transaction.

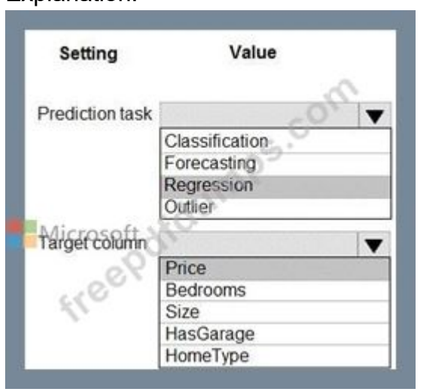
You need to use automated machine learning to generate the best model for predicting the sales price based on the features of the house.
Which values should you use? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.



Answer:



Explanation:



Box 1: Regression

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

NEW QUESTION: 86
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 (LEAVE A REPLY)
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: 87

You have the following code. The code prepares an experiment to run a script:

```
from azureml.core import Workspace, Experiment, Run, ScriptRunConfig

ws = Workspace.from_config()
script_config = ScriptRunConfig(source_directory='experiment_files',
                                script='experiment.py')

script_experiment = Experiment(workspace=ws, name='script-experiment')
```


The experiment must be run on local computer using the default environment.

You need to add code to start the experiment and run the script.

Which code segment should you use?

A. run = script_experiment.start_logging()

B. run = Run(experiment=script_experiment)

C. ws.get_run(run_id=experiment.id)

D. run = script_experiment.submit(config=script_config)

Answer: D (LEAVE A REPLY)

The experiment class submit method submits an experiment and return the active created run.

Syntax: submit(config, tags=None, **kwargs)

Reference:

https://docs.microsoft.com/en-us/python/api/azureml-core/azureml.core.experiment.experiment

NEW QUESTION: 88

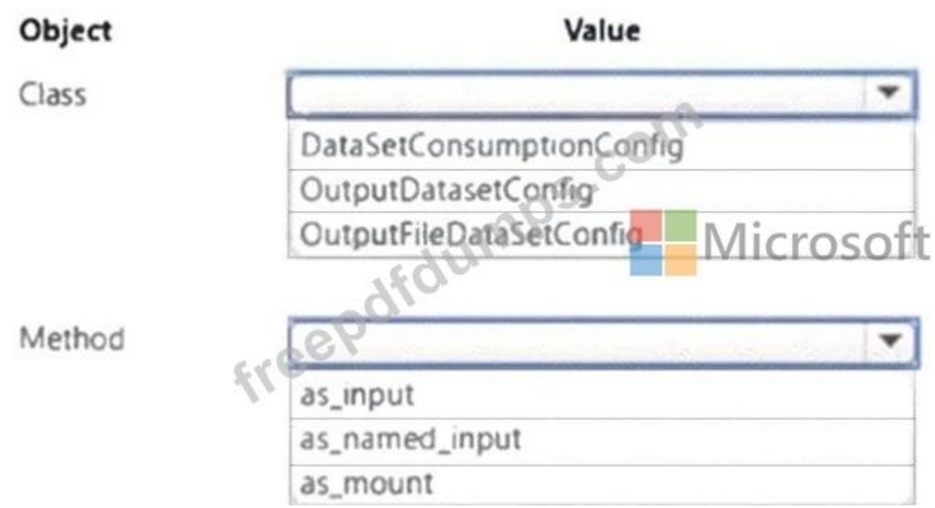
You plan to implement a two-step pipeline by using the Azure Machine Learning SDK for Python.

The pipeline will pass temporary data from the first step to the second step.

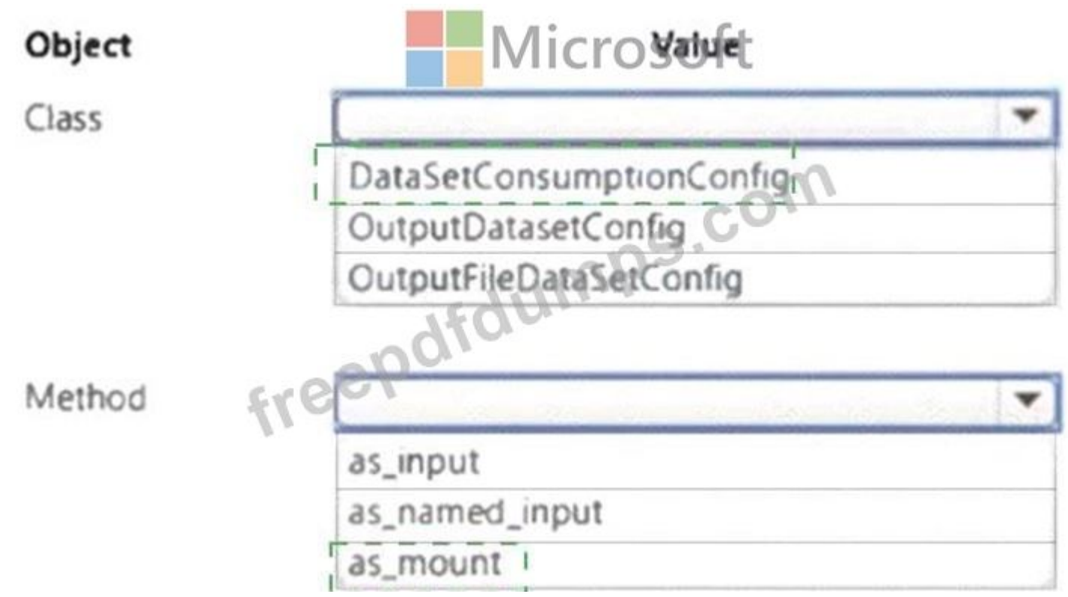
You need to identify the class and the corresponding method that should be used in the second step to access temporary data generated by the first step in the pipeline.

Which class and method should you identify? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point



Answer:



Explanation:

NEW QUESTION: 89

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.
After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.
You create an Azure Machine Learning service datastore in a workspace. The datastore contains the following files:
* /data/2018/Q1 .csv
* /data/2018/Q2.csv
* /data/2018/Q3.csv
* /data/2018/Q4.csv
* /data/2019/Q1.csv
All files store data in the following format:
id,M,f2,I
1,1,2,0
2,1,1,1
32,10
You run the following code:

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

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

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

Solution: Run the following code:

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

Does the solution meet the goal?
A. No
B. Yes
Answer: (SHOW ANSWER)

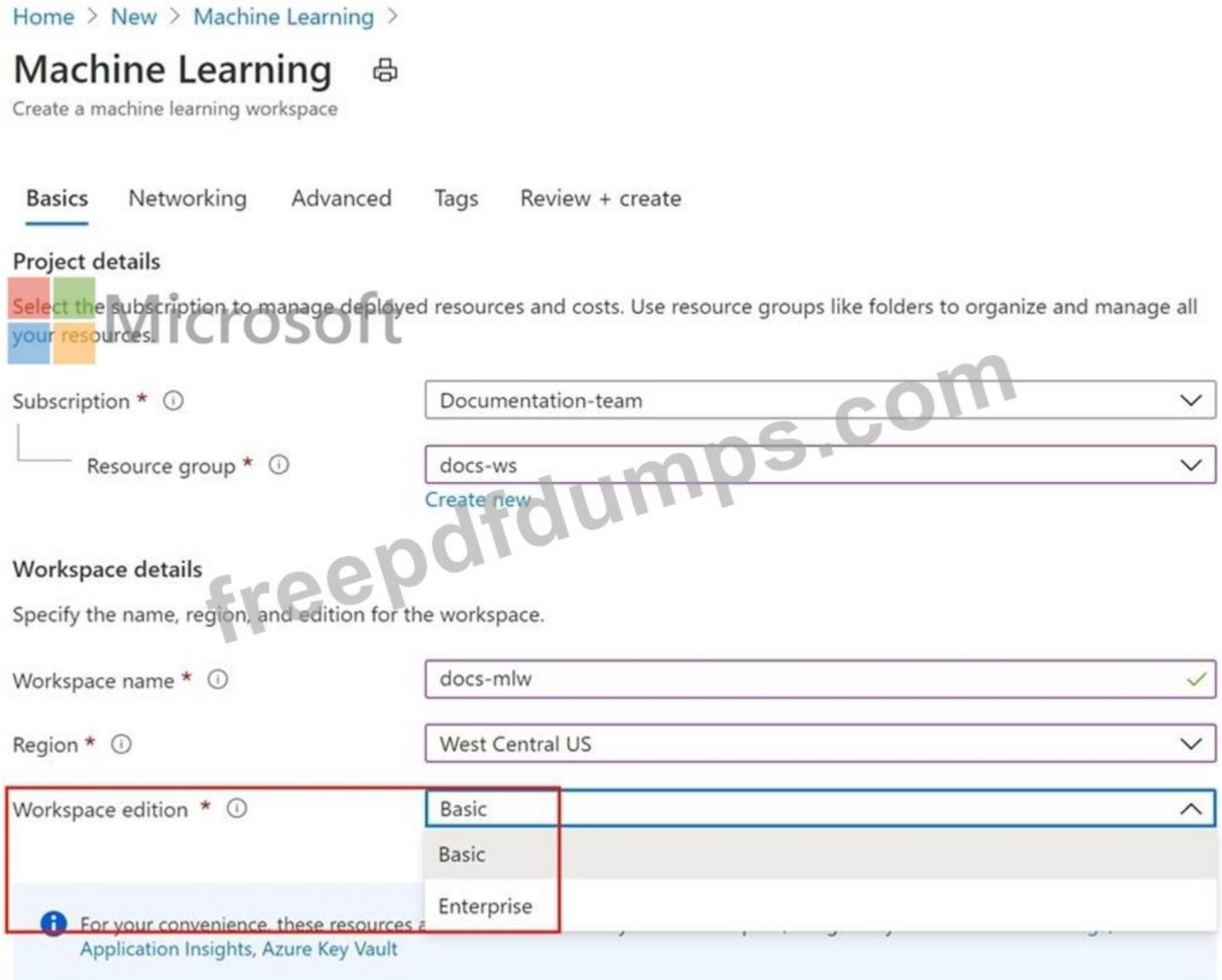
NEW QUESTION: 90
You are a data scientist creating a linear regression model.
You need to determine how closely the data fits the regression line.
Which metric should you review?
A. Coefficient of determination
B. Recall
C. Precision
D. Mean absolute error
E. Root Mean Square Error

Answer: A (LEAVE A REPLY)
Coefficient of determination, often referred to as R2, represents the predictive power of the model as a value between 0 and 1. Zero means the model is random (explains nothing); 1 means there is a perfect fit. However, caution should be used in interpreting R2 values, as low values can be entirely normal and high values can be suspect.
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/evaluate-model

NEW QUESTION: 91
You are developing deep learning models to analyze semi-structured, unstructured, and structured data types.
You have the following data available for model building:
Video recordings of sporting events
Transcripts of radio commentary about events
Logs from related social media feeds captured during sporting events
You need to select an environment for creating the model.
Which environment should you use?
A. Azure Cognitive Services
B. Azure Data Lake Analytics
C. Azure HDInsight with Spark MLib
D. Azure Machine Learning Studio
Answer: A (LEAVE A REPLY)
Azure Cognitive Services expand on Microsoft's evolving portfolio of machine learning APIs and enable developers to easily add cognitive features - such as emotion and video detection; facial, speech, and vision recognition; and speech and language understanding - into their applications. The goal of Azure Cognitive Services is to help developers create applications that can see, hear, speak, understand, and even begin to reason. The catalog of services within Azure Cognitive Services can be categorized into five main pillars - Vision, Speech, Language, Search, and Knowledge.
References:
https://docs.microsoft.com/en-us/azure/cognitive-services/welcome

NEW QUESTION: 92
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 azgroup create function with --name and --location parameters, and then the az ml workspace createfunction, 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: B,C,D (LEAVE A REPLY)
B: You can use an Azure Resource Manager template to create a workspace for Azure Machine Learning.
Example:
{"type": "Microsoft.MachineLearningServices/workspaces",
...
C: You can create a workspace for Azure Machine Learning with Azure CLI Install the machine learning extension.
Create a resource group: az group create --name <resource-group-name> --location <location> To create a new workspace where the services are automatically created, use the following command: az ml workspace create -w <workspace-name> -g <resource-group-name> D: You can create and manage Azure Machine Learning workspaces in the Azure portal.
Sign in to the Azure portal by using the credentials for your Azure subscription.
In the upper-left corner of Azure portal, select + Create a resource.
Use the search bar to find Machine Learning.
Select Machine Learning.
In the Machine Learning pane, select Create to begin.

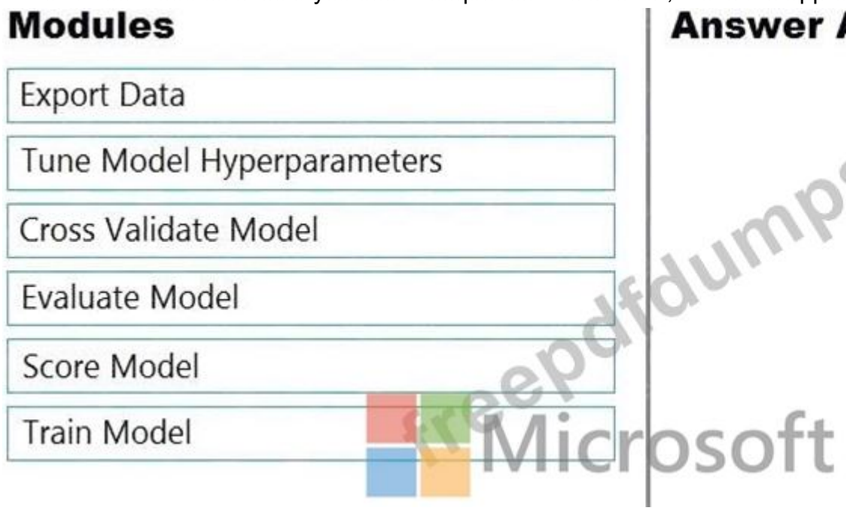


Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-create-workspace-template
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-manage-workspace-cli
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-manage-workspace

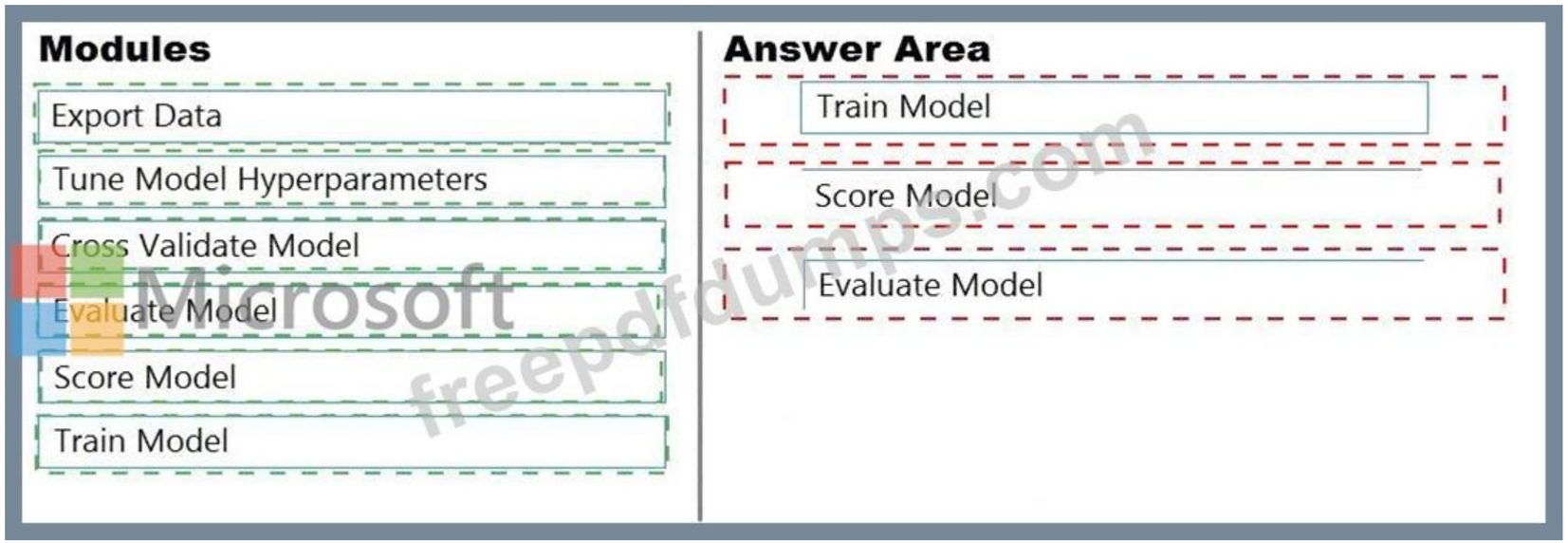
NEW QUESTION: 93

You need to select an environment that will meet the business and data requirements.
Which environment should you use?
A. Azure Cognitive Services
B. Azure HDInsight with Spark MLlib
C. Azure Machine Learning Studio
D. Microsoft Machine Learning Server
Answer: (SHOW ANSWER)

NEW QUESTION: 94
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.



Answer:



Explanation:

Step 1: Train Model

Two-Class Boosted Decision Tree

First, set up the boosted decision tree model.

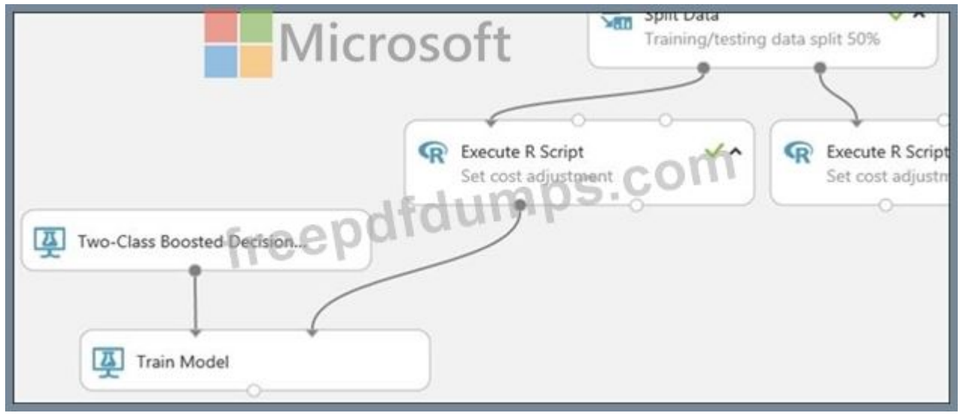
1. Find the Two-Class Boosted Decision Tree module in the module palette and drag it onto the canvas.

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:



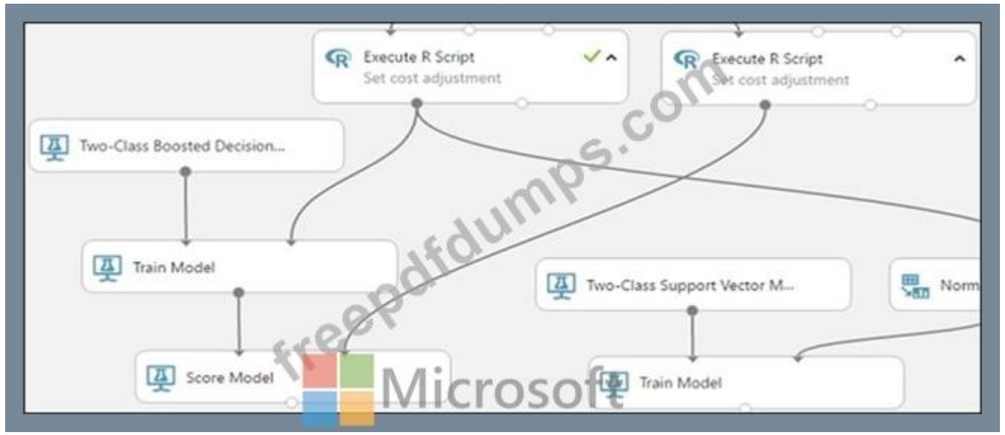
Step 2: Score Model

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

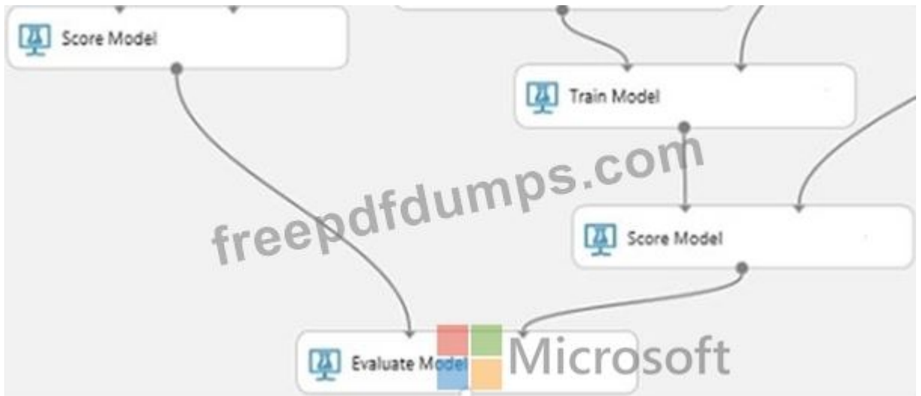
1. Find the Score Model module and drag it onto the canvas.
2. Connect the Train Model module that ' s connected to the Two-Class Boosted Decision Tree module to the left input port of the Score Model module.
3. Connect the right Execute R Script module (our testing data) to the right input port of the 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: 95

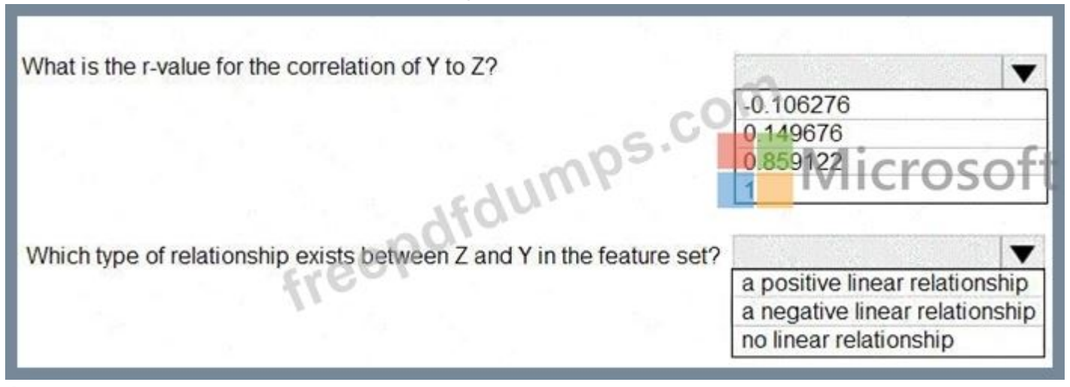
You have a feature set containing the following numerical features: X, Y, and Z.

The Poisson correlation coefficient (r-value) of X, Y, and Z features is shown in the following image:

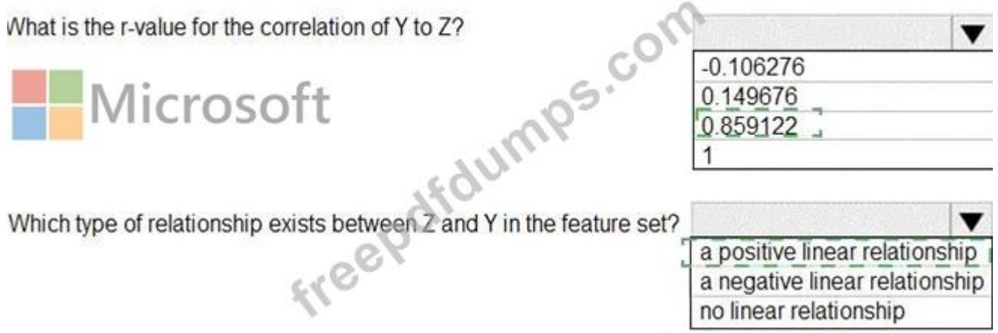
	X	Y	Z
X	1	0.149676	-0.106276
Y	0.149676	1	0.859122
Z	-0.106276	0.859122	1

Use the drop-down menus to select the answer choice that answers each question based on the information presented in the graphic.

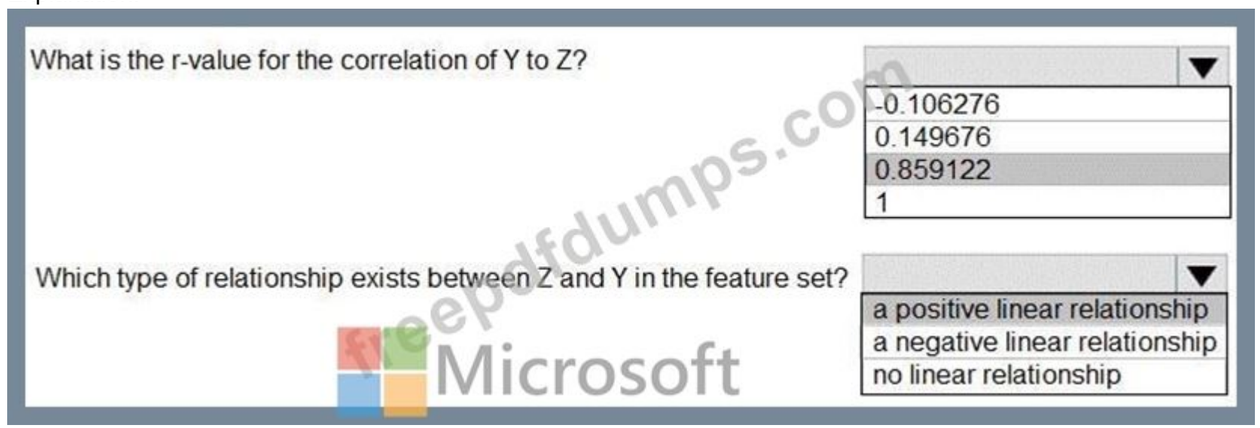
NOTE: Each correct selection is worth one point.



Answer:



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.

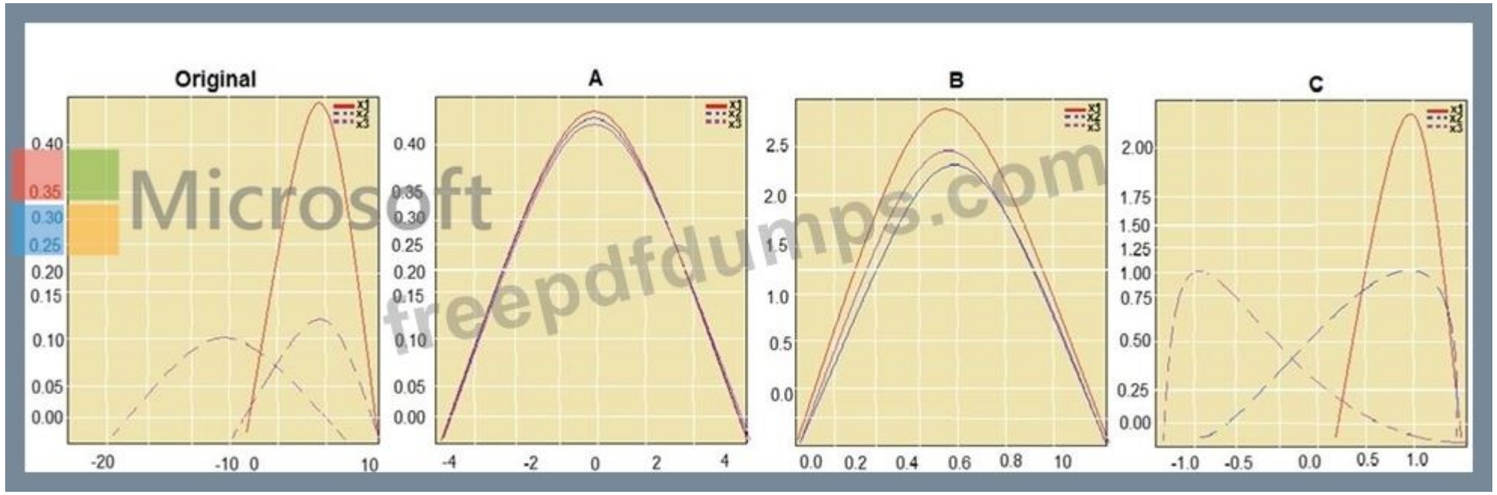
References:

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

NEW QUESTION: 96

You are performing feature scaling by using the scikit-learn Python library for x.1 x2, and x3 features.

Original and scaled data is shown in the following image.

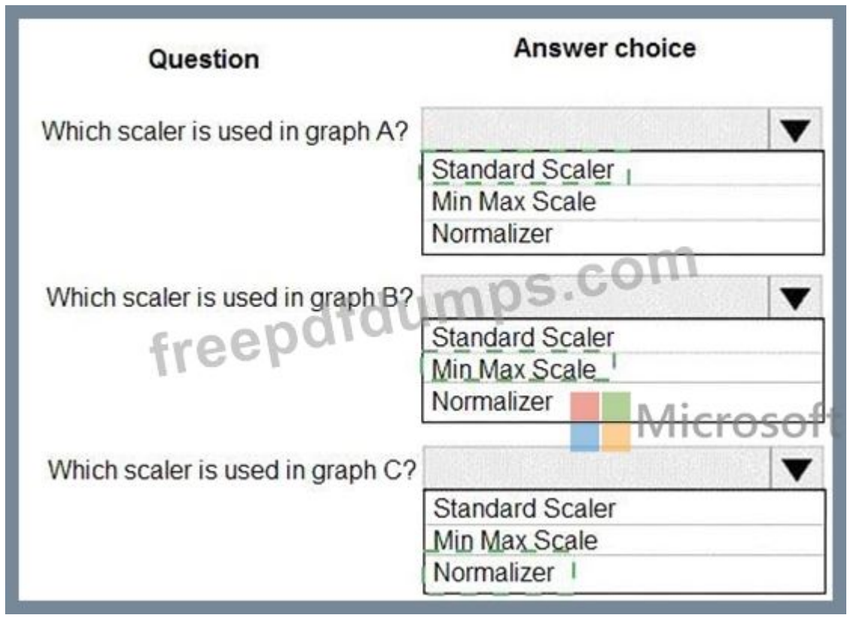


Use the drop-down menus to select the answer choice that answers each question based on the information presented in the graphic.

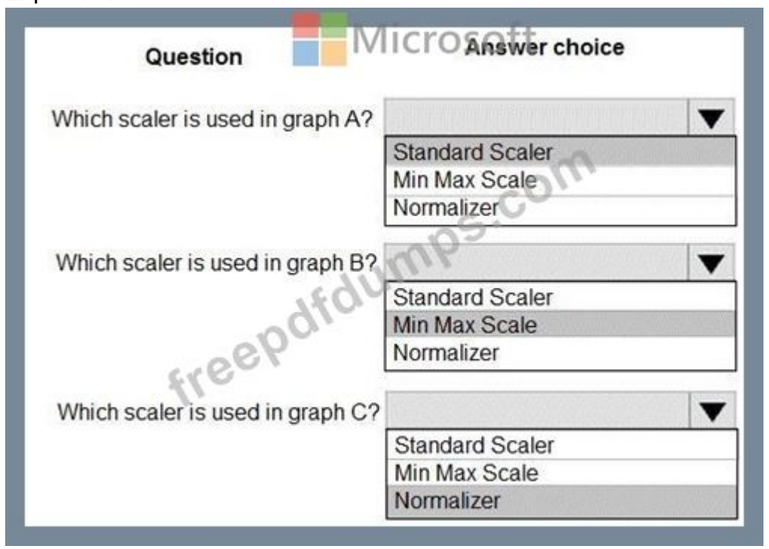
NOTE: Each correct selection is worth one point.

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

Answer:



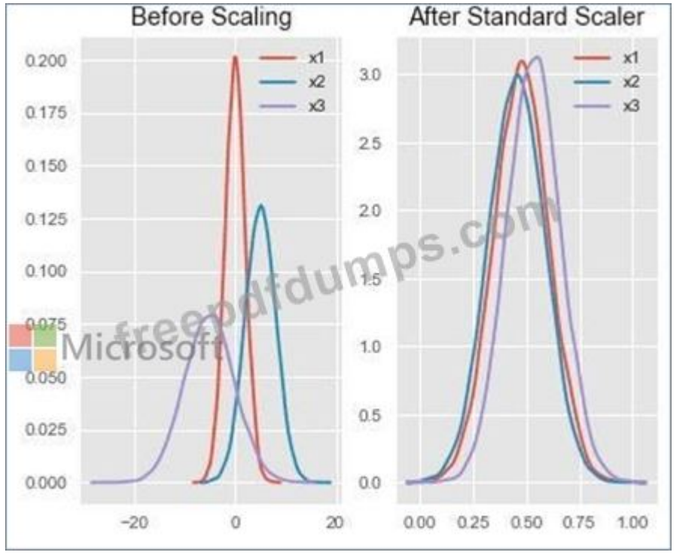
Explanation:



Box 1: StandardScaler

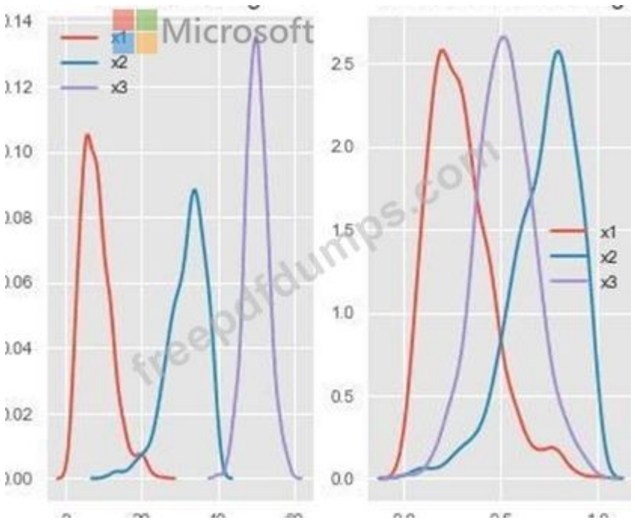
The StandardScaler assumes your data is normally distributed within each feature and will scale them such that the distribution is now centred around 0, with a standard deviation of 1.

Example:



All features are now on the same scale relative to one another.

Box 2: Min Max Scaler



Notice that the skewness of the distribution is maintained but the 3 distributions are brought into the same scale so that they overlap.

Box 3: Normalizer

References:

http://benalexkeen.com/feature-scaling-with-scikit-learn/

NEW QUESTION: 97

You have an Azure Machine Learning workspace.

You plan to use Azure Machine Learning Python SDK v2 to register a component in the workspace The component definition is stored in the local file ./components/train/train.yml.

You write code to connect to the workspace by using the ml_client object and import all required libraries You need to complete the remaining code.

How should you complete the code? to answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.

Answer:

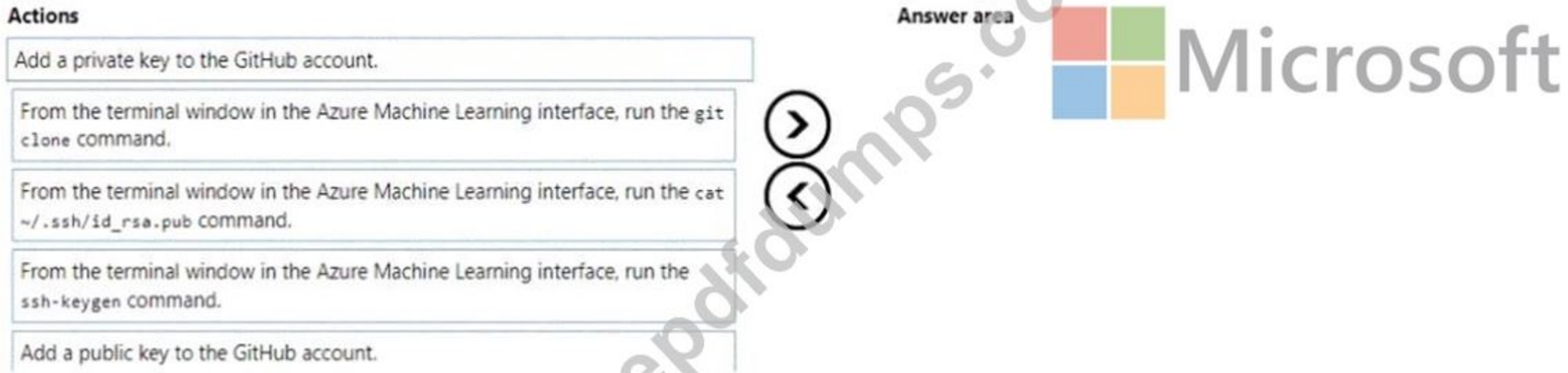
Explanation:

NEW QUESTION: 98
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 normalization with a QuantileIndex normalization.
Does the solution meet the GOAL?
A. Yes

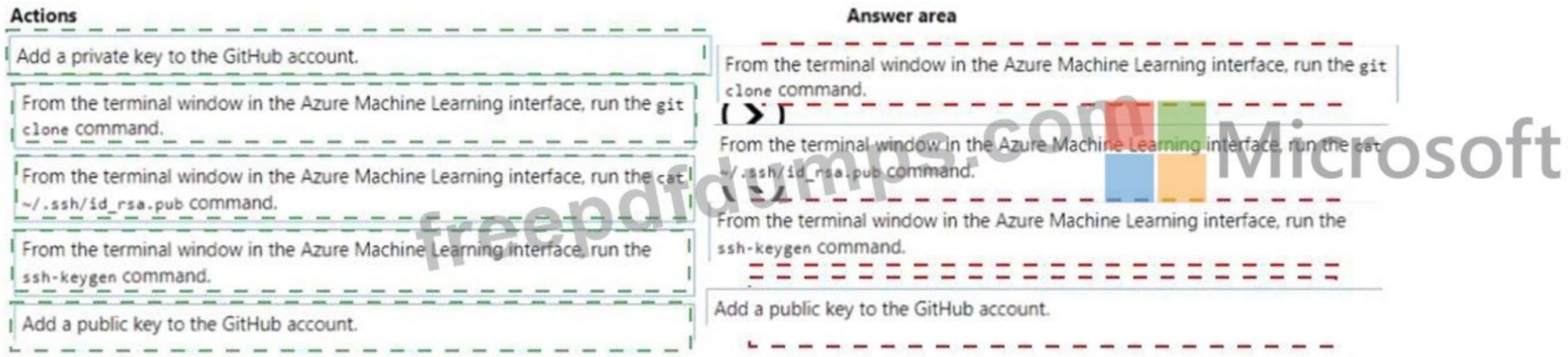
B. No
Answer: B (LEAVE A REPLY)
Use the Entropy MDL binning mode which has a target column.
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/group-data-into-bins

NEW QUESTION: 99
You manage an Azure Machine Learning workspace. You develop a regression model training pipeline by using Notebooks. You need to determine the appropriate evaluation metric for the experiment.
Which two metrics should you choose? Each correct answer presents a complete solution. Choose two.
NOTE: Each correct selection is worth one point.
A. recall
B. residuals
C. correlation
D. precision
Answer: (SHOW ANSWER)

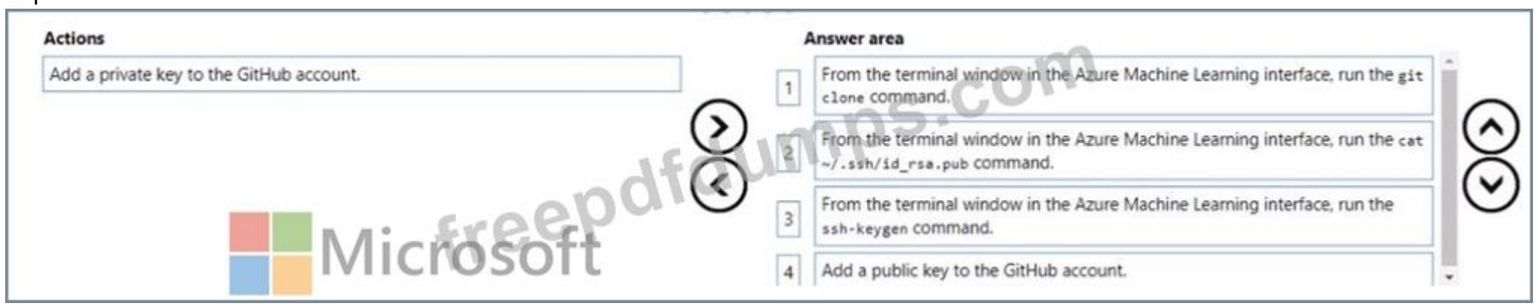
NEW QUESTION: 100
You have an existing GitHub repository containing Azure Machine Learning project files.
You need to clone the repository to your Azure Machine Learning shared workspace file system.
Which four actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.
NOTE: More than one order of answer choices is correct. You will receive credit for any of the correct orders you select.



Answer:



Explanation:



NEW QUESTION: 101

You create a script that trains a convolutional neural network model over multiple epochs and logs the validation loss after each epoch. The script includes arguments for batch size and learning rate.

You identify a set of batch size and learning rate values that you want to try.

You need to use Azure Machine Learning to find the combination of batch size and learning rate that results in the model with the lowest validation loss.

What should you do?

A. Run the script in an experiment based on an AutoMLConfig object

B. Create a PythonScriptStep object for the script and run it in a pipeline

C. Use the Automated Machine Learning interface in Azure Machine Learning studio

D. Run the script in an experiment based on a ScriptRunConfig object

E. Run the script in an experiment based on a HyperDriveConfig object

Answer: (SHOW ANSWER)

Reference:

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

NEW QUESTION: 102

You create a Python script that runs a training experiment in Azure Machine Learning. The script uses the Azure Machine Learning SDK for Python.

You must add a statement that retrieves the names of the logs and outputs generated by the script.

You need to reference a Python class object from the SDK for the statement.
Which class object should you use?
A. Run
B. ScripcRunConfig
C. Workspace
D. Experiment
Answer: A (LEAVE A REPLY)
A run represents a single trial of an experiment. Runs are used to monitor the asynchronous execution of a trial, log metrics and store output of the trial, and to analyze results and access artifacts generated by the trial.
The run Class get_all_logs method downloads all logs for the run to a directory.
Reference:
https://docs.microsoft.com/en-us/python/api/azureml-core/azureml.core.run(class)

NEW QUESTION: 103
You have an Azure Machine Learning workspace.
You plan to use the workspace to set up automated machine learning training for an image classification model.
You need to choose the primary metric to optimize the model training.
Which primary metric should you choose?
A. root_mean_squared_log_error
B. r2_score
C. mean_absolute_error
D. accuracy
Answer: D (LEAVE A REPLY)

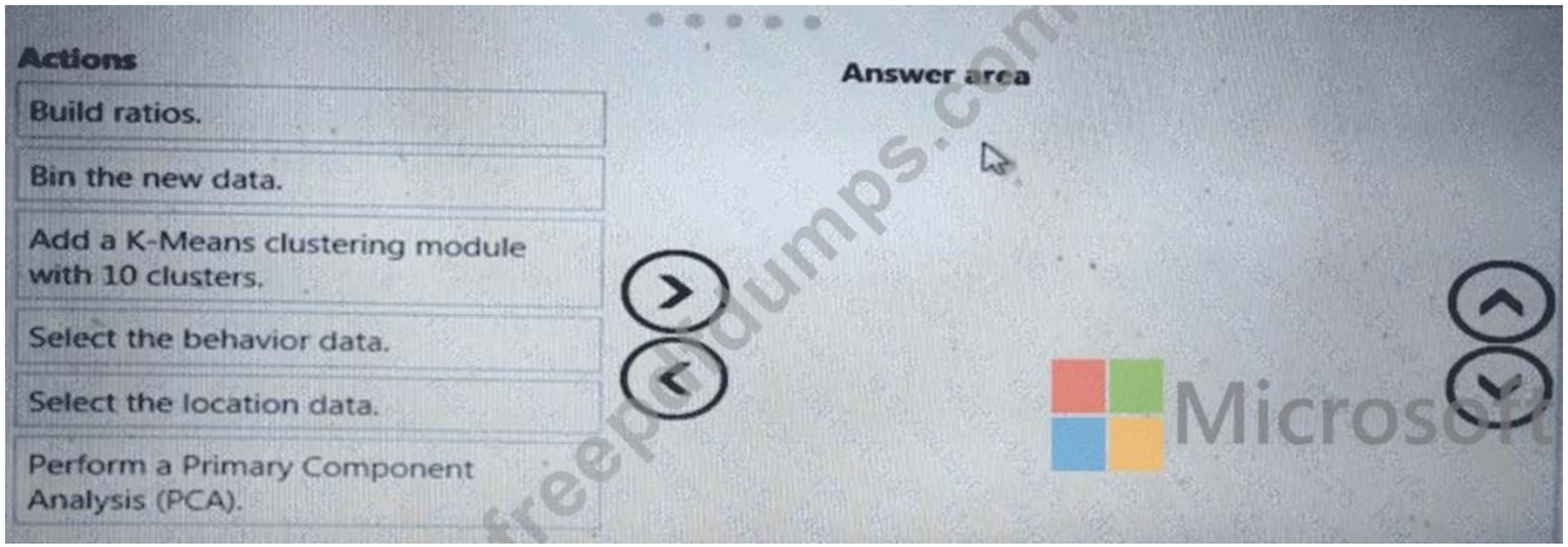
NEW QUESTION: 104
You use the Azure Machine Learning Python SDK to create a batch inference pipeline.
You must publish the batch inference pipeline so that business groups in your organization can use the pipeline. Each business group must be able to specify a different location for the data that the pipeline submits to the model for scoring.
You need to publish the pipeline.
What should you do?
A. Define a PipelineParameter object for the pipeline and use it to specify the business group-specific input dataset for each pipeline run.
B. Create multiple endpoints for the published pipeline service and have each business group submit jobs to its own endpoint.
C. Have each business group run the pipeline on local compute and use a local file for the input data.
D. Define a OutputFileDatasetConfig object for the pipeline and use the object to specify the business group-specific input dataset for each pipeline run.
Answer: D (LEAVE A REPLY)

NEW QUESTION: 105
You use the Azure Machine learning SDK foe Python to create a pipeline that includes the following step:
The output of the step run must be cached and reused on subsequent runs when the source.directory value has not changed.
You need to define the step.
What should you include in the step definition?
A. hash_path

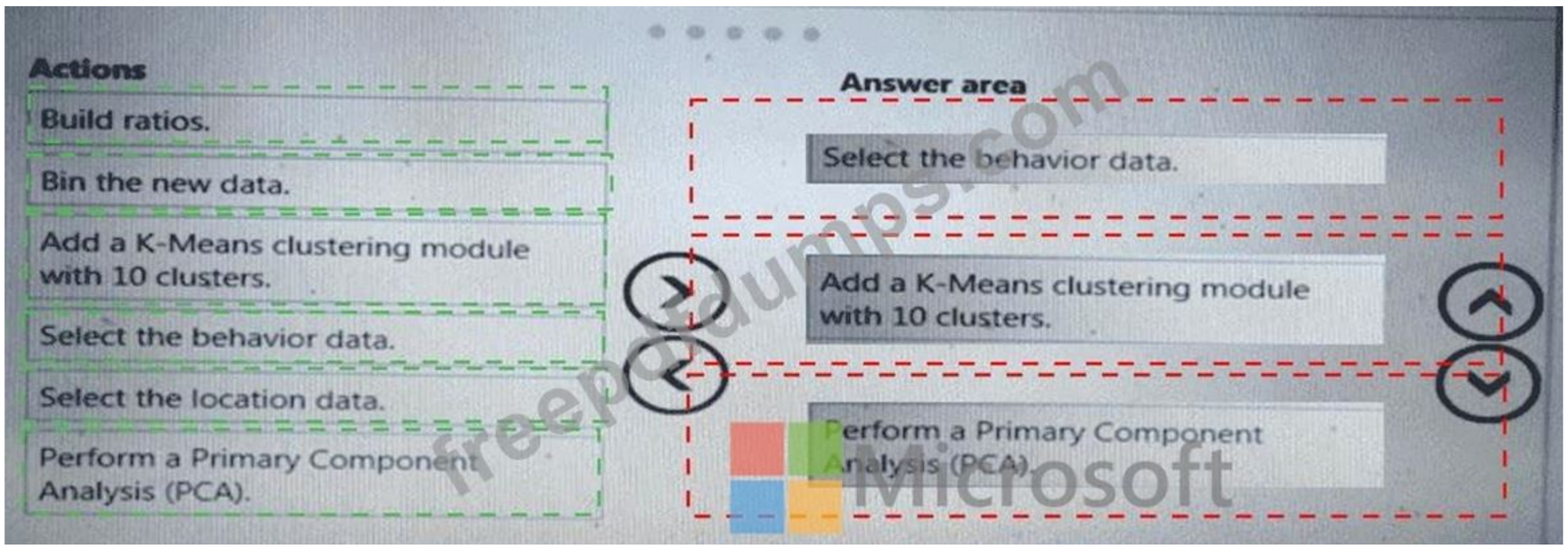
B. version
C. allow.reuse
D. data-as_input(name-..)
Answer: C (LEAVE A REPLY)

NEW QUESTION: 106
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 an Azure Machine Learning workspace. You connect to a terminal session from the Notebooks page in Azure Machine Learning studio.
You plan to add a new Jupyter kernel that will be accessible from the same terminal session.
You need to perform the task that must be completed before you can add the new kernel.
Solution: Create an environment.
Does the solution meet the goal?
A. No
B. Yes
Answer: B (LEAVE A REPLY)

NEW QUESTION: 107
You need to modify the inputs for the global penalty event model to address the bias and variance 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.



Answer:



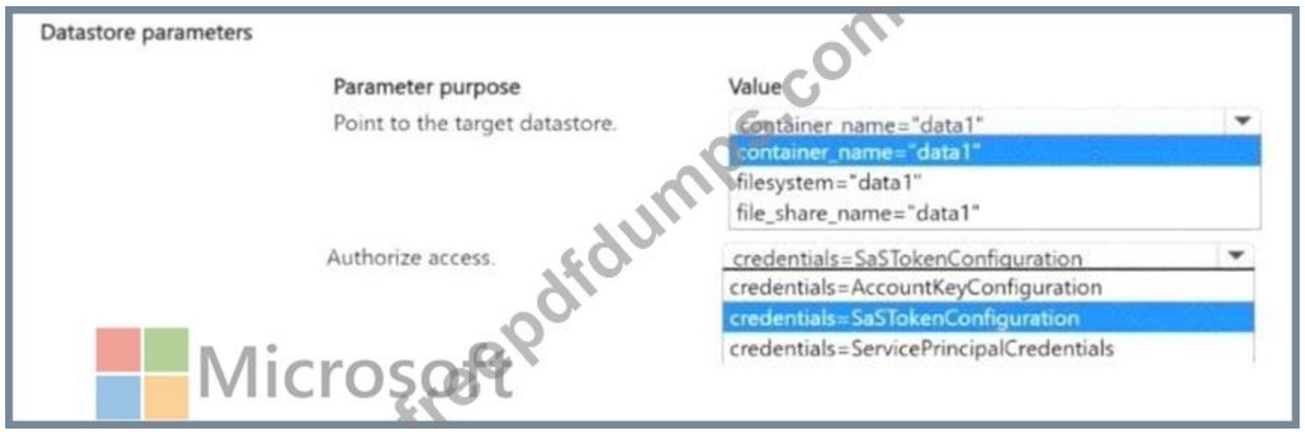
NEW QUESTION: 108

You manage an Azure Machine Learning workspace named Workspace1 and an Azure Blob Storage accessed by using the URL https://storage1.blob.core.wmdows.net/data1.

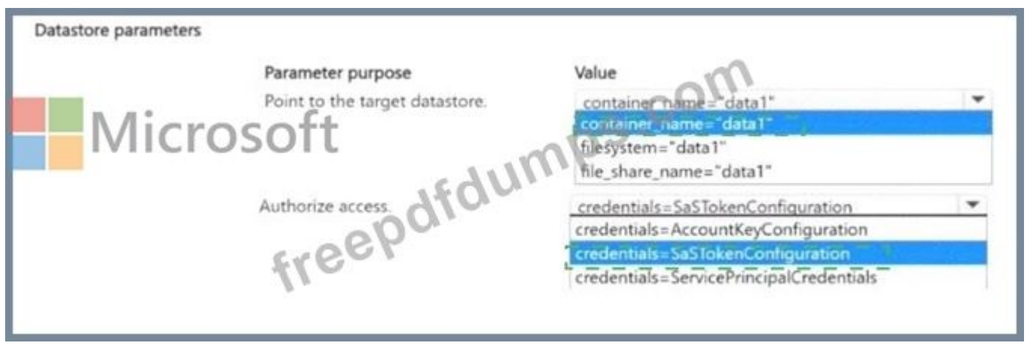
You plan to create an Azure Blob datastore in Workspace1. The datastore must target the Blob Storage by using Azure Machine Learning Python SDK v2. Access authorization to the datastore must be limited to a specific amount of time.

You need to select the parameters of the Azure Blob Datastore class that will point to the target datastore and authorize access to it.

Which parameters should you use? To answer, select the appropriate options in the answer area NOTE: Each correct selection is worth one point.



Answer:



Explanation:

NEW QUESTION: 109

You manage an Azure Machine Learning workspace. The Pylhon scrip! named scriptpy reads an argument named training_data. The trainlng.data argument specifies the path to the training data in a file named datasetl. csv.

You plan to run the scriptpy Python script as a command job that trains a machine learning model.

You need to provide the command to pass the path for the datasct as a parameter value when you submit the script as a training job.

Solution: python train.py --training_data training_data

Does the solution meet the goal?

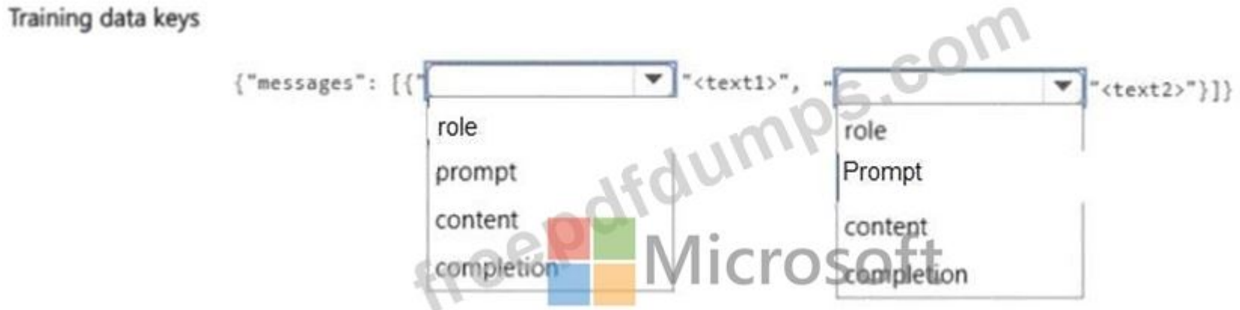
A. No

B. Yes

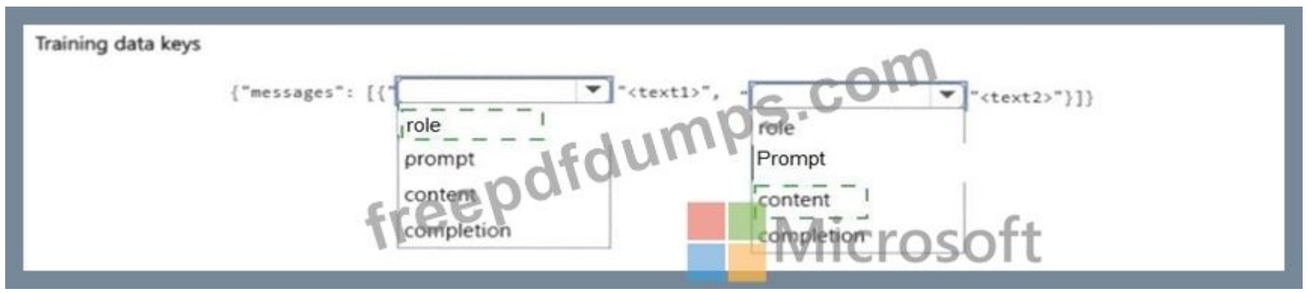
Answer: A (LEAVE A REPLY)

NEW QUESTION: 110
You have a comma-separated values (CSV) file containing data from which you want to train a classification model.
You are using the Automated Machine Learning interface in Azure Machine Learning studio to train the classification model. You set the task type to Classification.
You need to ensure that the Automated Machine Learning process evaluates only linear models.
What should you do?
A. Add all algorithms other than linear ones to the blocked algorithms list.
B. Set the Exit criterion option to a metric score threshold.
C. Clear the option to perform automatic featurization.
D. Clear the option to enable deep learning.
E. Set the task type to Regression.
Answer: C (LEAVE A REPLY)
Automatic featurization can fit non-linear models.
Reference:
https://econml.azurewebsites.net/spec/estimation/dml.html
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-use-automated-ml-for-ml-models

NEW QUESTION: 111
You manage an Azure OpenAI deployment of the gpt-4o base model.
You plan to fine-tune the deployed model.
You need to prepare a file that contains training data.
Which keys should you include in each line of the training data file? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.



Answer:



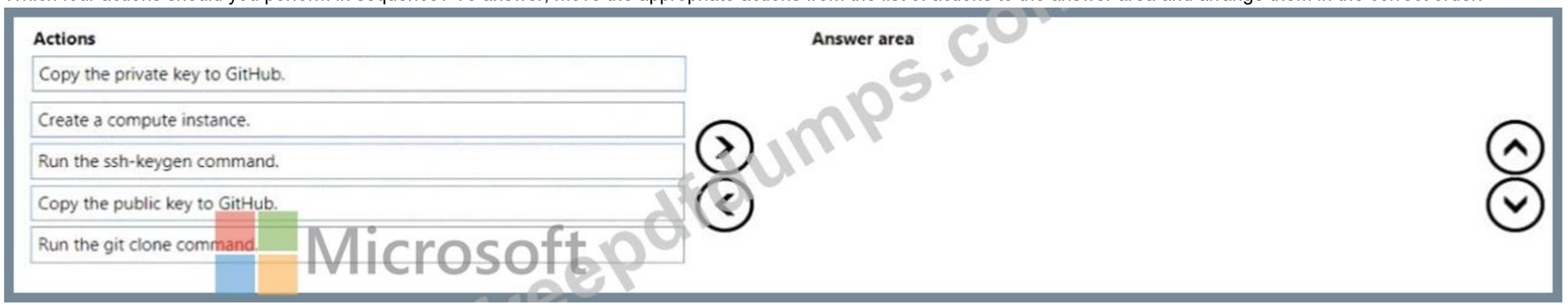
Explanation:

NEW QUESTION: 112

You create an Azure Machine Learning workspace.

You need to use the shared file system of the workspace to store a clone of a private Git repository.

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.



Answer:

Explanation:

NEW QUESTION: 113

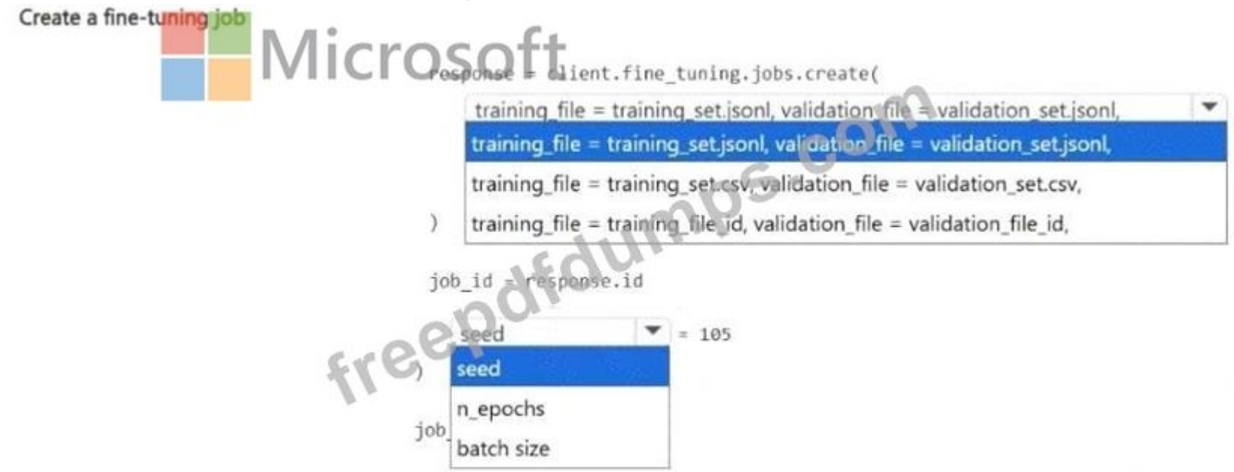
You manage an Azure AI Foundry project.

You plan to fine-tune a base model by using pre-uploaded training and validation data. You must specify a hyperparameter to ensure the job is reproducible.

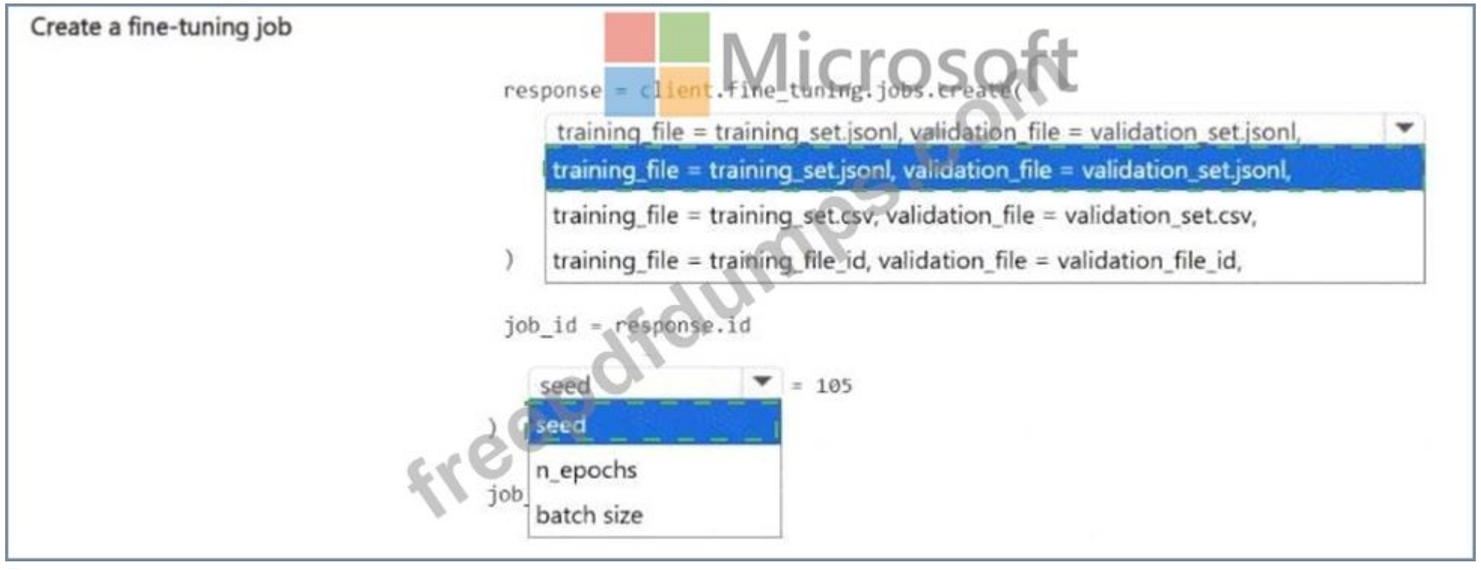
You need to submit the fine-tuning training job.

How should you complete the Python code segment? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.



Answer:



Explanation:

NEW QUESTION: 114
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 (LEAVE A REPLY)
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: 115
You plan to use automated machine learning to train a regression model. You have data that has features which have missing values, and categorical features with few distinct values.
You need to configure automated machine learning to automatically impute missing values and encode categorical features as part of the training task.
Which parameter and value pair should you use in the AutoMLConfig class?

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

NEW QUESTION: 116
You have an Azure Machine Learning workspace. You build a deep learning model.
You need to publish a GPU-enabled model as a web service.
Which two compute targets can you use? Each correct answer presents a complete solution.
NOTE: Each correct selection is worth one point.
A. Azure Container Instances (ACI)
B. Local web service
C. Azure Kubernetes Service (AKS)
D. Azure Machine Learning compute clusters
Answer: (SHOW ANSWER)

NEW QUESTION: 117
You have an Azure Machine Learning workspace named Workspace 1 Workspace! has a registered Mlflow model named model 1 with PyFunc flavor You plan to deploy model1 to an online endpoint named endpoint1 without egress connectivity by using Azure Machine learning Python SDK vl You have the following code:

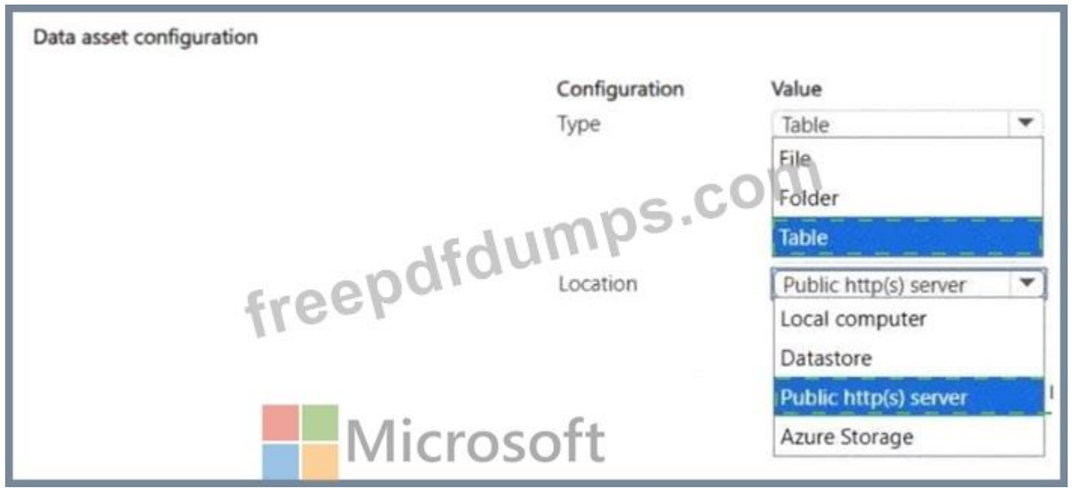
You need to add a parameter to the ManagedOnllneDeployment object to ensure the model deploys successfully Solution: Add the scoring_script parameter.
Does the solution meet the goal?
A. Yes
B. No
Answer: A (LEAVE A REPLY)

NEW QUESTION: 118
You manage an Azure Machine Learning workspace.
You must create and configure a compute cluster for a training job by using Python SDK v2.
You need to create a persistent Azure Machine Learning compute resource, specifying the fewest possible properties.
Which two properties should you define? Each correct answer presents part of the solution.
NOTE: Each correct selection is worth one point.
A. name
B. Min_instances
C. size
D. max_instances
E. type
Answer: B,D (LEAVE A REPLY)

NEW QUESTION: 119
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 an Azure Machine Learning model.
You plan to deploy the model to an online endpoint.
You need to ensure that applications will be able to use the authentication method with a non-expiring artifact to access the model.
Solution:
Create a managed online endpoint and set the value of its auto_mode parameter to key. Deploy the model to the inline endpoint.
Does the solution meet the goal?
A. No
B. Yes
Answer: (SHOW ANSWER)

NEW QUESTION: 120
You manage an Azure Machine Learning workspace and a GitHub repository. The GitHub repository contains a CSV file located at httpsy/raw.githubusercontent.com/account1/repo1/main/doc1/data1.csv. The CSV file includes embedded newlines.
You plan to consume the content of the CSV file in the workspace. The solution must minimize the possibility of misaligned field values when reading the file content.
You need to create a data asset that references the CSV file.
Which data asset configuration values should you use? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.

Answer:



Explanation:

NEW QUESTION: 121

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 tin- 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 Principal Components Analysis (PCA) sampling mode.
Does the solution meet the goal?
A. Yes
B. No
Answer: (SHOW ANSWER)
Instead use the Synthetic Minority Oversampling Technique (SMOTE) sampling mode.
Note: SMOTE is used to increase the number of underepresented cases in a dataset used for machine learning.
SMOTE is a better way of increasing the number of rare cases than simply duplicating existing cases.

NEW QUESTION: 122
You plan to use Hyperdrive to optimize the hyperparameters selected when training a model. You create the following code to define options for the hyperparameter experiment

```
import azureml.train.hyperdrive.parameter_expressions as pe
from azureml.train.hyperdrive import GridParameterSampling, HyperDriveConfig

param_sampling = GridParameterSampling({
   "max_depth" : pe.choice(6, 7, 8, 9),
   "learning_rate" : pe.choice(0.05, 0.1, 0.15)
   })
hyperdrive_run_config = HyperDriveConfig(
   estimator = estimator,
   hyperparameter_sampling = param_sampling,
   policy = None,
   primary_metric_name = "auc",
   primary_metric_goal = PrimaryMetricGoal.MAXIMIZE,
```

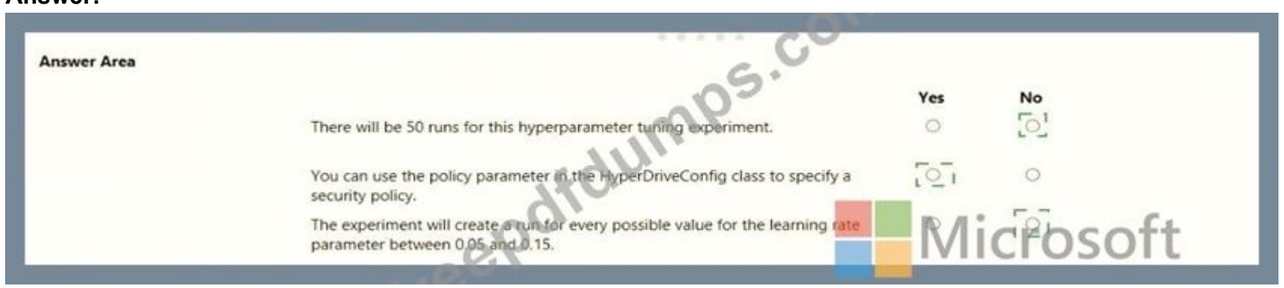
```
  estimator = estimator,
  hyperparameter_sampling = param_sampling,
  policy = None,
  primary_metric_name = "auc",
  primary_metric_goal = PrimaryMetricGoal.MAXIMIZE,
  max_total_runs = 50,
  max_concurrent_runs = 4)
```

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

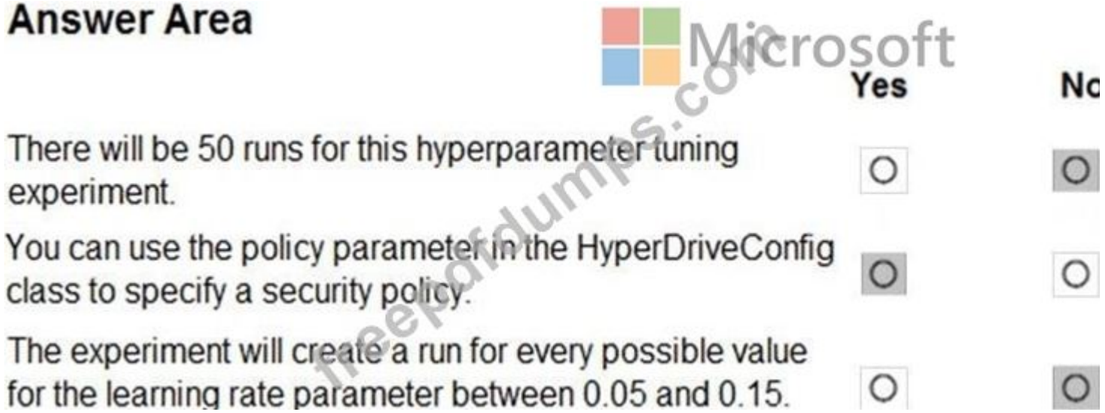
Answer Area

	Yes	No
There will be 50 runs for this hyperparameter tuning experiment.	○	○
You can use the policy parameter in the HyperDriveConfig class to specify a security policy.	○	○
The experiment will create a run for every possible value for the learning rate parameter between 0.05 and 0.15.	○	○

Answer:



Explanation:



Box 1: No

max_total_runs (50 here)

The maximum total number of runs to create. This is the upper bound; there may be fewer runs when the sample space is smaller than this value.

Box 2: Yes

Policy EarlyTerminationPolicy

The early termination policy to use. If None - the default, no early termination policy will be used.

Box 3: No

Discrete hyperparameters are specified as a choice among discrete values. choice can be:

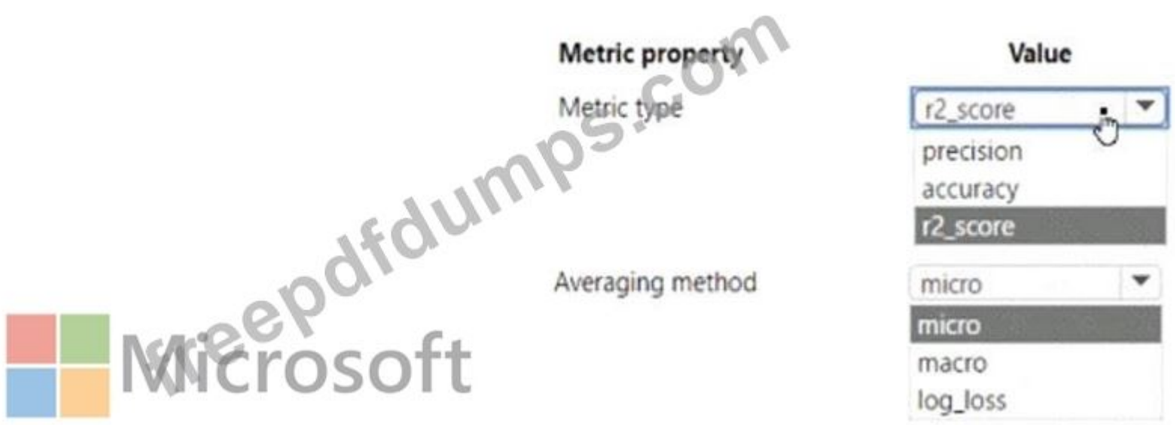
one or more comma-separated values

a range object

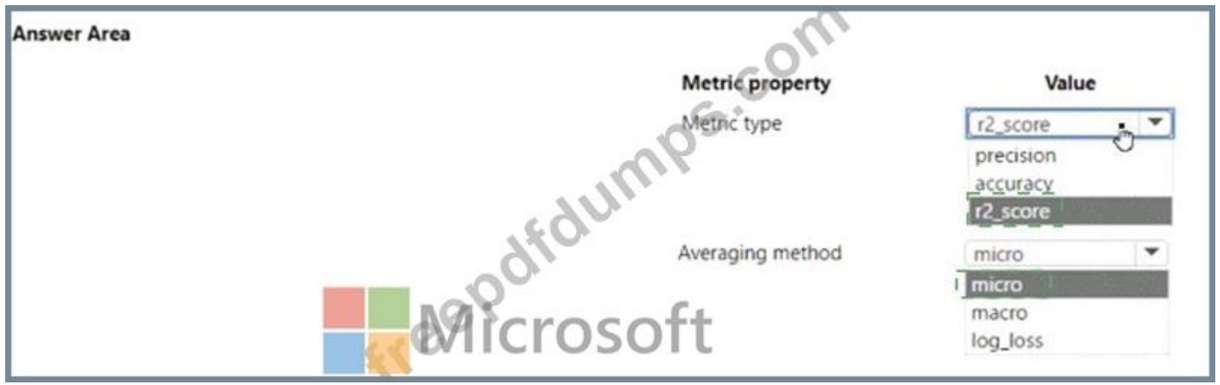
any arbitrary list object
Reference:
https://docs.microsoft.com/en-us/python/api/azureml-train-core/azureml.train.hyperdrive.hyperdriveconfig
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-tune-hyperparameters

NEW QUESTION: 123
You ate reviewing model benchmarks in Azure Al Foundry.
You must use an embedding model that can assess rank-order relevance based on cosine similarity. You need to select the applicable embedding model. Which model metric should you focus on?
A. Spearman correlation
B. V measure
C. F1 score
D. Mean average precision
Answer: (SHOW ANSWER)

NEW QUESTION: 124
You create an Azure Machine Learning workspace. You train a classification model by using automated machine learning (automated ML) in Azure Machine Learning studio. The training data contains multiple classes that have significantly different numbers of samples.
You must use a metric type to avoid labeling negative samples as positive and an averaging method that will minimize the class imbalance.
You need to configure the metric type and the averaging method.
Which configurations should you use? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.
nswer Area



Answer:



Explanation:

NEW QUESTION: 125

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.

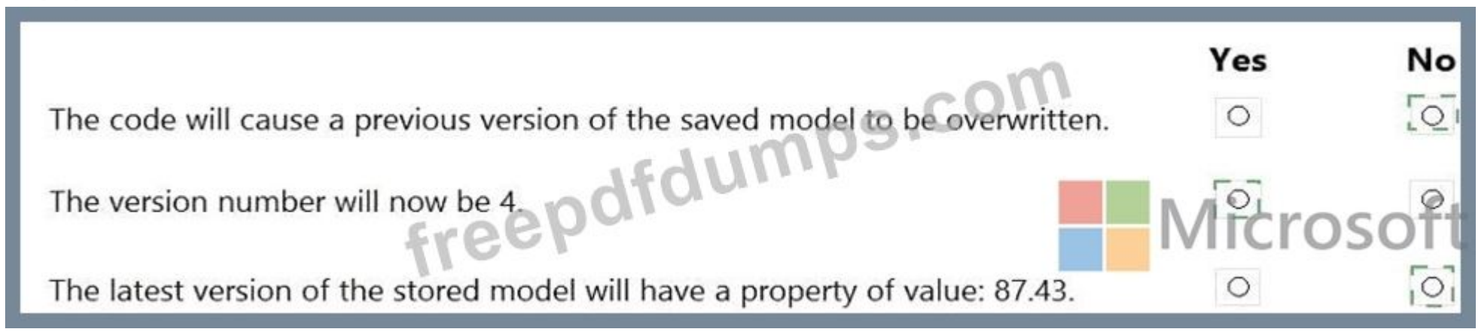
After the run has been submitted and completed, you run the following code:

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

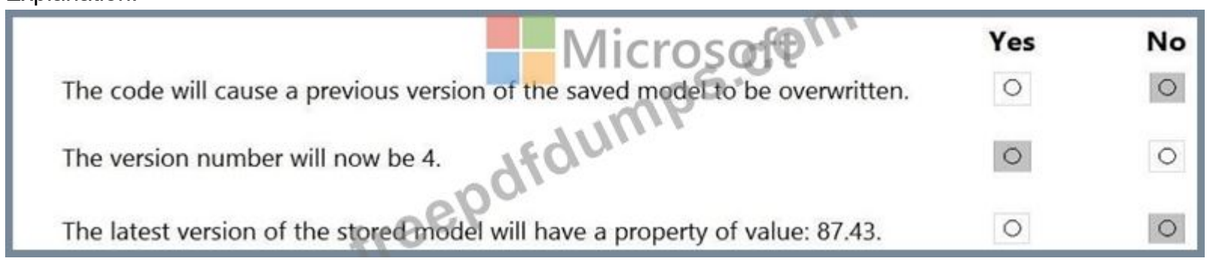
For each of the following statements, select Yes if the statement is true. Otherwise, select No.

NOTE: Each correct selection is worth one point.

Answer:



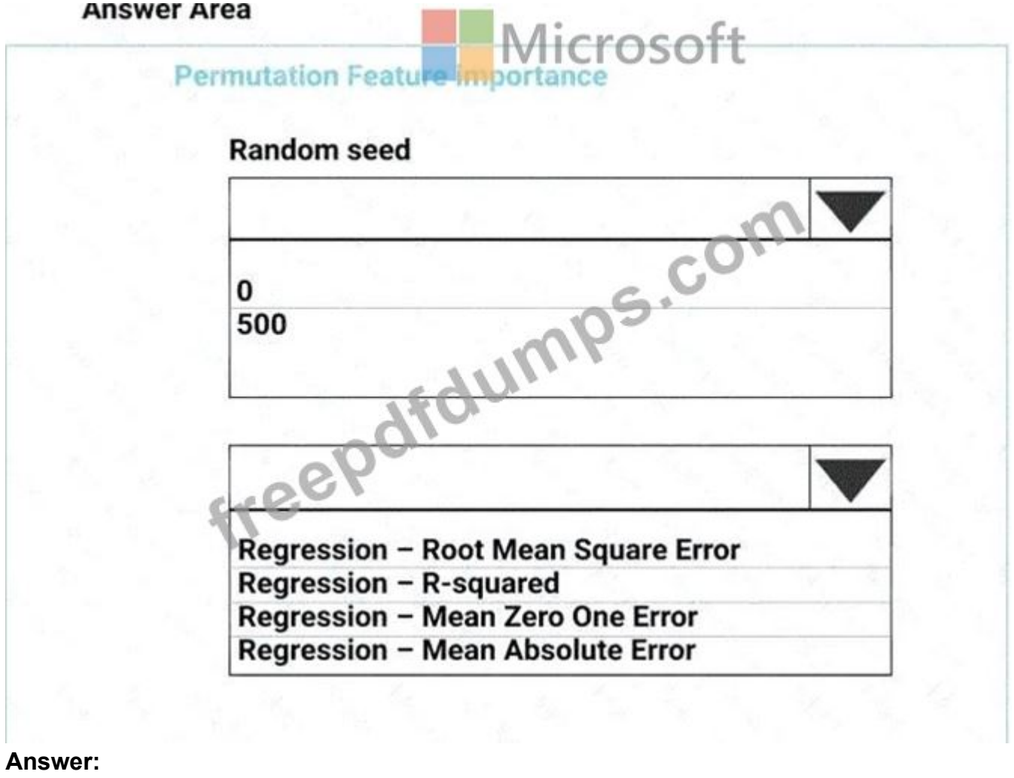
Explanation:



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

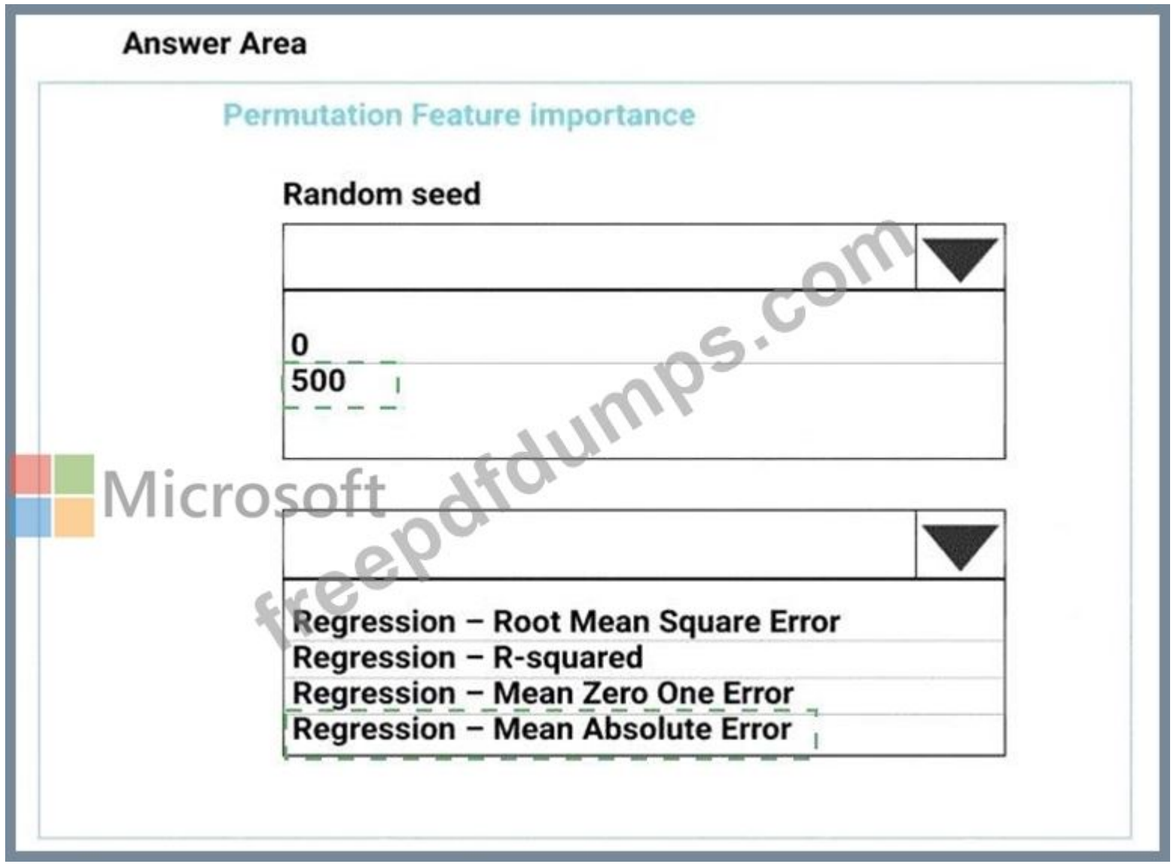
NEW QUESTION: 126
You need to configure the Permutation Feature Importance module for the model training requirements.
What should you do? To answer, select the appropriate options in the dialog box in the answer area.
NOTE: Each correct selection is worth one point.

Answer Area

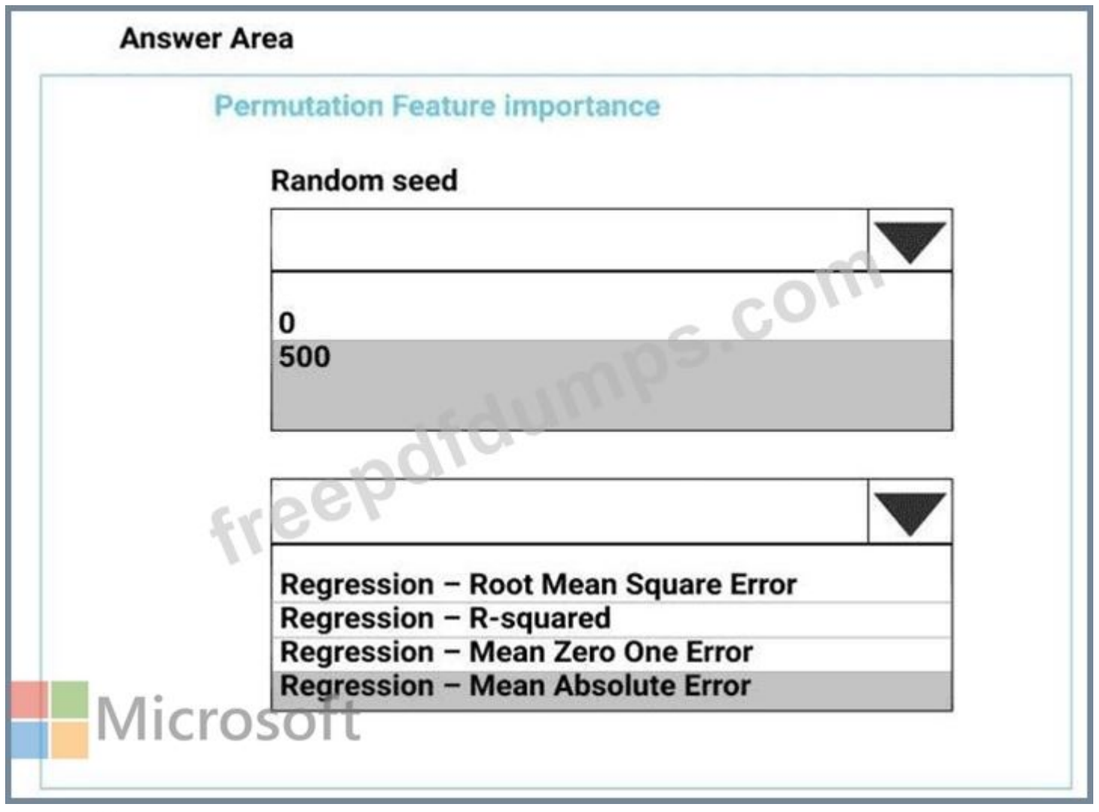


Answer:

Answer Area



Explanation:



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

NEW QUESTION: 127
You are using the Azure Machine Learning Service to automate hyperparameter exploration of your neural network classification model.
You must define the hyperparameter space to automatically tune hyperparameters using random sampling according to following requirements:
The learning rate must be selected from a normal distribution with a mean value of 10 and a standard deviation of 3.
Batch size must be 16, 32 and 64.
Keep probability must be a value selected from a uniform distribution between the range of 0.05 and 0.1.
You need to use the param_sampling method of the Python API for the Azure Machine Learning Service.
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 azureml.train.hyperdrive import RandomParameterSamplin
param_sampling = RandomParameterSampling( {
"learning_rate" :                                          ,
"batch_size":                                              ,
"keep_probability" :                                       ,
    }
```

- uniform(10,3)
- normal(10,3)
- choice(10,3)
- Loguniform(10,3)

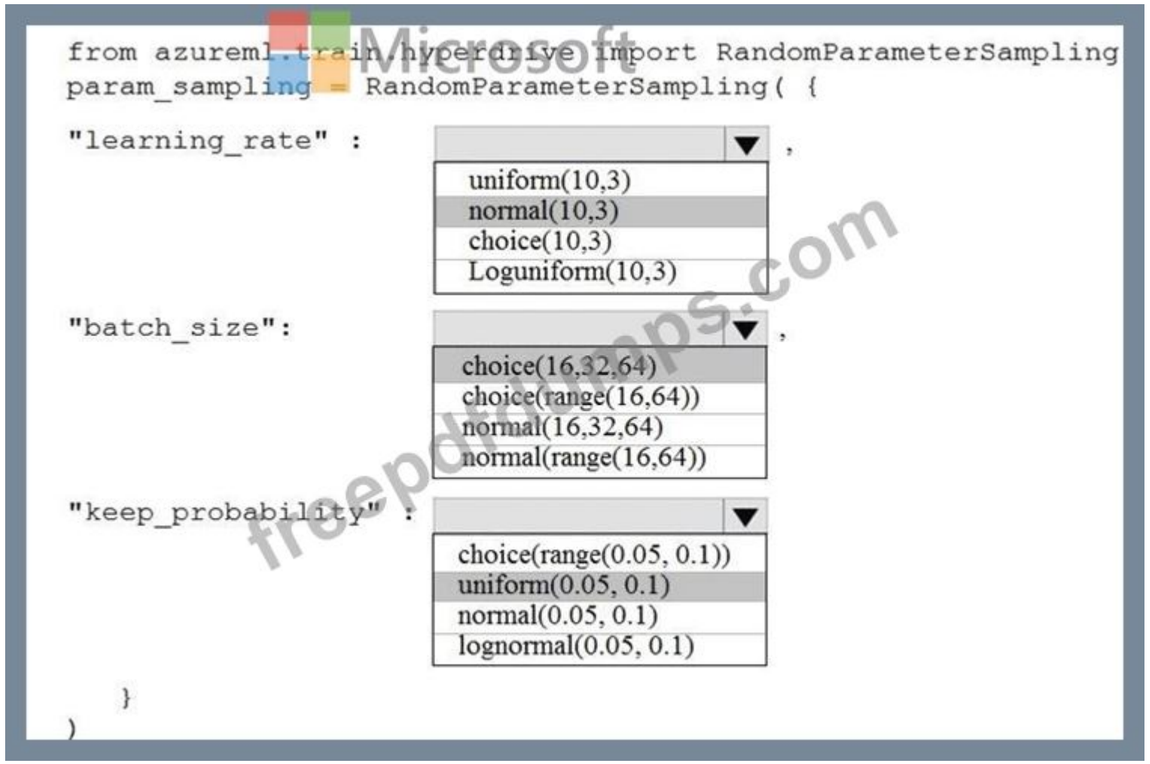
- choice(16,32,64)
- choice(range(16,64))
- normal(16,32,64)
- normal(range(16,64))

- choice(range(0.05, 0.1))
- uniform(0.05, 0.1)
- normal(0.05, 0.1)
- lognormal(0.05, 0.1)

Answer:

```
from azureml.train.hyperdrive import RandomParameterSampling
param_sampling = RandomParameterSampling( {
"learning_rate" :  [uniform(10,3) | normal(10,3) | choice(10,3) | Loguniform(10,3)] ,
"batch_size":  [choice(16,32,64) | choice(range(16,64)) | normal(16,32,64) | normal(range(16,64))] ,
"keep_probability" :  [choice(range(0.05, 0.1)) | uniform(0.05, 0.1) | normal(0.05, 0.1) | lognormal(0.05, 0.1)]
   }
```


Explanation:



In random sampling, hyperparameter values are randomly selected from the defined search space. Random sampling allows the search space to include both discrete and continuous hyperparameters.
Example:
from azureml.train.hyperdrive import RandomParameterSampling
param_sampling = RandomParameterSampling({
" learning_rate " : normal(10, 3),
" keep_probability " : uniform(0.05, 0.1),
" batch_size " : choice(16, 32, 64)
}
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/service/how-to-tune-hyperparameters

NEW QUESTION: 128
You have an Azure Machine Learning workspace.
You plan to implement automated hyperparameter tuning for model training in the workspace.
You need to select the sweep jobs parameter sampling method that will randomize the selection of hyperparameters from the search space but allow for reproducing search results.
Which sampling method should you use?
A. Sobol
B. Grid
C. Random
D. Bayesian

Answer: B (LEAVE A REPLY)

NEW QUESTION: 129
You register a model that you plan to use in a batch inference pipeline.
The batch inference pipeline must use a ParallelRunStep step to process files in a file dataset. The script has the ParallelRunStep step runs must process six input files each time the inferencing function is called.
You need to configure the pipeline.
Which configuration setting should you specify in the ParallelRunConfig object for the PrallelRunStep step?
A. process_count_per_node= "6"
B. node_count= "6"
C. mini_batch_size= "6"
D. error_threshold= "6"
Answer: B (LEAVE A REPLY)
node_count is the number of nodes in the compute target used for running the ParallelRunStep.
Reference:
https://docs.microsoft.com/en-us/python/api/azureml-contrib-pipeline-steps/azureml.contrib.pipeline.steps.parallelrunconfig?view=azure-ml-py

NEW QUESTION: 130
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: (SHOW ANSWER)
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: 131
You define a datastore named ml-data for an Azure Storage blob container. In the container, you have a folder named train that contains a file named data.csv. You plan to use the file to train a model by using the Azure Machine Learning SDK.
You plan to train the model by using the Azure Machine Learning SDK to run an experiment on local compute.
You define a DataReference object by running the following code:

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

You need to load the training data.

Which code segment should you use?

```
import os
import argparse
import pandas as pd

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

```
import os
import argparse
import pandas as pd

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

```
import pandas as pd
import os
import argparse
import pandas as pd

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

```
import os
import argparse
import pandas as pd

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

A. Option A

B. Option B

C. Option C

D. Option D

E. Option E

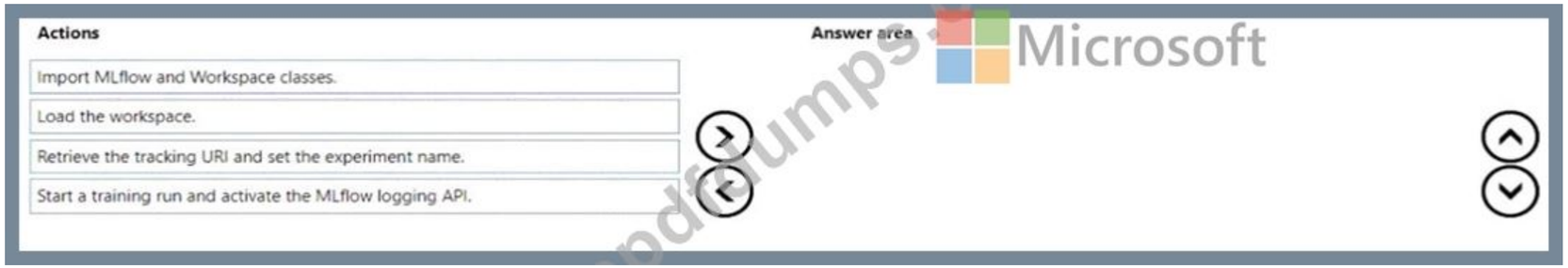
Answer: (SHOW ANSWER)

Example:
data_folder = args.data_folder
# Load Train and Test data
train_data = pd.read_csv(os.path.join(data_folder, ' data.csv'))
Reference:
https://www.element61.be/en/resource/azure-machine-learning-services-complete-toolbox-ai

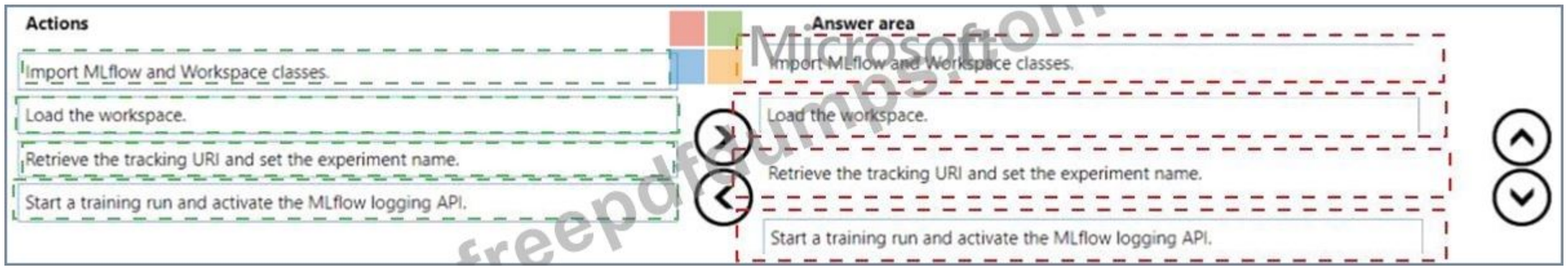
NEW QUESTION: 132
You arc creating a new experiment in Azure Machine Learning Studio. You have a small dataset that has missing values in many columns. The data does not require the application of predictors for each column. You plan to use the Clean Missing Data module to handle the missing data.
You need to select a data cleaning method.
Which method should you use?
A. Replace using MICE
B. Synthetic Minority
C. Replace using Probabilistic PAC
D. Normalization
Answer: C (LEAVE A REPLY)

NEW QUESTION: 133
You are performing feature engineering on a dataset.
You must add a feature named CityName and populate the column value with the text London.
You need to add the new feature to the dataset.
Which Azure Machine Learning Studio module should you use?
A. Edit Metadata
B. Preprocess Text
C. Execute Python Script
D. Latent Dirichlet Allocation
Answer: A (LEAVE A REPLY)
Typical metadata changes might include marking columns as features.
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/edit-metadata

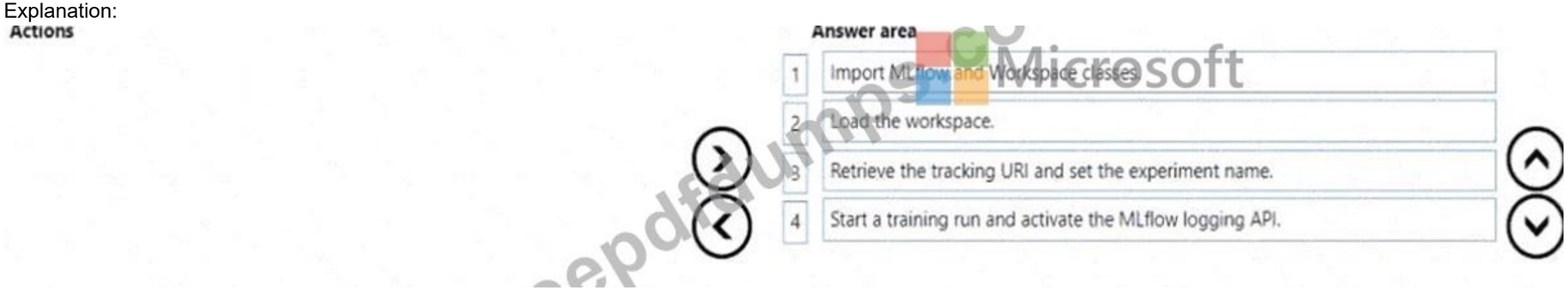
NEW QUESTION: 134
You have an Azure Machine Learning workspace. You are running an experiment on your local computer.
You need to use MLflow Tracking to store metrics and artifacts from your local experiment runs in the workspace.
In which order should you perform the actions? To answer, move all actions from the list of actions to the answer area and arrange them in the correct order.



Answer:



Explanation:



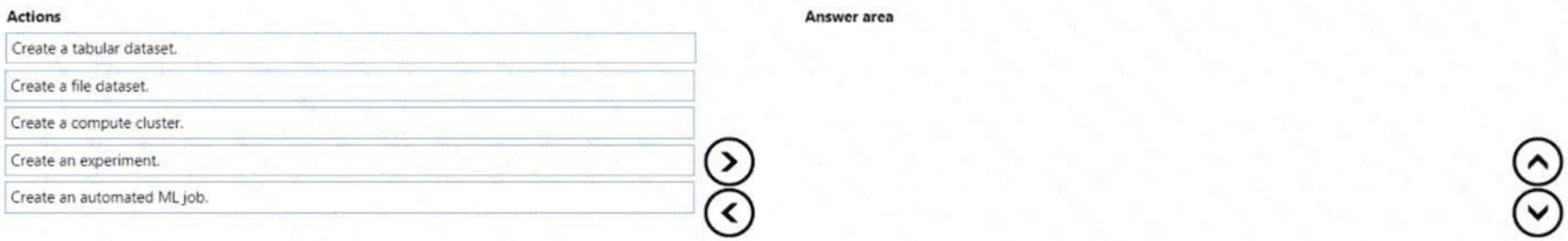
NEW QUESTION: 135

You create an Azure Machine Learning workspace. You are training a classification model with no-code AutoML in Azure Machine Learning studio.

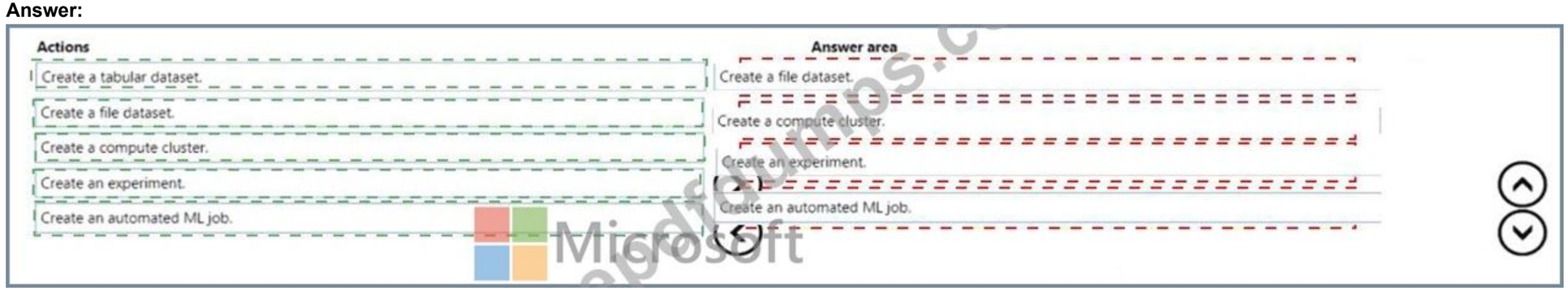
The model must predict if a client of a financial institution will subscribe to a fixed-term deposit. You must preview the data profile in Azure Machine Learning studio once the dataset is created.

You need to train the model.

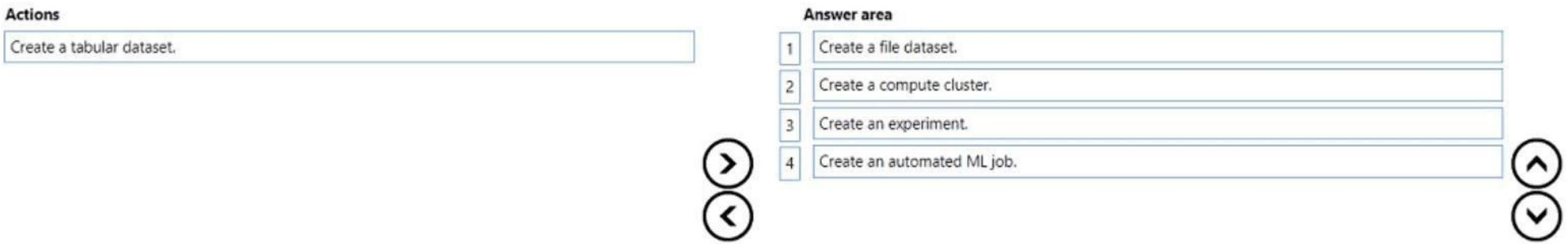
Which four actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.



Answer:



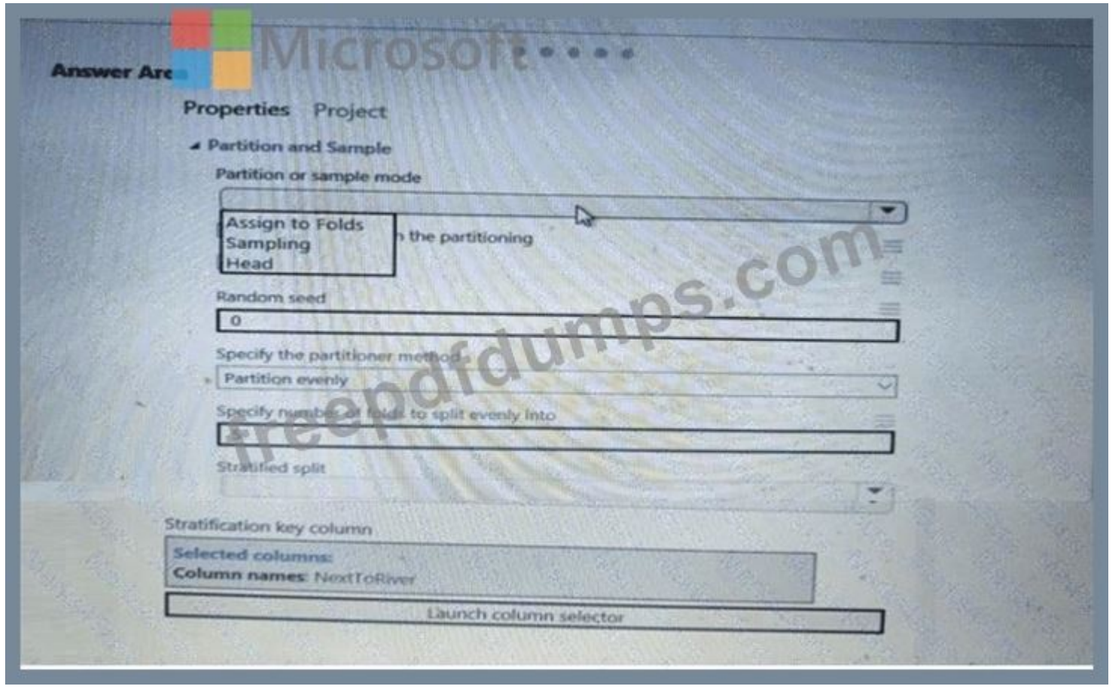
Explanation:



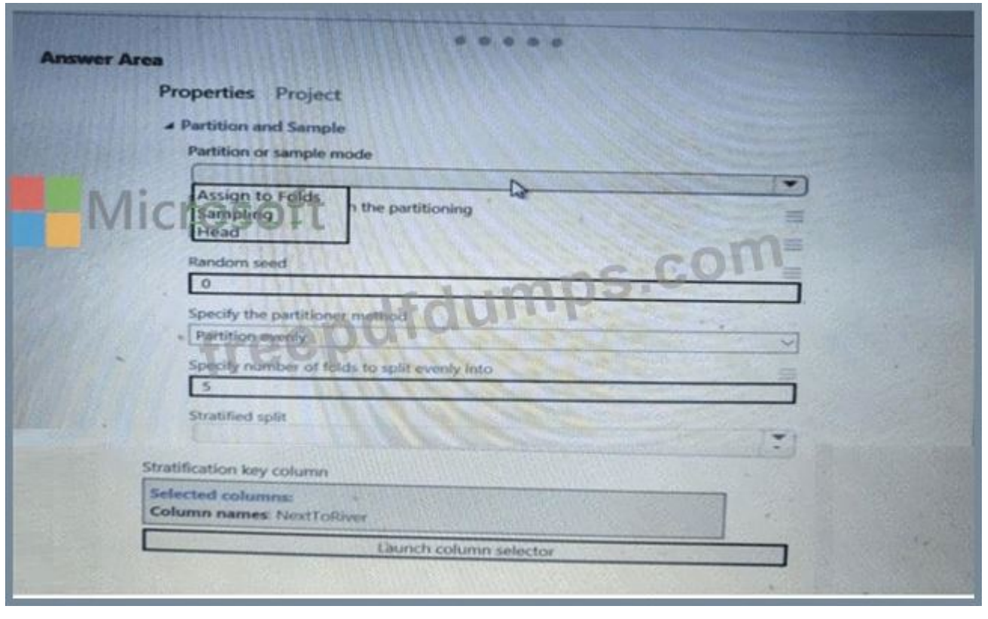
NEW QUESTION: 136

You need to identify the methods for dividing the data according, to the testing requirements.

Which properties should you select? To answer, select the appropriate option-, m the answer area. NOTE: Each correct selection is worth one point.


Answer Are
Properties Project
Partition and Sample
Partition or sample mode
Assign to Folds
Sampling
Head
the partitioning
Random seed
0
Specify the partitioner method
Partition evenly
Specify number of folds to split evenly into
Stratified split
Stratification key column
Selected columns:
Column names: NextToRiver
Launch column selector
freepdfdumps.com

Answer:


Answer Area
Properties Project
Partition and Sample
Partition or sample mode
Assign to Folds
Sampling
Head
the partitioning
Random seed
0
Specify the partitioner method
Partition evenly
Specify number of folds to split evenly into
5
Stratified split
Stratification key column
Selected columns:
Column names: NextToRiver
Launch column selector
Microsoft
freepdfdumps.com

NEW QUESTION: 137
You ate reviewing model benchmarks in Azure Al Foundry.
You must use a large language model based on the proficiency of the model to generate the most linguistically correct text. You need to select the model benchmark. Which benchmark metric should you focus on?
A. coherence
B. precision
C. fluency
D. accuracy
Answer: (SHOW ANSWER)

NEW QUESTION: 138
You download a .csv file from a notebook in an Azure Machine Learning workspace to a data/sample.csv folder on a compute instance. The file contains 10,000 records. You must generate the summary statistics for the data in the file. The statistics must include the following for each numerical column:
* number of non-empty values
* average value
* standard deviation
* minimum and maximum values
* 25th. 50th. and 75th percentiles
You need to complete the Python code that will generate the summary statistics.
Which code segments should you use? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.

Answer:

Explanation:

NEW QUESTION: 139

You plan to explore demographic data for home ownership in various cities. The data is in a CSV file with the following format:

age,city,income,home_owner

21,Chicago,50000,0

35,Seattle,120000,1

23,Seattle,65000,0

45,Seattle,130000,1

18,Chicago,48000,0

You need to run an experiment in your Azure Machine Learning workspace to explore the data and log the results. The experiment must log the following information:

the number of observations in the dataset

a box plot of income by home_owner

a dictionary containing the city names and the average income for each city You need to use the appropriate logging methods of the experiment's run object to log the required information.

How should you complete the code? To answer, drag the appropriate code segments to the correct locations.

Each code segment may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.

Code segments

- log
- log_list
- log_row
- log_table
- log_image

Answer Area

```
from azureml.core import Experiment, Run
import pandas as pd
import matplotlib.pyplot as plt
# Create an Azure ML experiment in workspace
experiment = Experiment(workspace = ws, name = "demo-experiment")
# Start logging data from the experiment
run = experiment.start_logging()
# load the dataset
data = pd.read_csv('research/demographics.csv')
# Log the number of observations
row_count = (len(data))
run. Segment ("observations", row_count)
# Log box plot for income by home_owner
fig = plt.figure(figsize=(9, 6))
ax = fig.gca()
data.boxplot(column = 'income', by = "home_owner", ax = ax)
ax.set_title('income by home_owner')
ax.set_ylabel('income')
run. Segment (name = 'income_by_home_owner', plot = fig)
# Create a dataframe of mean income per city
mean_inc_df = data.groupby('city')['income'].agg(np.mean).to_frame().reset_index()
# Convert to a dictionary
mean_inc_dict = mean_inc_df.to_dict('dict')
# Log city names and average income dictionary
run. Segment (name="mean_income_by_city", value= mean_inc_dict)
# Complete tracking and get link to details
run.complete()
```


Answer:

Code segments

- log
- log_list
- log_row
- log_table
- log_image

Answer Area

```
from azureml.core import Experiment, Run
import pandas as pd
import matplotlib.pyplot as plt
# Create an Azure ML experiment in workspace
experiment = Experiment(workspace = ws, name = "demo-experiment")
# Start logging data from the experiment
run = experiment.start_logging()
# load the dataset
data = pd.read_csv('research/demographics.csv')
# Log the number of observations
row_count = (len(data))
run.log ("observations", row_count)
# Log box plot for income by home_owner
fig = plt.figure(figsize=(9, 6))
ax = fig.gca()
data.boxplot(column = 'income', by = "home_owner", ax = ax)
ax.set_title('income by home_owner')
ax.set_ylabel('income')
run.log_image (name = 'income_by_home_owner', plot = fig)
# Create a dataframe of mean income per city
mean_inc_df = data.groupby('city')['income'].agg(np.mean).to_frame().reset_index()
# Convert to a dictionary
mean_inc_dict = mean_inc_df.to_dict('dict')
# Log city names and average income dictionary
run.log_table (name="mean_income_by_city", value= mean_inc_dict)
# Complete tracking and get link to details
run.complete()
```


Explanation:

```
from azureml.core import Experiment, Run
import pandas as pd
import matplotlib.pyplot as plt
# Create an Azure ML experiment in workspace
experiment = Experiment(workspace = ws, name = "demo-experiment")
# Start logging data from the experiment
run = experiment.start_logging()
# load the dataset
data = pd.read_csv('research/demographics.csv')
# Log the number of observations
row_count = (len(data))
run. log          ("observations", row_count)
# Log box plot for income by home_owner
fig = plt.figure(figsize=(9, 6))
ax = fig.gca()
data.boxplot(column = 'income', by = "home_owner", ax = ax)
ax.set_title('income by home_owner')
ax.set_ylabel('income')
run. log_image    (name = 'income_by_home_owner', plot = fig)
# Create a dataframe of mean income per city
mean_inc_df = data.groupby('city')['income'].agg(np.mean).to_frame().reset_index()
# Convert to a dictionary
mean_inc_dict = mean_inc_df.to_dict('dict')
# Log city names and average income dictionary
run. log_table    (name="mean_income_by_city", value= mean_inc_dict)
# Complete tracking and get link to details
```


Box 1: log

The number of observations in the dataset.

run.log(name, value, description= ' ')

Scalar values: Log a numerical or string value to the run with the given name. Logging a metric to a run causes that metric to be stored in the run record in the experiment. You can log the same metric multiple times within a run, the result being considered a vector of that metric.

Example: run.log(" accuracy " , 0.95)

Box 2: log_image

A box plot of income by home_owner.

log_image Log an image to the run record. Use log_image to log a .PNG image file or a matplotlib plot to the run. These images will be visible and comparable in the run record.

Example: run.log_image(" ROC " , plot=plt)

Box 3: log_table

A dictionary containing the city names and the average income for each city.

log_table: Log a dictionary object to the run with the given name.

NEW QUESTION: 140

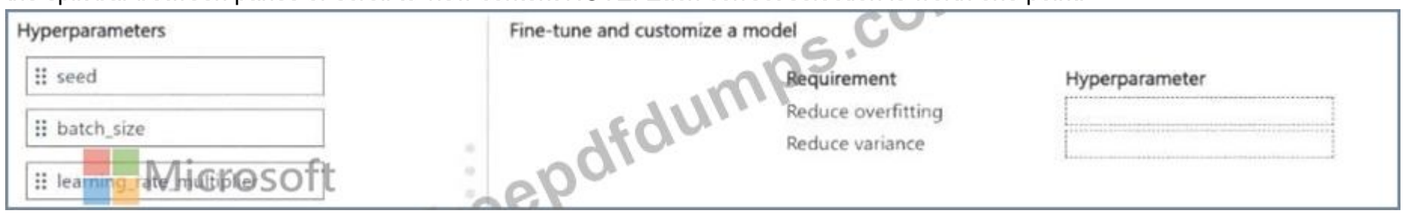
You are with a time series dataset in Azure Machine Learning Studio.

You need to split your dataset into training and testing subsets by using the Split Data module.

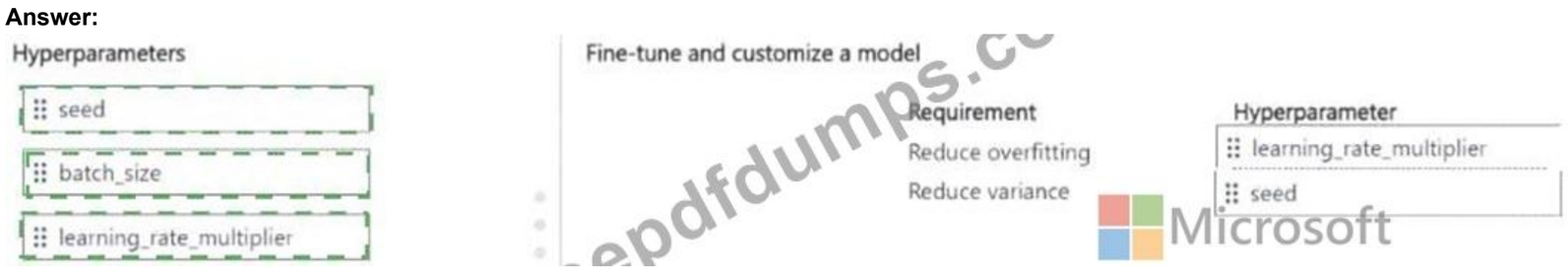
Which splitting mode should you use?

A. Regular Expression Split
B. Split Rows with the Randomized split parameter set to true
C. Relative Expression Split
D. Recommender Split
Answer: B (LEAVE A REPLY)
Split Rows: Use this option if you just want to divide the data into two parts. You can specify the percentage of data to put in each split, but by default, the data is divided 50-50.
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/split-data

NEW QUESTION: 141
You manage an Azure AI Foundry project. You fine-tune the base model
During evaluation, you observe that the model is overfitting and its responses are highly varying You need to improve the fine-tuned model.
Which hyperparameters should you use? To answer, move the appropriate hyper para meters to the correct requirements. You may use each hyperparameter once, more than once, or not at all. You may need to move the split bar between panes or scroll to view content NOTE: Each correct selection is worth one point.



Answer:



Explanation:

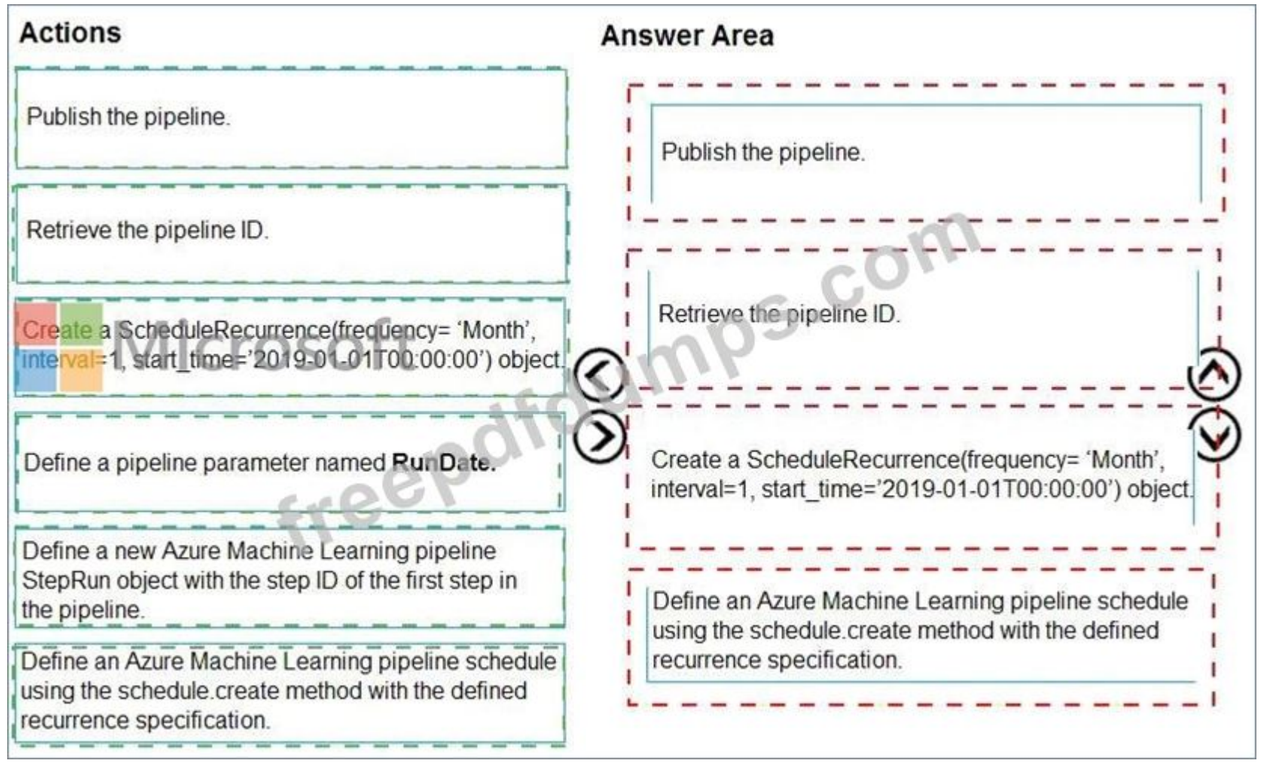
NEW QUESTION: 142
You create a multi-class image classification deep learning model.
The model must be retrained monthly with the new image data fetched from a public web portal. You create an Azure Machine Learning pipeline to fetch new data, standardize the size of images, and retrain the model.
You need to use the Azure Machine Learning SDK to configure the schedule for the pipeline.
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

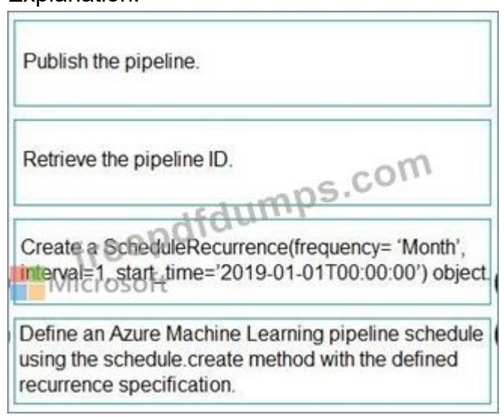
- Publish the pipeline.
- Retrieve the pipeline ID.
- Create a ScheduleRecurrence(frequency= 'Month', interval=1, start_time='2019-01-01T00:00:00') object.
- Define a pipeline parameter named **RunDate**.
- Define a new Azure Machine Learning pipeline StepRun object with the step ID of the first step in the pipeline.
- Define an Azure Machine Learning pipeline schedule using the schedule.create method with the defined recurrence specification.

Answer Area

Answer:



Explanation:



Step 1: Publish the pipeline.

To schedule a pipeline, you ' ll need a reference to your workspace, the identifier of your published pipeline, and the name of the experiment in which you wish to create the schedule.

Step 2: Retrieve the pipeline ID.

Needed for the schedule.

Step 3: Create a ScheduleRecurrence..

To run a pipeline on a recurring basis, you ' ll create a schedule. A Schedule associates a pipeline, an experiment, and a trigger.

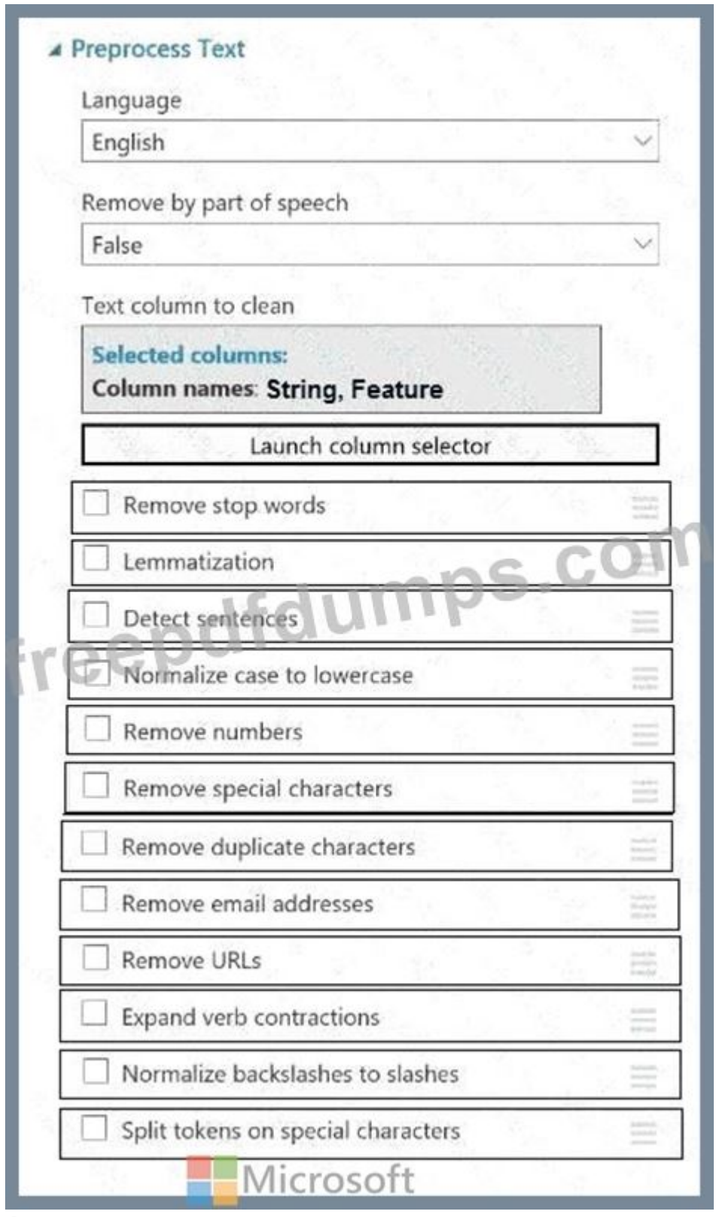
First create a schedule. Example: Create a Schedule that begins a run every 15 minutes:

recurrence = ScheduleRecurrence(frequency= " Minute " , interval=15)

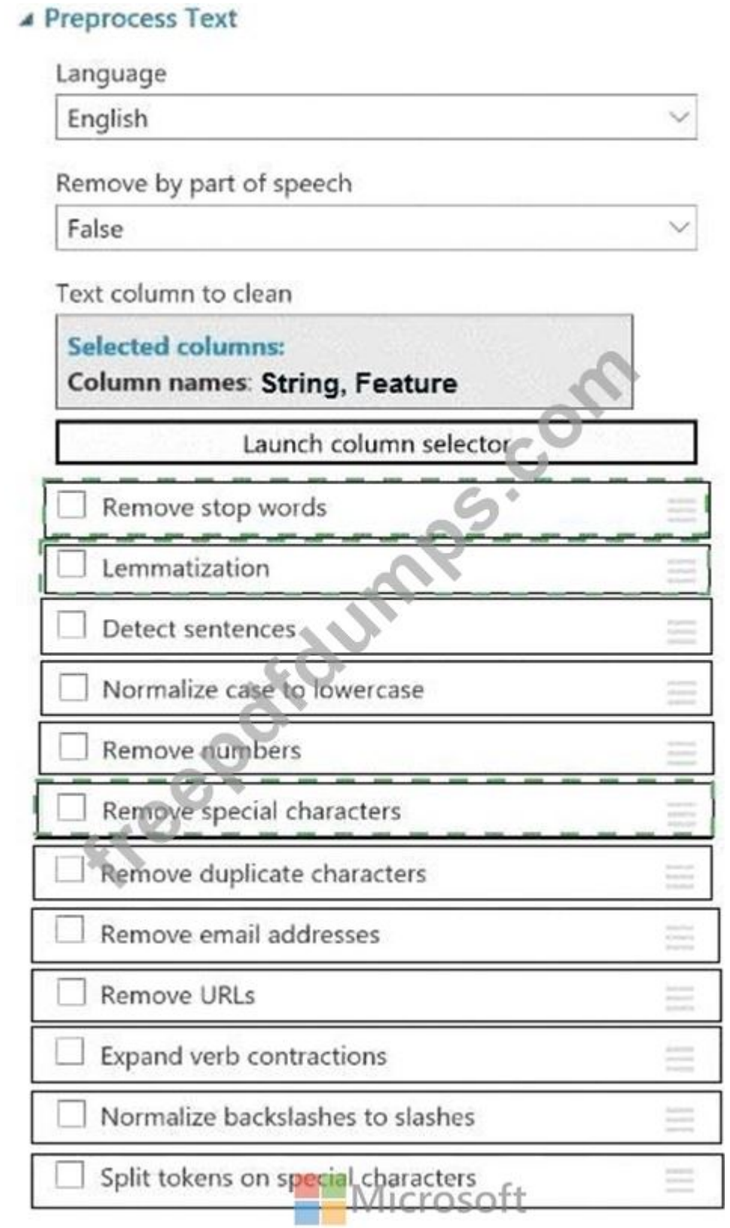
Step 4: Define an Azure Machine Learning pipeline schedule..
Example, continued:
recurring_schedule = Schedule.create(ws, name= " MyRecurringSchedule " , description= " Based on time " , pipeline_id=pipeline_id, experiment_name=experiment_name, recurrence=recurrence) Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-schedule-pipelines

NEW QUESTION: 143
You are moving a large dataset from Azure Machine Learning Studio to a Weka environment.
You need to format the data for the Weka environment.
Which module should you use?
A. Convert to CSV
B. Convert to Dataset
C. Convert to ARFF
D. Convert to SVMLight
Answer: C (LEAVE A REPLY)
Use the Convert to ARFF module in Azure Machine Learning Studio, to convert datasets and results in Azure Machine Learning to the attribute-relation file format used by the Weka toolset. This format is known as ARFF.
The ARFF data specification for Weka supports multiple machine learning tasks, including data preprocessing, classification, and feature selection. In this format, data is organized by entites and their attributes, and is contained in a single text file.
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/convert-to-arff

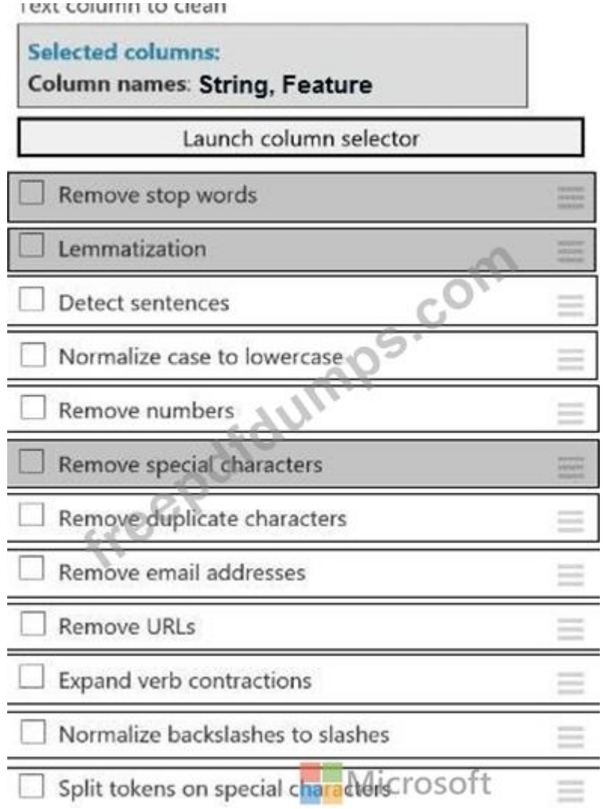
NEW QUESTION: 144
You plan to preprocess text from CSV files. You load the Azure Machine Learning Studio default stop words list.
You need to configure the Preprocess Text module to meet the following requirements:
Ensure that multiple related words from a single canonical form.
Remove pipe characters from text.
Remove words to optimize information retrieval.
Which three options should you select? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.


Preprocess Text
Language
English
Remove by part of speech
False
Text column to clean
Selected columns:
Column names: String, Feature
Launch column selector
Remove stop words
Lemmatization
Detect sentences
Normalize case to lowercase
Remove numbers
Remove special characters
Remove duplicate characters
Remove email addresses
Remove URLs
Expand verb contractions
Normalize backslashes to slashes
Split tokens on special characters
Microsoft

Answer:



Explanation:



Box 1: Remove stop words

Remove words to optimize information retrieval.

Remove stop words: Select this option if you want to apply a predefined stopword list to the text column.

Stop word removal is performed before any other processes.

Box 2: Lemmatization

Ensure that multiple related words from a single canonical form.

Lemmatization converts multiple related words to a single canonical form Box 3: Remove special characters Remove special characters: Use this option to replace any non-alphanumeric special characters with the pipe | character.

References:

https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/preprocess-text

NEW QUESTION: 145

You manage an Azure AI Foundry project.

You develop a Prompt flow that includes a large language model (LLM) node and an upstream node with a single output. You need to link the LLM node input with the output of the upstream node by using a YAML flow configuration. Which flow configuration should you use?

A. {{upstream.node.nameoutput})

B. $(upstream_node_name.output)

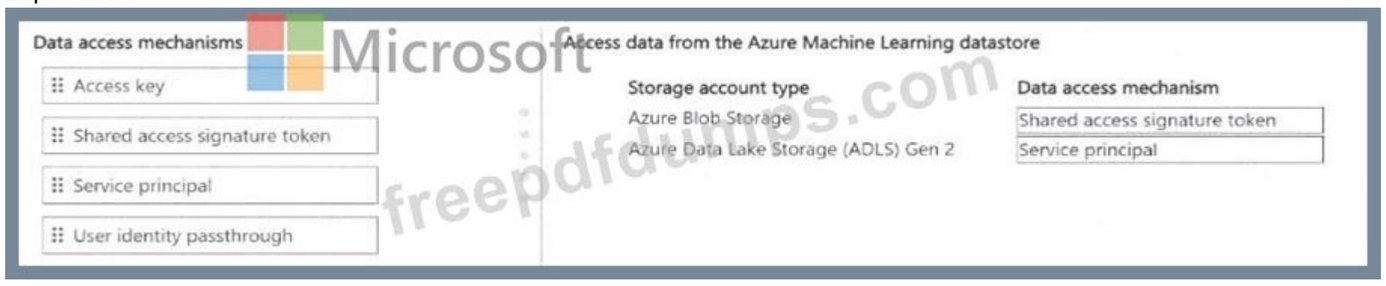
C. (% upstream node_name,output%}

D. <#upstream_node_nameoutput#>
Answer: (SHOW ANSWER)

NEW QUESTION: 146
You manage an Azure Machine Learning workspace That has an Azure Machine Learning datastore.
Data must be loaded from the following sources:
* a credential-less Azure Blob Storage
* an Azure Data Lake Storage (ADLS) Gen 2 which is not a credential-less datastore You need to define the authentication mechanisms to access data in the Azure Machine Learning datastore.
Which data access mechanism should you use? To answer, move the appropriate data access mechanisms to the correct storage types. You may use each data access mechanism once, more than once, or not at all.
You may need to move the split bar between panes or scroll to view content.
NOTE: Each correct selection is worth one point.

Answer:

Explanation:



NEW QUESTION: 147
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

Answer:

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)
```

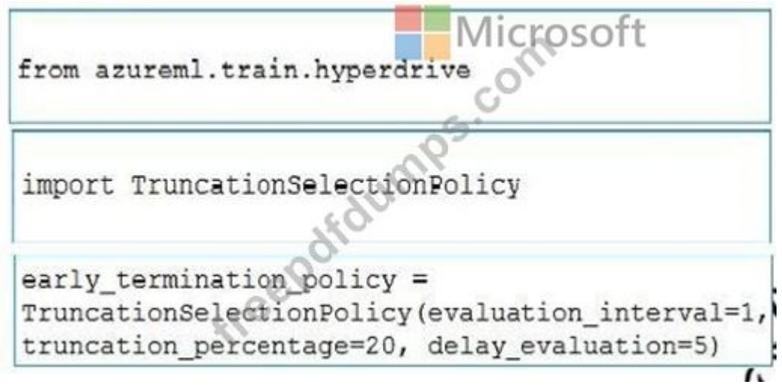
Answer Area

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

```
import TruncationSelectionPolicy
```

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


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: 148
You use the Azure Machine Learning designer to create and run a training pipeline.
The pipeline must be run every night to inference predictions from a large volume of files. The folder where the files will be stored is defined as a dataset.
You need to publish the pipeline as a REST service that can be used for the nightly inferencing run.
What should you do?
A. Create a batch inference pipeline
B. Set the compute target for the pipeline to an inference cluster
C. Create a real-time inference pipeline
D. Clone the pipeline
Answer: A (LEAVE A REPLY)
Azure Machine Learning Batch Inference targets large inference jobs that are not time-sensitive. Batch Inference provides cost-effective inference compute scaling, with unparalleled throughput for asynchronous applications. It is optimized for high-throughput, fire-and-forget inference over large collections of data.
You can submit a batch inference job by pipeline_run, or through REST calls with a published pipeline.
Reference:
https://github.com/Azure/MachineLearningNotebooks/blob/master/how-to-use-azureml/machine-learning-pipelines/parallel-run/README.md

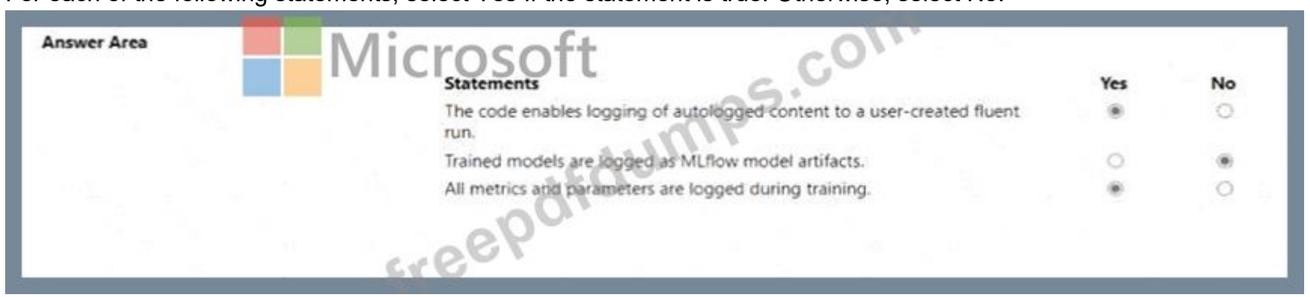
NEW QUESTION: 149
You are developing a data science workspace that uses an Azure Machine Learning service.
You need to select a compute target to deploy the workspace.
What should you use?
A. Azure Data Lake Analytics
B. Azure Databrick .
C. Apache Spark for HDInsight.
D. Azure Container Service

Answer: (SHOW ANSWER)
Azure Container Instances can be used as compute target for testing or development. Use for low-scale CPU-based workloads that require less than 48 GB of RAM.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/service/how-to-deploy-and-where

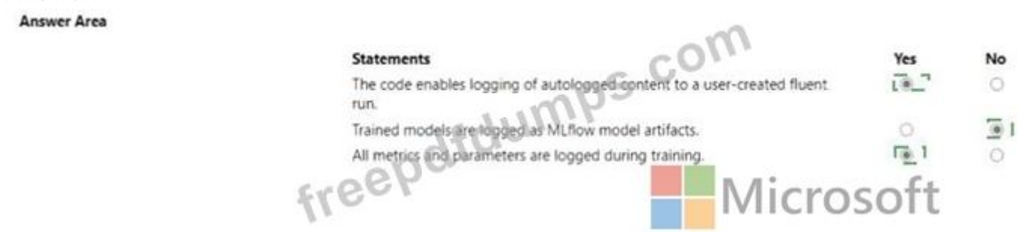
NEW QUESTION: 150
You manage an Azure Machine Learning workspace. You submit a training job with the Azure Machine Learning Python SDK v2. You must use MLflow to log metrics, model parameters, and model artifacts automatically when training a model.
You start by writing the following code segment:

```
import mlflow
mlflow.autolog(log_models=False, exclusive=True)
```

For each of the following statements, select Yes If the statement is true. Otherwise, select No.



Answer:



Explanation:

NEW QUESTION: 151
You run an automated machine learning experiment in an Azure Machine Learning workspace. Information about the run is listed in the table below:

Experiment	Run ID	Status	Created on	Duration
auto_ml_classification	AutoML_1234567890-123	Completed	11/11/2019 11:00:00 AM	00:27:11

You need to write a script that uses the Azure Machine Learning SDK to retrieve the best iteration of the experiment run. Which Python code segment should you use?

A.
```
from azureml.core import Workspace
from azureml.train.automl.run import AutoMLRun
ws = Workspace.from_config()
automl_ex = ws.experiments.get('auto_ml_classification')
automl_run = AutoMLRun(automl_ex, 'AutoML_1234567890-123')
best_iter = automl_run.get_output()[0]
```

B.
```
from azureml.core import Workspace
from azureml.train.automl.run import AutoMLRun
ws = Workspace.from_config()
automl_ex = ws.experiments.get('auto_ml_classification')
best_iter = list(automl_ex.get_runs())[0]
```

C.
```
from azureml.core import Workspace
from azureml.train.automl.run import AutoMLRun
ws = Workspace.from_config()
automl_ex = ws.experiments.get('auto_ml_classification')
best_iter = list(automl_ex.get_runs())[0]
```

D.
```
from azureml.core import Workspace
```

Answer: (SHOW ANSWER)

The get_output method on automl_classifier returns the best run and the fitted model for the last invocation.

Overloads on get_output allow you to retrieve the best run and fitted model for any logged metric or for a particular iteration.

In [ ]:

best_run, fitted_model = local_run.get_output()

Reference:

https://notebooks.azure.com/azureml/projects/azureml-getting-started/html/how-to-use-azureml/automated- machine-learning/classification-with-deployment/auto-ml-classification-with-deployment.ipynb

NEW QUESTION: 152

You have a dataset that contains 2,000 rows. You are building a machine learning classification model by using Azure Learning Studio. You add a Partition and Sample module to the experiment.

You need to configure the module. You must meet the following requirements:

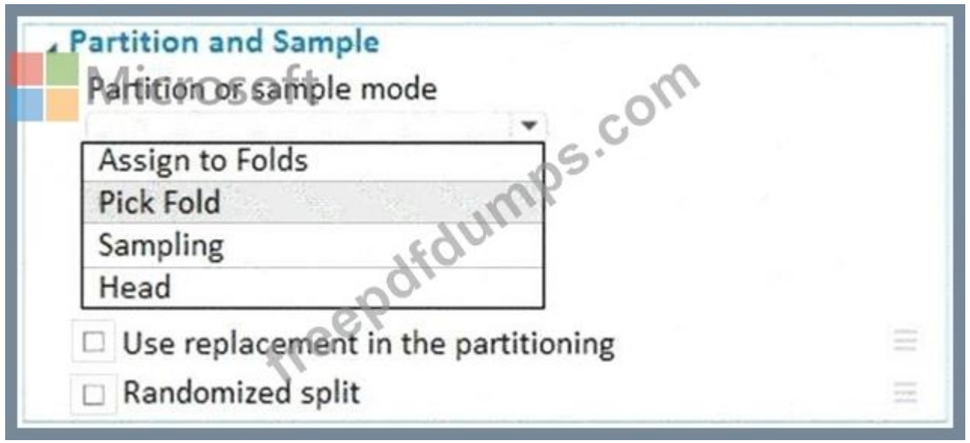
Divide the data into subsets

Assign the rows into folds using a round-robin method

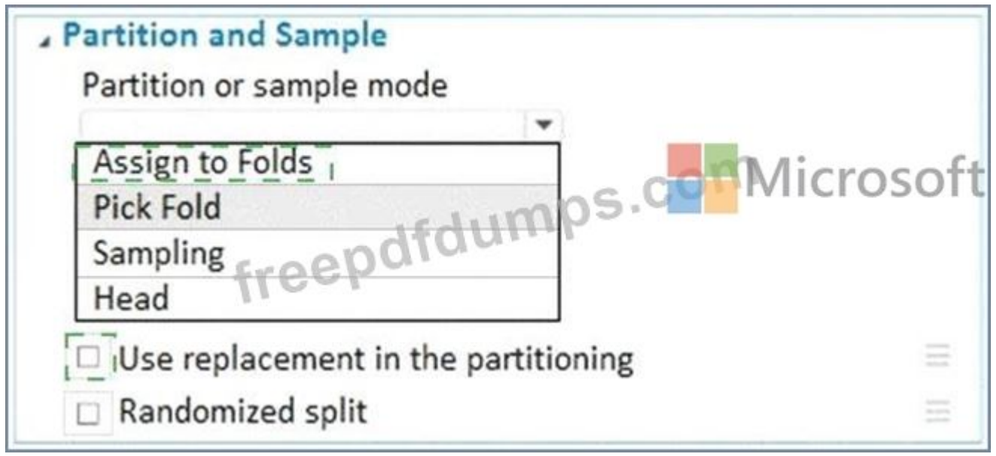
Allow rows in the dataset to be reused

How should you configure the module? To answer, select the appropriate options in the dialog box in the answer area.

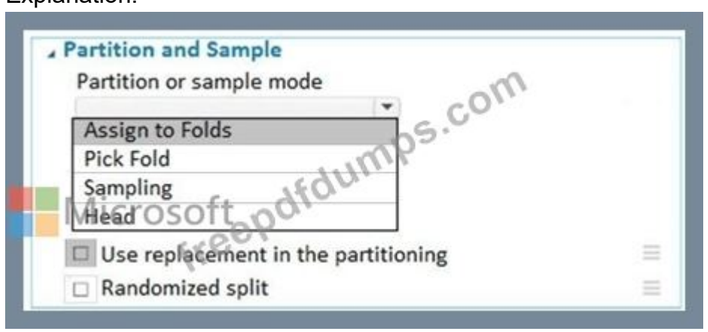
NOTE: Each correct selection is worth one point.



Answer:



Explanation:



Use the Split data into partitions option when you want to divide the dataset into subsets of the data. This option is also useful when you want to create a custom number of folds for cross-validation, or to split rows into several groups.

Add the Partition and Sample module to your experiment in Studio (classic), and connect the dataset.

For Partition or sample mode, select Assign to Folds.

Use replacement in the partitioning: Select this option if you want the sampled row to be put back into the pool of rows for potential reuse. As a result, the same row might be assigned to several folds.

If you do not use replacement (the default option), the sampled row is not put back into the pool of rows for potential reuse. As a result, each row can be assigned to only one fold.

Randomized split: Select this option if you want rows to be randomly assigned to folds.
If you do not select this option, rows are assigned to folds using the round-robin method.
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/partition-and-sample

NEW QUESTION: 153
You manage an Azure Machine Learning workspace.
You need to define an environment from a Docker image by using the Azure Machine Learning Python SDK v2.
Which parameter should you use?
A. image
B. conda_file
C. properties
D. build
Answer: (SHOW ANSWER)

NEW QUESTION: 154
You manage an Azure Machine Learning workspace.
You train a model interactively with a Jupyter Notebook in the workspace During training, a dataset is created with accuiacy and loss metrics for each epoch.
You need to configure model tracking with MLflow to log the dataset created during the training.
How should you complete the code segment? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.

Answer:

Explanation:

NEW QUESTION: 155

An organization uses Azure Machine Learning service and wants to expand their use of machine learning.

You have the following compute environments. The organization does not want to create another compute environment.

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

You need to determine which compute environment to use for the following scenarios.

Which compute types should you use? To answer, drag the appropriate compute environments to the correct scenarios. Each compute environment may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.

Answer:

Explanation:

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

Box 1: nb_server

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

Box 2: mlc_cluster

With Azure Machine Learning, you can train your model on a variety of resources or environments, collectively referred to as compute targets. A compute target can be a local machine or a cloud resource, such as an Azure Machine Learning Compute, Azure HDInsight or a remote virtual machine.

Reference:

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

https://docs.microsoft.com/en-us/azure/machine-learning/how-to-set-up-training-targets

NEW QUESTION: 156

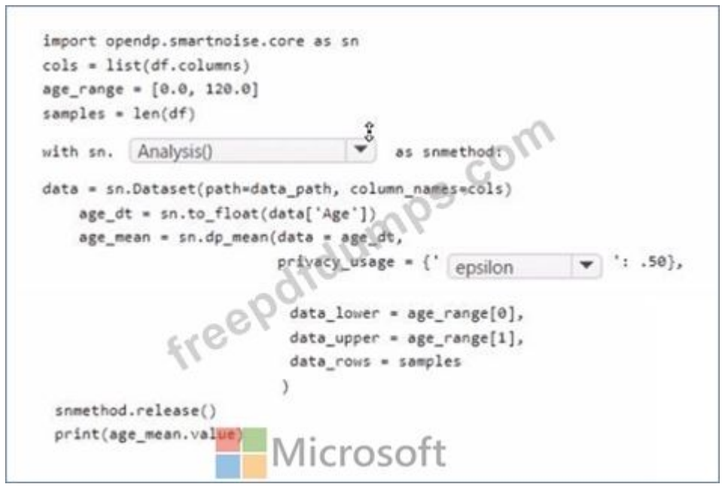
You load data from a notebook in an Azure Machine Learning workspace into a pandas dataframe named df.

The data contains 10.000 patient records. Each record includes the Age property for the corresponding patient.

You must identify the mean age value from the differentially private data generated by SmartNoise SDK.
You need to complete the Python code that will generate the mean age value from the differentially private data.
Which code segments should you use? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.

Answer:

Explanation:



NEW QUESTION: 157

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

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

You are using Azure Machine Learning to run an experiment that trains a classification model.

You want to use Hyperdrive to find parameters that optimize the AUC metric for the model. You configure a HyperDriveConfig for the experiment by running the following code:

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

You plan to use this configuration to run a script that trains a random forest model and then tests it with validation dat a. The label values for the validation data are stored in a variable named y_test variable, and the predicted probabilities from the model are stored in a variable named y_predicted.

You need to add logging to the script to allow Hyperdrive to optimize hyperparameters for the AUC metric.

Solution: Run the following code:

```
import numpy as np
from sklearn.metrics import roc_auc_score
# code to train model omitted
auc = roc_auc_score(y_test, y_predicted)
print(np.float(auc))
```

Does the solution meet the goal?

A. Yes

B. No

Answer: (SHOW ANSWER)

Use a solution with logging.info(message) instead.

Note: Python printing/logging example:

logging.info(message)
Destination: Driver logs, Azure Machine Learning designer
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-debug-pipelines

NEW QUESTION: 158
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: (SHOW ANSWER)
train_step is missing.
Reference:
https://docs.microsoft.com/en-us/python/api/azureml-pipeline-core/azureml.pipeline.core.pipelinedata?
view=azure-ml-py

NEW QUESTION: 159
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.
After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.
You create an Azure Machine Learning service datastore in a workspace. The datastore contains the following files:
* /data/2018/Q1 .csv
* /data/2018/Q2.csv
* /data/2018/Q3.csv
* /data/2018/Q4.csv
* /data/2019/Q1.csv
All files store data in the following format:

id,f1,f2,l
1,1,2,0
2,1,1,1
3.2.1.0
You run the following code:

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


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

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

Solution: Run the following code:

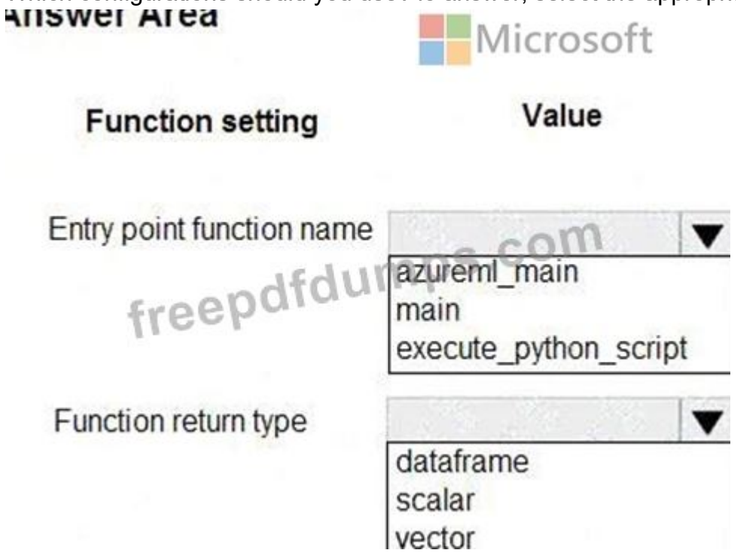
```
from azureml.core import Dataset
paths = [(data_store,'data/2018/*.csv'),(data_store,'data/2019/*.csv')]
training_data = Dataset.Tabular.from_delimited_files(paths)
```

Does the solution meet the goal?
A. Yes
B. No
Answer: A (LEAVE A REPLY)
Use two file paths.
Use Dataset.Tabular_from_delimeted as the data isn ' t cleansed.
Note:
A TabularDataset represents data in a tabular format by parsing the provided file or list of files. This provides you with the ability to materialize the data into a pandas or Spark DataFrame so you can work with familiar data preparation and training libraries without having to leave your notebook. You can create a TabularDataset object from .csv, .tsv, .parquet, .jsonl files, and from SQL query results.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-create-register-datasets

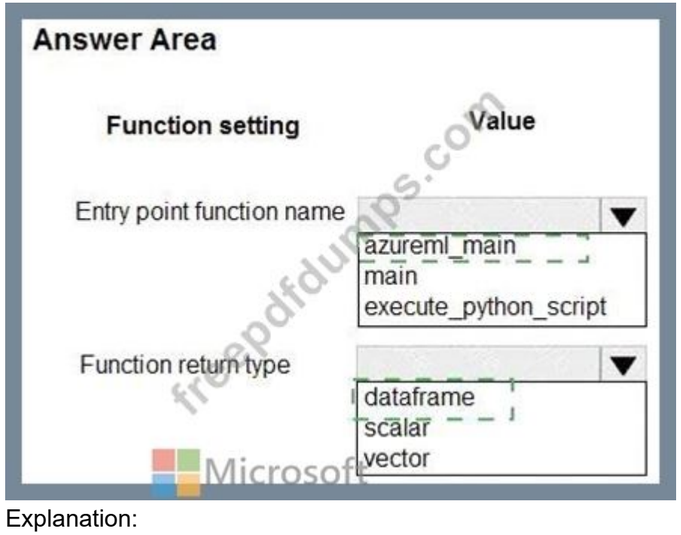
NEW QUESTION: 160
You use an Azure Machine Learning workspace.
You have a trained model that must be deployed as a web service. Users must authenticate by using Azure Active Directory.
What should you do?
A. Deploy the model to Azure Kubernetes Service (AKS). During deployment, set the token_auth_enabled parameter of the target configuration object to true
B. Deploy the model to Azure Container Instances. During deployment, set the auch_enabled parameter of the target configuration object to true
C. Deploy the model to Azure Container Instances. During deployment, set the coken_auch_enabled parameter of the target configuration object to true
D. Deploy the model to Azure Kubernetes Service (AKS). During deployment, set the auch. enabled parameter of the target configuration object to true
Answer: A (LEAVE A REPLY)
To control token authentication, use the token_auth_enabled parameter when you create or update a deployment Token authentication is disabled by default when you deploy to Azure Kubernetes Service.
Note: The model deployments created by Azure Machine Learning can be configured to use one of two authentication methods:
key-based: A static key is used to authenticate to the web service.
token-based: A temporary token must be obtained from the Azure Machine Learning workspace (using Azure Active Directory) and used to authenticate to the web service.

Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-authenticate-web-service

NEW QUESTION: 161
You create an Azure Machine Learning dataset containing automobile price data The dataset includes 10,000 rows and 10 columns You use Azure Machine Learning Designer to transform the dataset by using an Execute Python Script component and custom code.
The code must combine three columns to create a new column.
You need to configure the code function.
Which configurations should you use? lo answer, select the appropriate options in the answer area NOTE: Each correct selection is worth one point.



Answer:



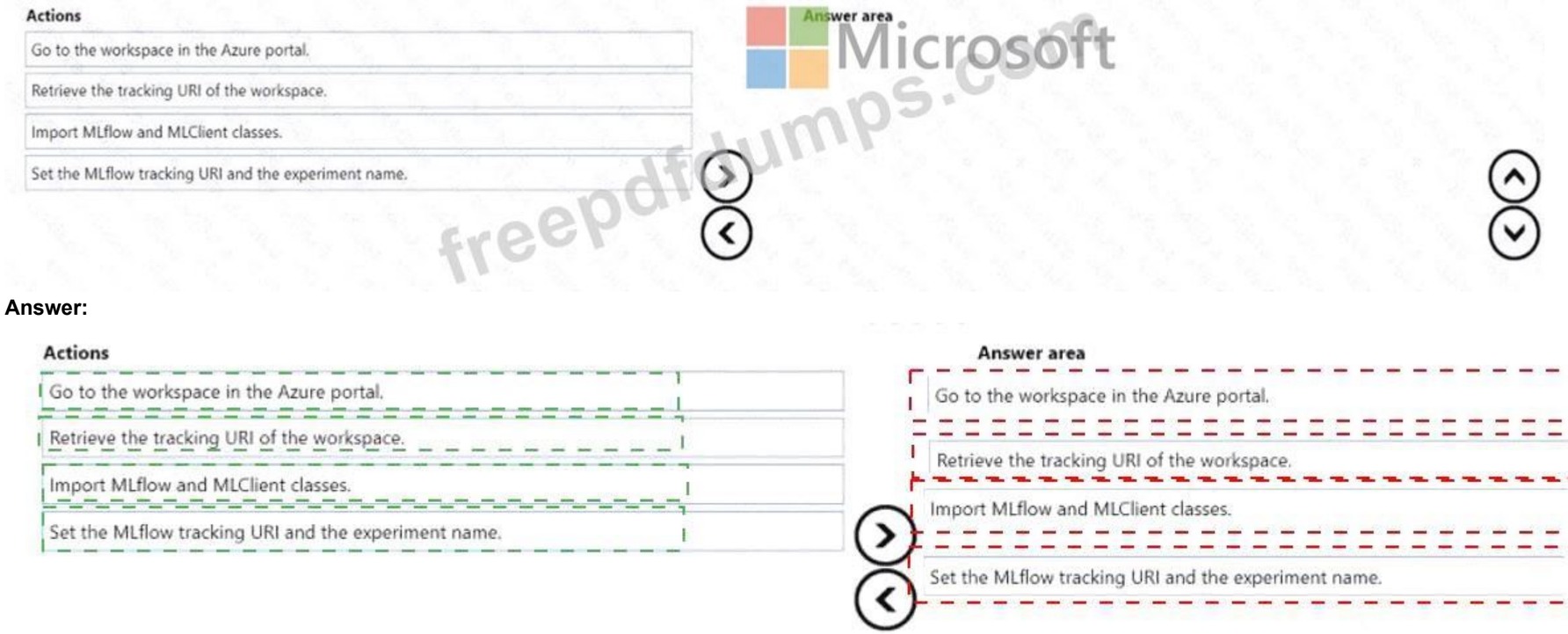
Explanation:

NEW QUESTION: 162

You have an Azure Machine Learning workspace. You are running an experiment on your local computer.

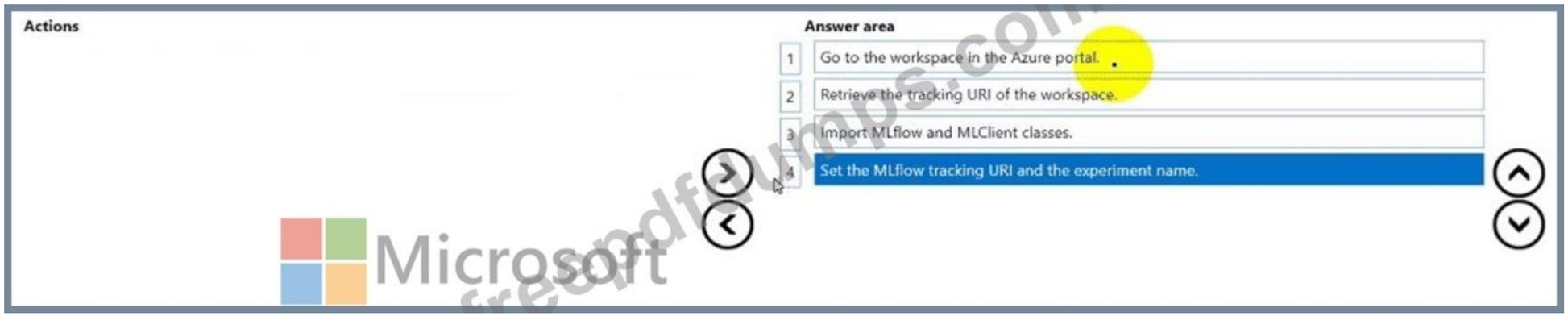
You need to ensure that you can use MLflow Tracking with Azure Machine Learning Python SDK v2 to store metrics and artifacts from your local experiment runs in the workspace.

In which order should you perform the actions? To answer, move all actions from the list of actions to the answer area and arrange them in the correct order.



Answer:

Explanation:



NEW QUESTION: 163
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: (SHOW ANSWER)
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.

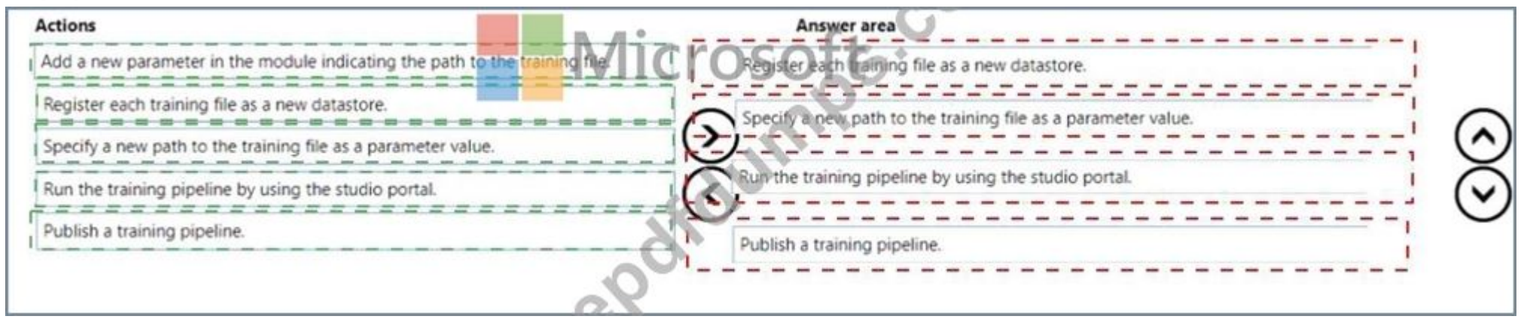
NEW QUESTION: 164
You use Azure Machine Learning designer to create a training pipeline for a regression model.
You need to prepare the pipeline for deployment as an endpoint that generates predictions asynchronously for a dataset of input data values.
What should you do?
A. Clone the training pipeline.
B. Create a batch inference pipeline from the training pipeline.
C. Create a real-time inference pipeline from the training pipeline.
D. Replace the dataset in the training pipeline with an Enter Data Manually module.
Answer: C (LEAVE A REPLY)
You must first convert the training pipeline into a real-time inference pipeline. This process removes training modules and adds web service inputs and outputs to handle requests.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/tutorial-designer-automobile-price-deploy
https://docs.microsoft.com/en-us/azure/machine-learning/algorithm-module-reference/enter-data-manually

NEW QUESTION: 165
You are analyzing a dataset containing historical data from a local taxi company. You arc developing a regression a regression model.
You must predict the fare of a taxi trip.
You need to select performance metrics to correctly evaluate the- regression model.
Which two metrics can you use? Each correct answer presents a complete solution.
NOTE: Each correct selection is worth one point.
A. an F1 score that is high
B. an R Squared value dose to 1
C. an R-Squared value close to 0
D. a Root Mean Square Error value that is high
E. a Root Mean Square Error value that is low
F. an F 1 score that is low.
Answer: B,E (LEAVE A REPLY)
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/evaluate-model

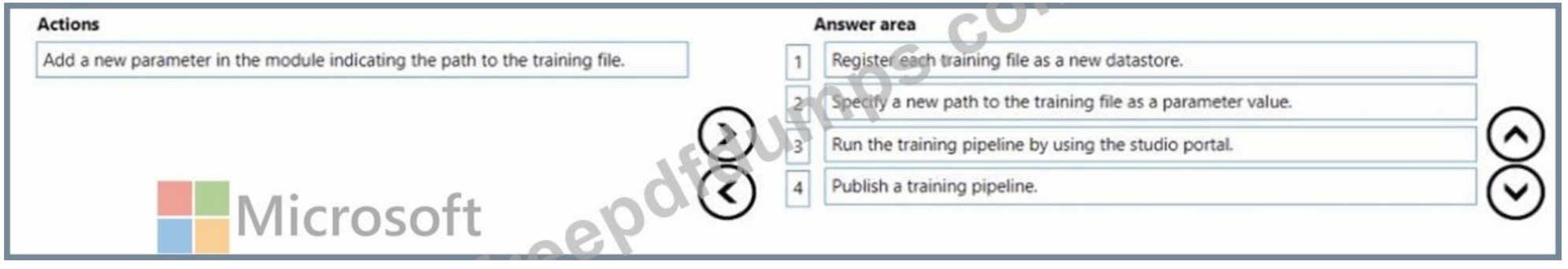
NEW QUESTION: 166
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.
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 an Azure Machine Learning model.
You plan to deploy the model to an online end point.
You need to ensure that applications will be able to use the authentication method with a non-expiring artifact to access the model.
Solution:
Create a Kubernetes online endpoint and set the value of its auth-mode parameter to amyl Token. Deploy the model to the online endpoint.
Does the solution meet the goal?
A. Yes
B. No
Answer: (SHOW ANSWER)

NEW QUESTION: 167
You use a training pipeline in the Azure Machine Learning designer. You register a datastore named ds1. The datastore contains multiple training data files. You use the Import Data module with the configured datastore.
You need to retrain a model on a different set of data files.
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.

Answer:



Explanation:



NEW QUESTION: 168

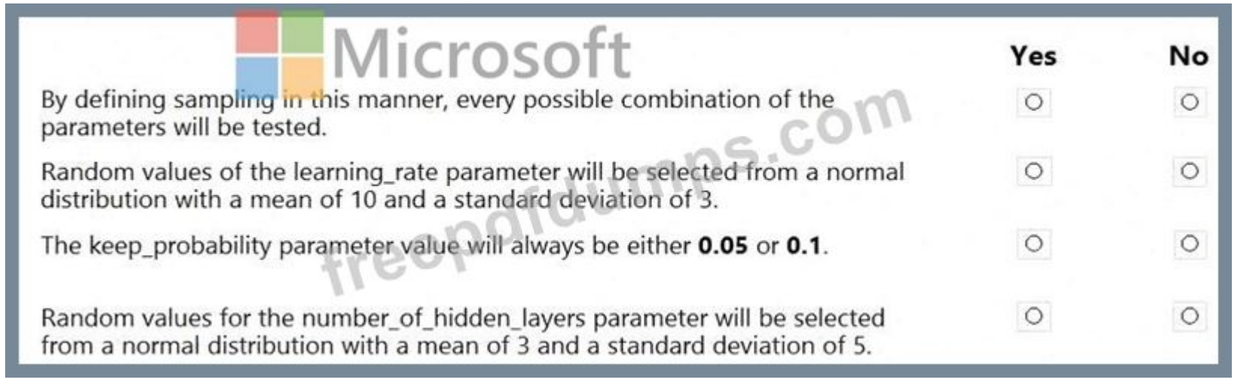
You are using the Hyperdrive feature in Azure Machine Learning to train a model.

You configure the Hyperdrive experiment by running the following code:

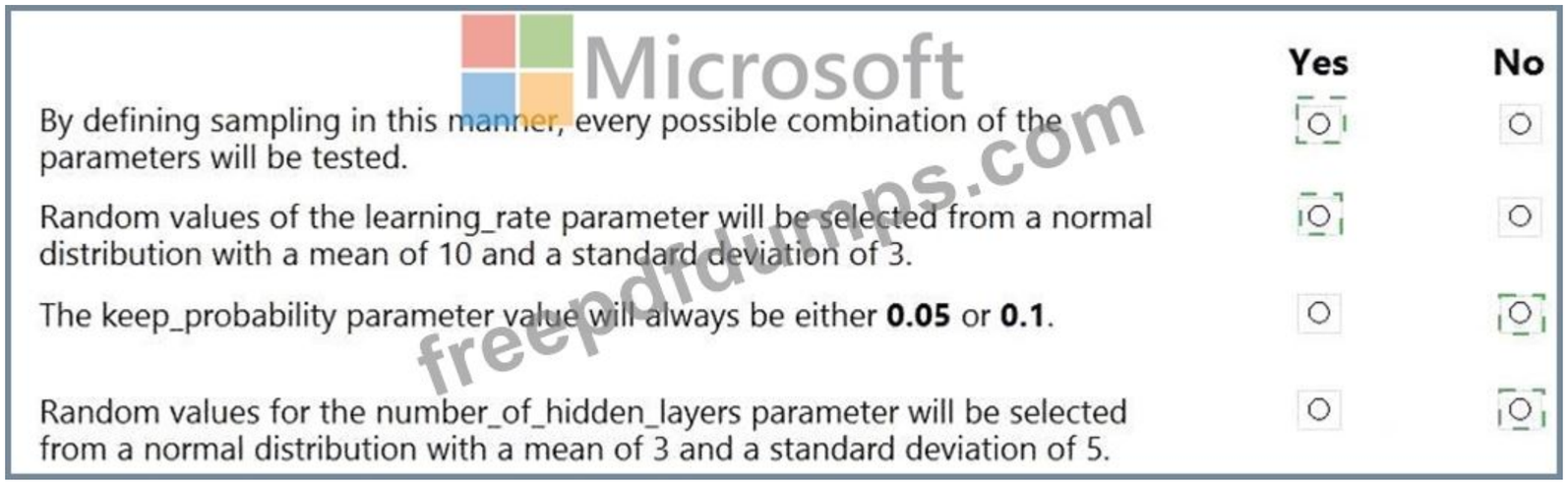
```
from azureml.train.hyperdrive import RandomParameterSampling
param_sampling = RandomParameterSampling( {
      "learning_rate": normal(10, 3),
      "keep_probability": uniform(0.05, 0.1),
      "batch_size": choice(16, 32, 64, 128)
      "number_of_hidden_layers": choice(range(3,5))
   }
)
```

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

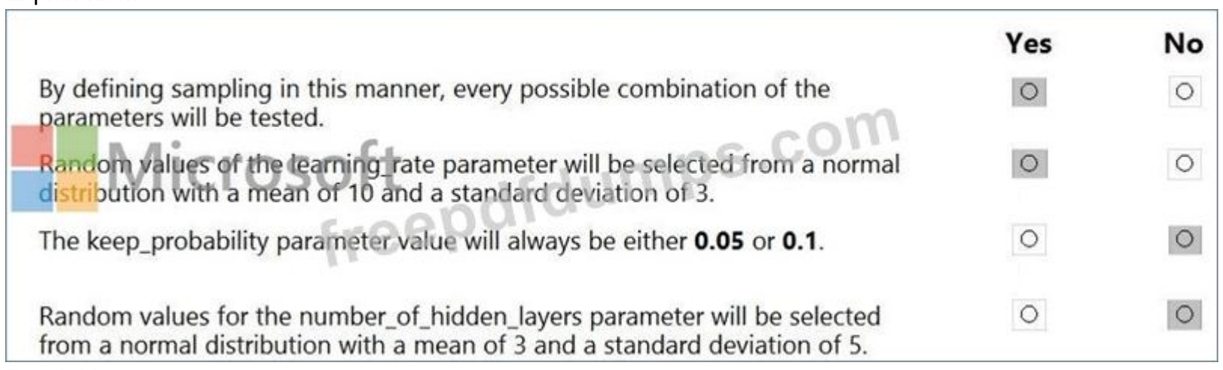
NOTE: Each correct selection is worth one point.



Answer:



Explanation:



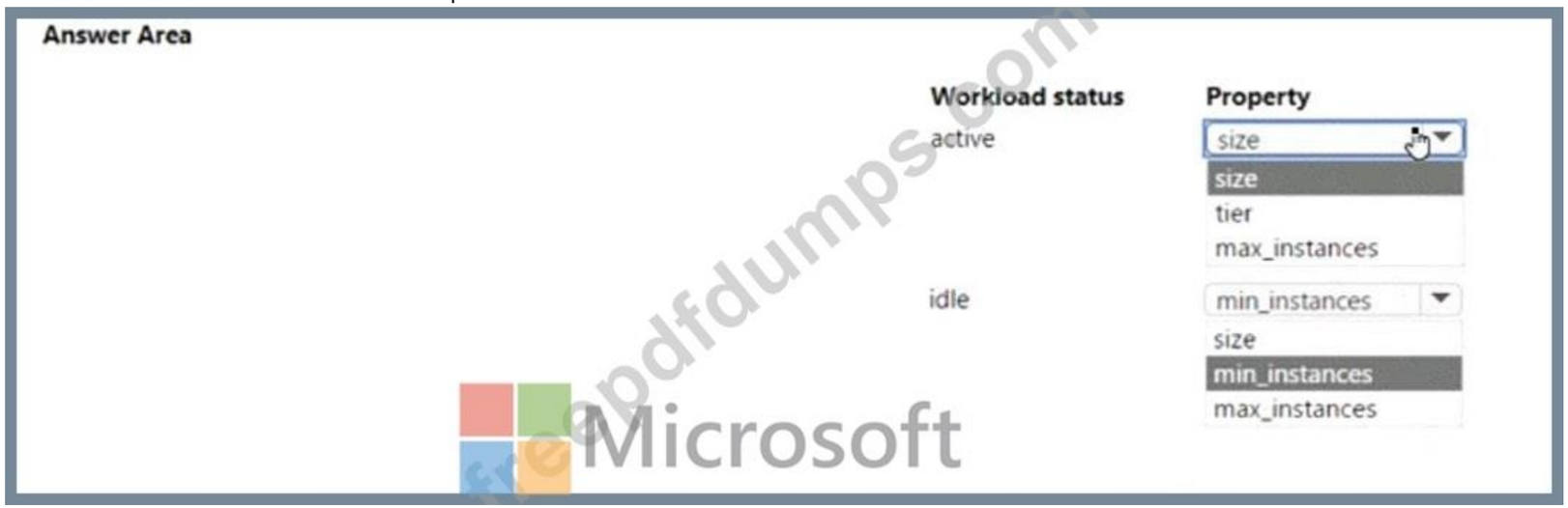
Box 1: Yes

In random sampling, hyperparameter values are randomly selected from the defined search space. Random sampling allows the search space to include both discrete and continuous hyperparameters.

Box 2: Yes

learning_rate has a normal distribution with mean value 10 and a standard deviation of 3.
Box 3: No
keep_probability has a uniform distribution with a minimum value of 0.05 and a maximum value of 0.1.
Box 4: No
number_of_hidden_layers takes on one of the values [3, 4, 5].
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-tune-hyperparameters

NEW QUESTION: 169
You manage an Azure Machine Learning workspace by using the Python SDK v2.
You must create a compute cluster in the workspace. The compute cluster must run workloads and properly handle interruptions. You start by calculating the maximum amount of compute resources required by the workloads and size the cluster to match the calculations.
The cluster definition includes the following properties and values:
* name= " mlcluster1"
* size= " STANDARD.DS3.v2 "
* min_instances=1
* maxjnstances=4
* tier= " dedicated "
The cost of the compute resources must be minimized when a workload is active Of idle. Cluster property changes must not affect the maximum amount of compute resources available to the workloads run on the cluster.
You need to modify the cluster properties to minimize the cost of compute resources.
Which properties should you modify? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.



Answer:

Explanation:

NEW QUESTION: 170

You manage an Azure Machine Learning workspace. The development environment for managing the workspace is configured to use Python SDK v2 in Azure Machine Learning Notebooks.

A Synapse Spark Compute is currently attached and uses system-assigned identity.

You need to use Python code to update the Synapse Spark Compute to use a user-assigned identity.

Solution: Pass the UserAssignedIdentity class object to the SynapseSparkCompute class.

Does the solution meet the goat?

A. Yes

B. No

Answer: B (LEAVE A REPLY)

NEW QUESTION: 171

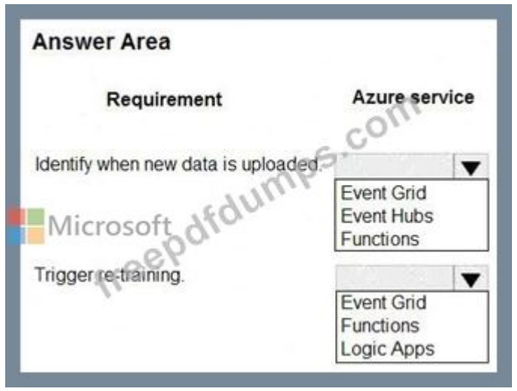
You train a model by using Azure Machine Learning. You use Azure Blob Storage to store production data.

The model must be re-trained when new data is uploaded to Azure Blob Storage. You need to minimize development and coding.

You need to configure Azure services to develop a re-training solution.

Which Azure services should you use? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.



Answer:

Explanation:

To set up a solution for retraining a model when new data is uploaded to Azure Blob Storage, you can use a combination of Azure services. Here are the appropriate services to use for each requirement:

Identify when new data is uploaded:

Event Grid: Azure Event Grid can be used to monitor Azure Blob Storage for events, such as the upload of new data. It is designed to handle events in a scalable way, making it a suitable choice for identifying new data uploads.

Trigger re-training:

Logic Apps: Azure Logic Apps can be used to create workflows that automate the process of retraining your model. It allows you to create workflows with minimal code and can integrate with various Azure services, including Azure Machine Learning and Event Grid.

So, the correct selections are:

Identify when new data is uploaded: Event Grid

Trigger re-training: Logic Apps

NEW QUESTION: 172

You create an Azure Machine Learning workspace. You use Azure Machine Learning designer to create a pipeline within the workspace. You need to submit a pipeline run from the designer.

What should you do first?

A. Create a compute cluster.

B. Select a model.

C. Create an experiment.
D. Create an attached compute resource.
Answer: (SHOW ANSWER)

NEW QUESTION: 173
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 (LEAVE A REPLY)
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

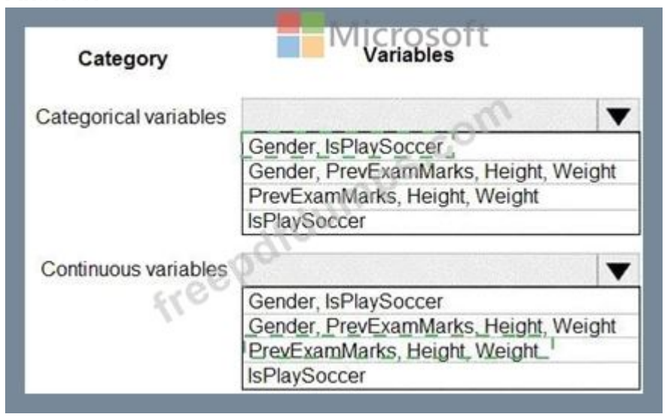
NEW QUESTION: 174
You run Azure Machine Learning training experiments. The training scripts directory contains 100 files that includes a file named. amlignore. The directory also contains subdirectories named. /outputs and./logs.
There are 20 files in the training scripts directory that must be excluded from the snapshot to the compute targets. You create a file named. gift ignore in the root of the directory. You add the names of the 20 files to the. gift ignore file. These 20 files continue to be copied to the compute targets.
You need to exclude the 20 files. What should you do?
A. Copy the contents of the file named. gift ignore to the file named. amlignore.
B. Move the file named. gift ignore to the. /outputs directory.
C. Add the contents of the file named. amlignore to the file named. gift ignore.
D. Move the file named. gift ignore to the. /logs directory.
Answer: (SHOW ANSWER)

NEW QUESTION: 175
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:

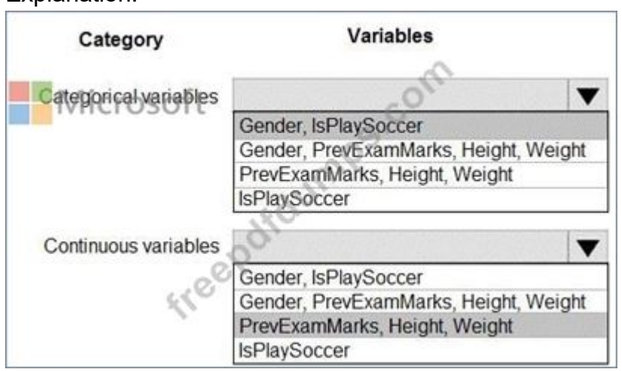
Name	Description
IsPlaySoccer	Values can be 1 and 0.
Gender	Values can be M or F.
PrevExamMarks	Stores values from 0 to 100
Height	Stores values in centimeters
Weight	Stores values in kilograms

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.

Answer:



Explanation:



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

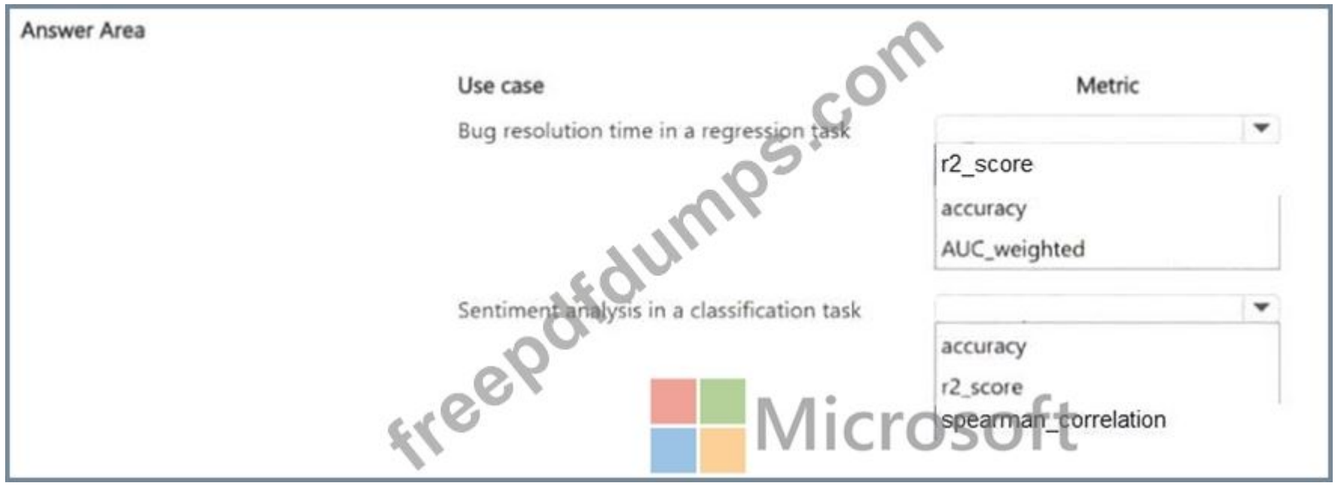
NEW QUESTION: 176

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 True.
Deploy the model to the service.
Does the solution meet the goal?
A. Yes
B. No
Answer: A (LEAVE A REPLY)
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: 177
You manage an Azure Machine Learning workspace.
You must provide explanations for the behavior of the models with feature importance measures.
You need to configure a Responsible AI dashboard in Azure Machine Learning.
Which dashboard component should you configure?
A. Fairness assessment
B. Casual inference
C. Interpretability
D. Counterfactual what-if
Answer: C (LEAVE A REPLY)

NEW QUESTION: 178
You create multiple machine learning models by using automated machine learning.
You need to configure a primary metric for each use case.
Which metrics should you configure? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.



Answer:

Explanation:

NEW QUESTION: 179

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.

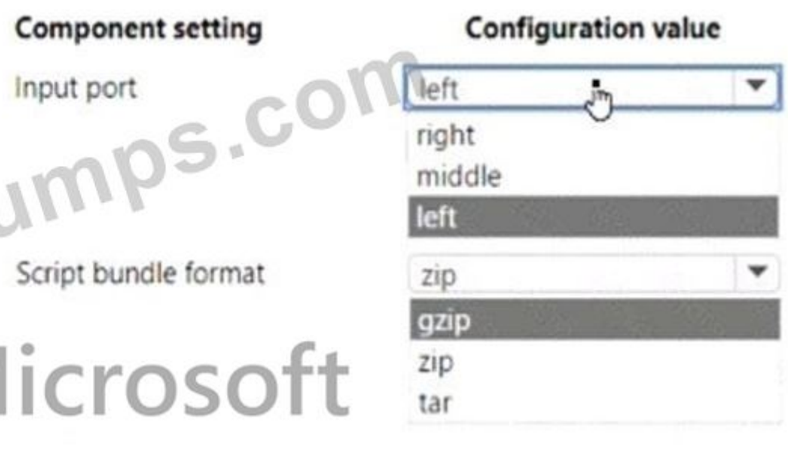
Which three tasks will you be able to perform? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

D

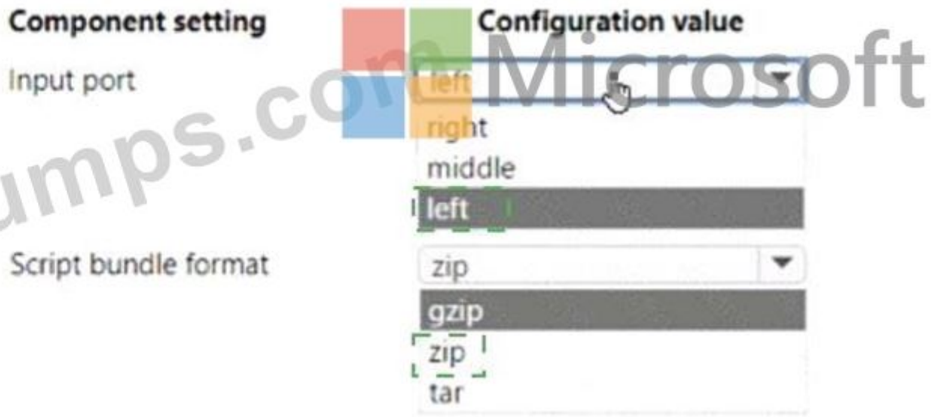
A. Create a Compute Instance and use it to run code in Jupyter notebooks.
B. Create an Azure Kubernetes Service (AKS) inference cluster.
C. Use the designer to train a model by dragging and dropping pre-defined modules.
D. Create a tabular dataset that supports versioning.
E. Use the Automated Machine Learning user interface to train a model.
Answer: A,B,D (LEAVE A REPLY)
Reference:
https://azure.microsoft.com/en-us/pricing/details/machine-learning/

NEW QUESTION: 180
You create an Azure Machine Learning dataset. You use the Azure Machine Learning designer to transform the dataset by using an Execute Python Script component and custom code.
You must upload the script and associated libraries as a script bundle.
You need to configure the Execute Python Script component.
Which configurations should you use? To answer, select the appropriate options in the answer area.
NOTE Each correct selection is worth one point.
Answer Area

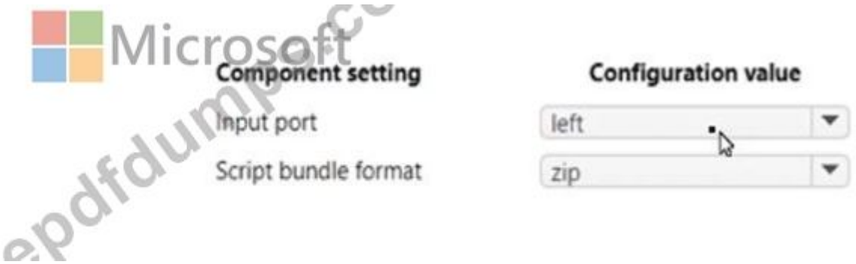


Answer:

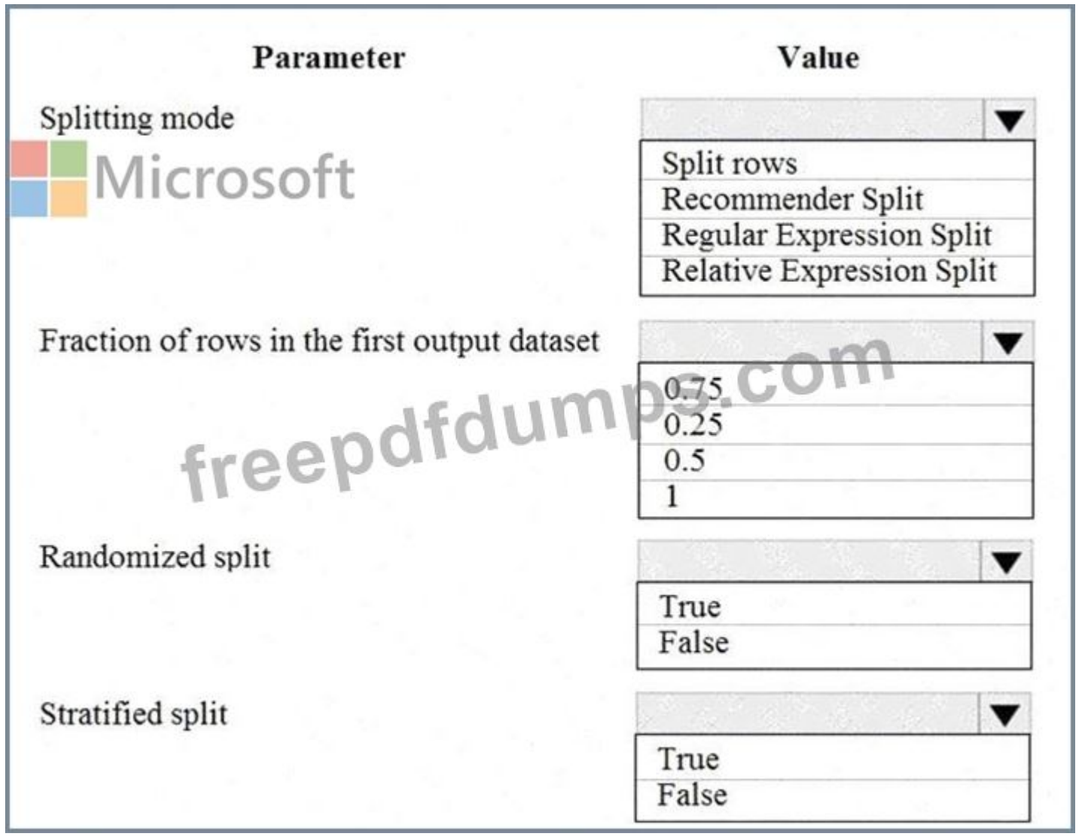
Answer Area



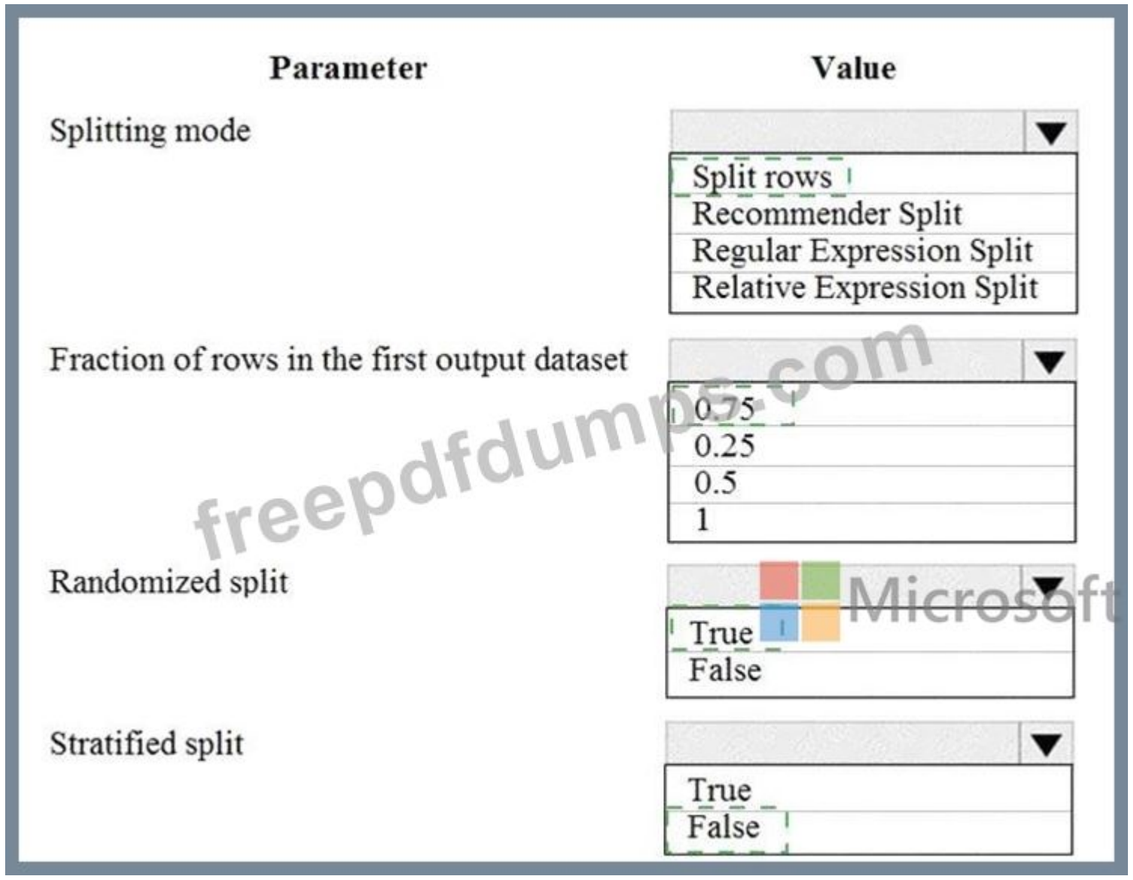
Explanation:
Answer Area



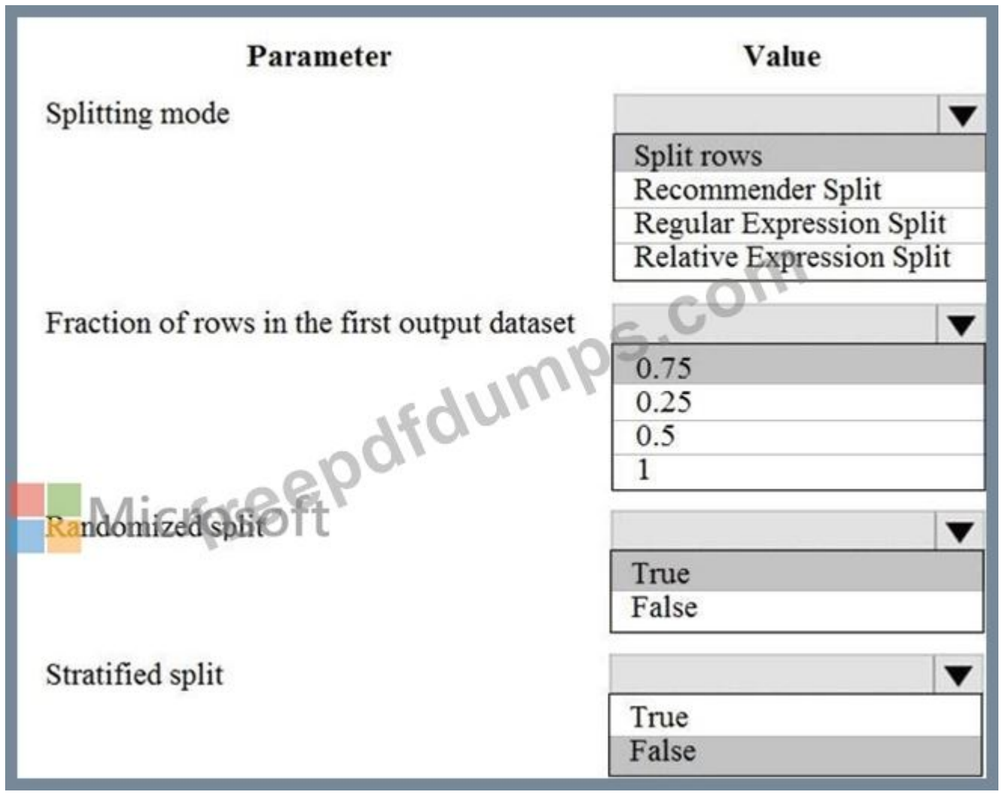
NEW QUESTION: 181
You are performing a classification task in Azure Machine Learning Studio.
You must prepare balanced testing and training samples based on a provided data set.
You need to split the data with a 0.75:0.25 ratio.
Which value should you use for each parameter? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.



Answer:



Explanation:



Box 1: Split rows

Use the Split Rows option if you just want to divide the data into two parts. You can specify the percentage of data to put in each split, but by default, the data is divided 50-50.

You can also randomize the selection of rows in each group, and use stratified sampling. In stratified sampling, you must select a single column of data for which you want values to be apportioned equally among the two result datasets.

Box 2: 0.75

If you specify a number as a percentage, or if you use a string that contains the " % " character, the value is interpreted as a percentage. All percentage values must be within the range (0, 100), not including the values 0 and 100.

Box 3: Yes

To ensure splits are balanced.

Box 4: No

If you use the option for a stratified split, the output datasets can be further divided by subgroups, by selecting a strata column.

Reference:

https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/split-data

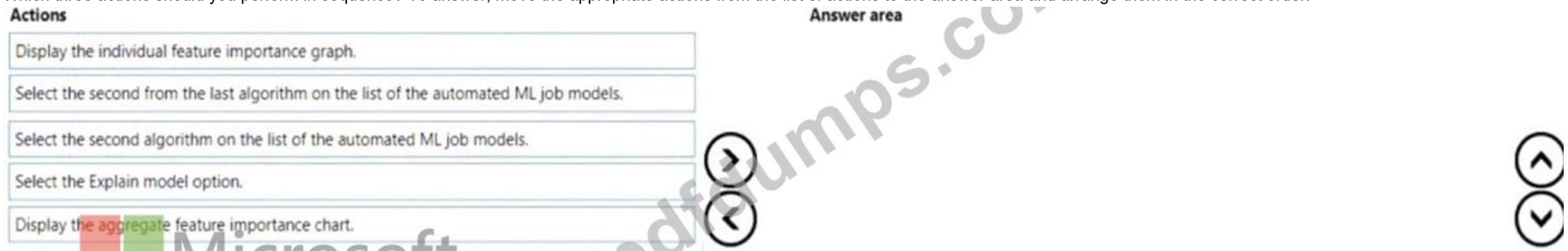
NEW QUESTION: 182

You create an Azure Machine Learning workspace. You are training a classification model with no-code AutoML in Azure Machine Learning studio.

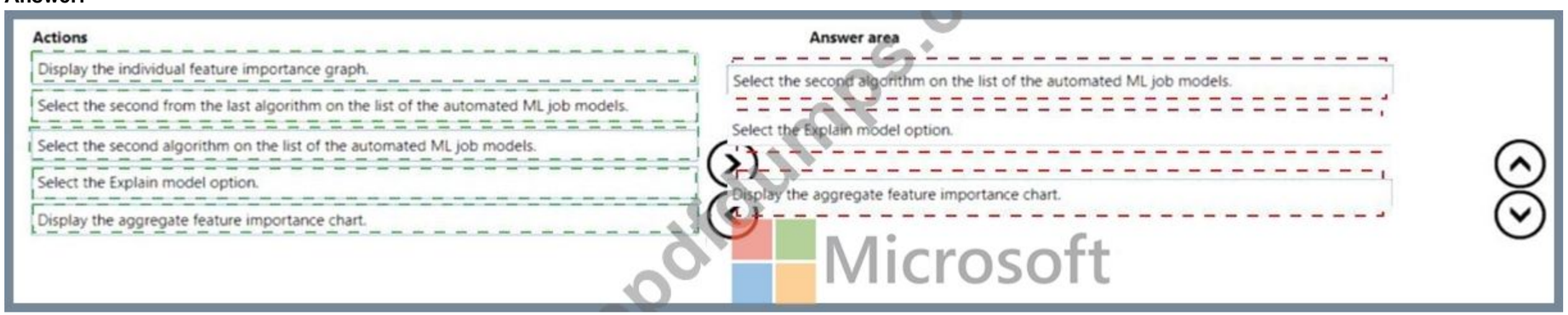
The model must predict if a client of a financial institution will subscribe to a fixed-term deposit. You must identify the feature that has the most influence on the predictions of the model for the second highest scoring algorithm. You must minimize the effort and time to identify the feature.

You need to complete the identification.

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.



Answer:



Explanation:

Actions

- Display the individual feature importance graph.
- Select the second from the last algorithm on the list of the automated ML job models.

Answer area

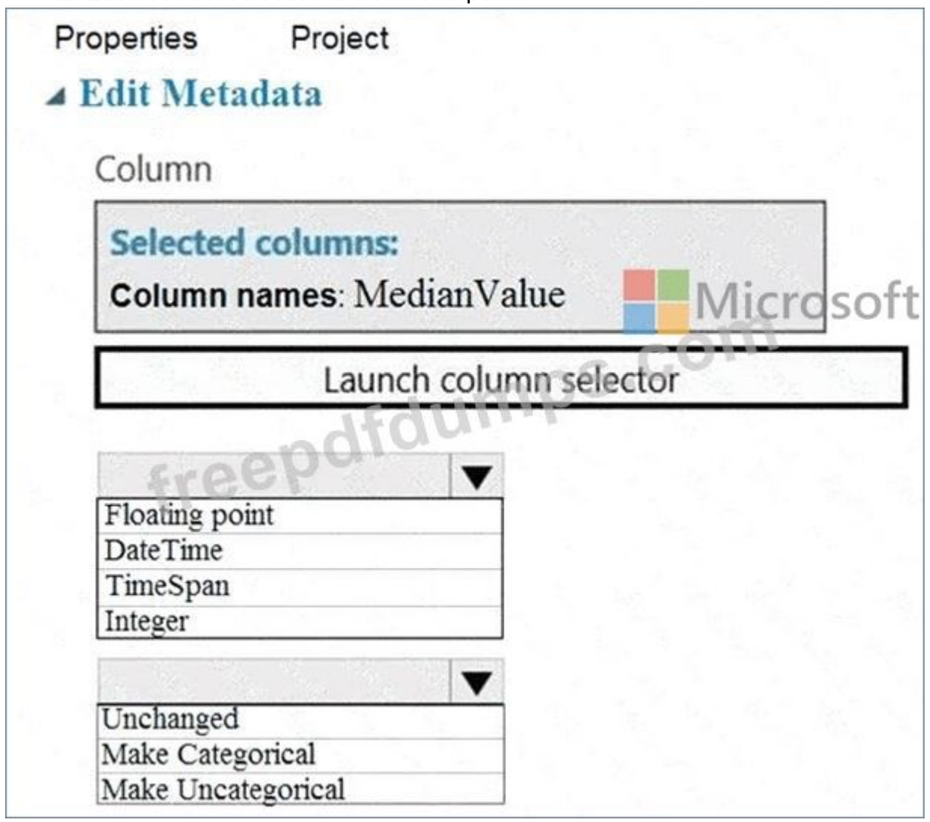
1. Select the second algorithm on the list of the automated ML job models.
2. Select the Explain model option.
3. Display the aggregate feature importance chart.

NEW QUESTION: 183

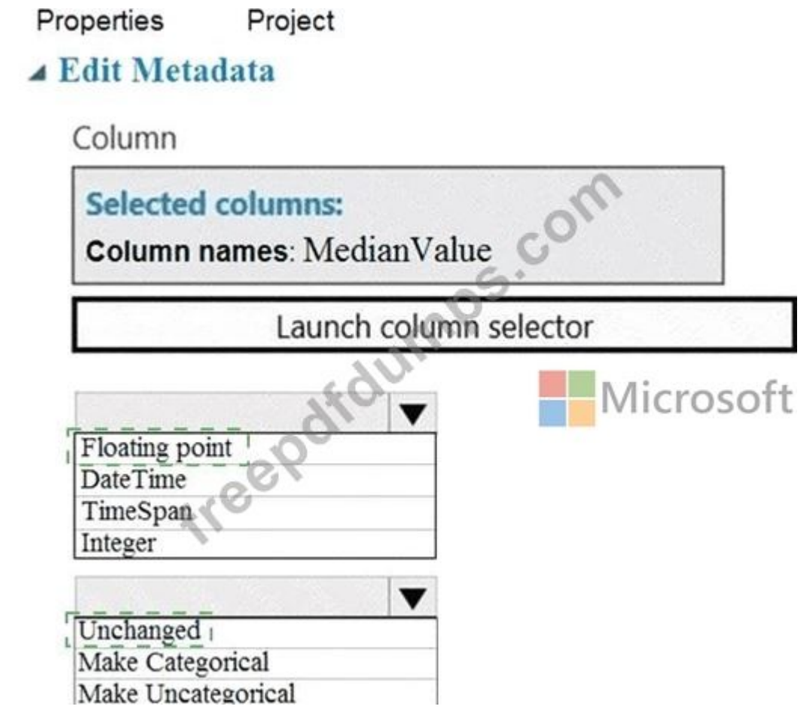
You need to configure the Edit Metadata module so that the structure of the datasets match.

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

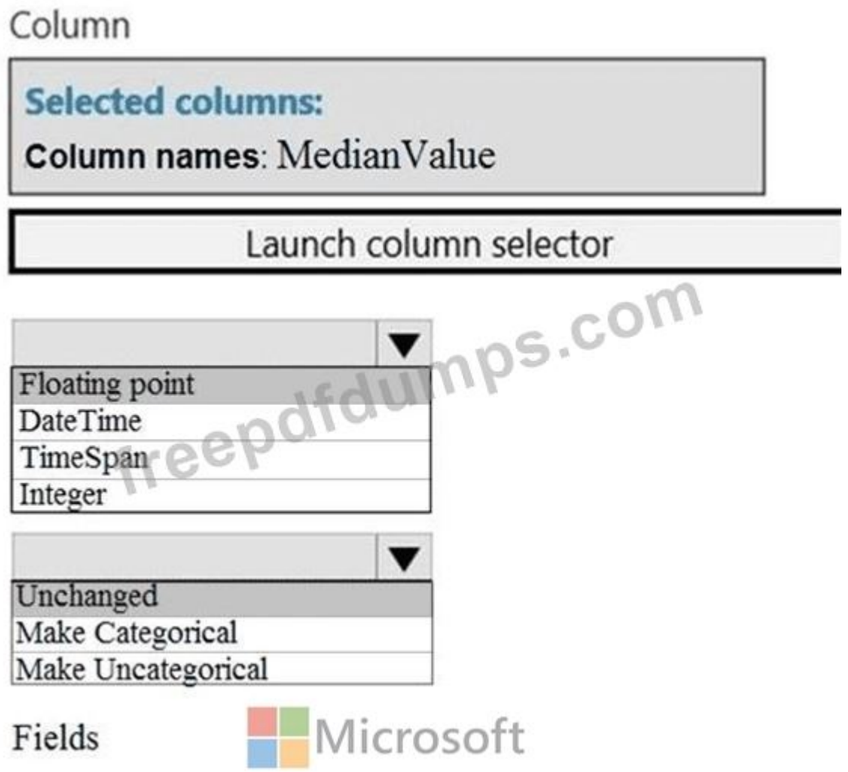
NOTE: Each correct selection is worth one point.



Answer:



Explanation:



Box 1: Floating point
Need floating point for Median values.
Scenario: An initial investigation shows that the datasets are identical in structure apart from the MedianValue column. The smaller Paris dataset contains the MedianValue in text format, whereas the larger London dataset contains the MedianValue in numerical format.
Box 2: Unchanged
Note: Select the Categorical option to specify that the values in the selected columns should be treated as categories.
For example, you might have a column that contains the numbers 0,1 and 2, but know that the numbers actually mean " Smoker " , " Non smoker " and " Unknown " . In that case, by flagging the column as categorical you can ensure that the values are not used in numeric calculations, only to group data.

NEW QUESTION: 184
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 (LEAVE A REPLY)

Instead use the Synthetic Minority Oversampling Technique (SMOTE) sampling mode.

Note: SMOTE is used to increase the number of underepresented cases in a dataset used for machine learning.

SMOTE is a better way of increasing the number of rare cases than simply duplicating existing cases.

References:

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

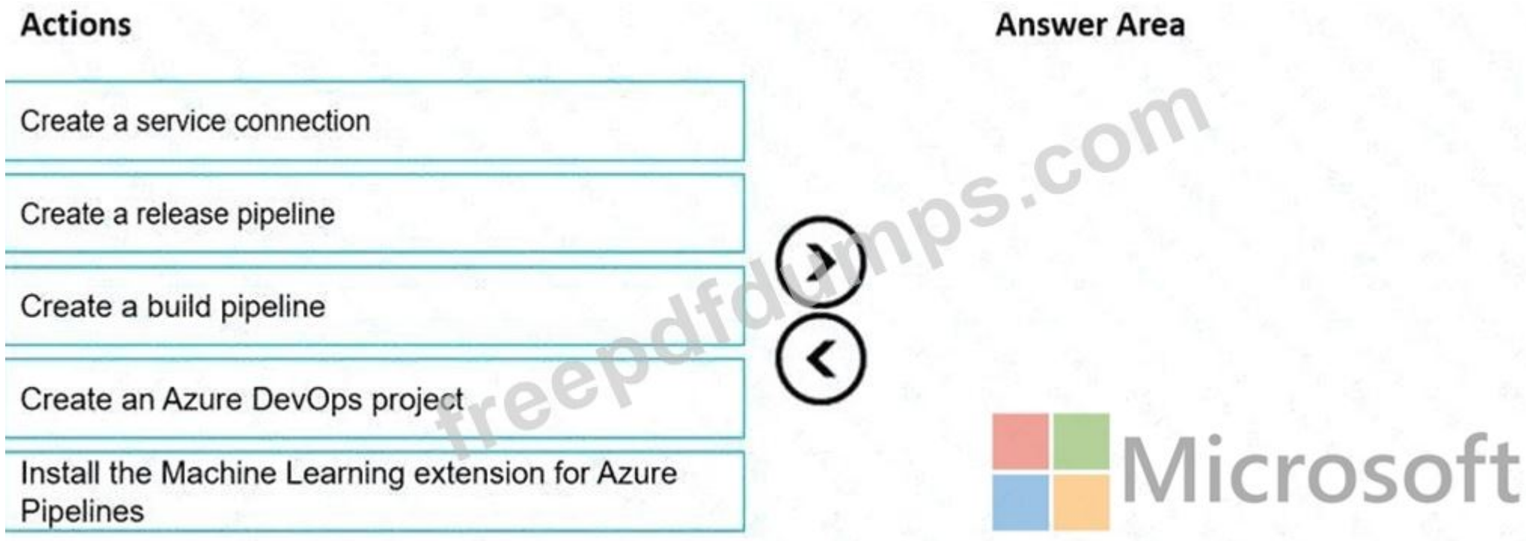
NEW QUESTION: 185

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.

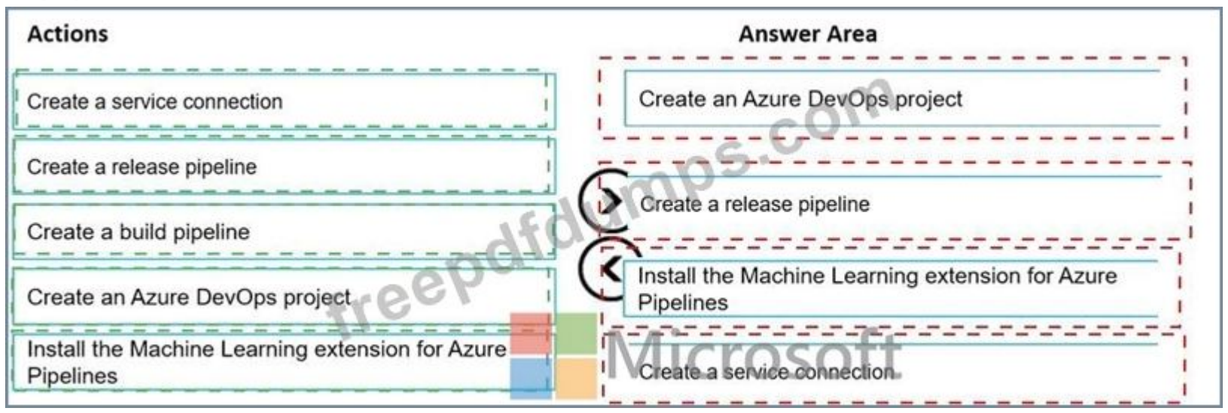
All new versions of the model registered in the workspace must automatically be deployed to the target environment.

You need to configure Azure Pipelines to deploy the model.

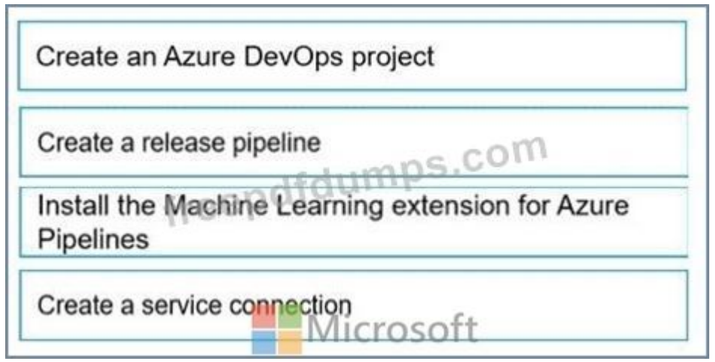
Which four actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.



Answer:



Explanation:



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.

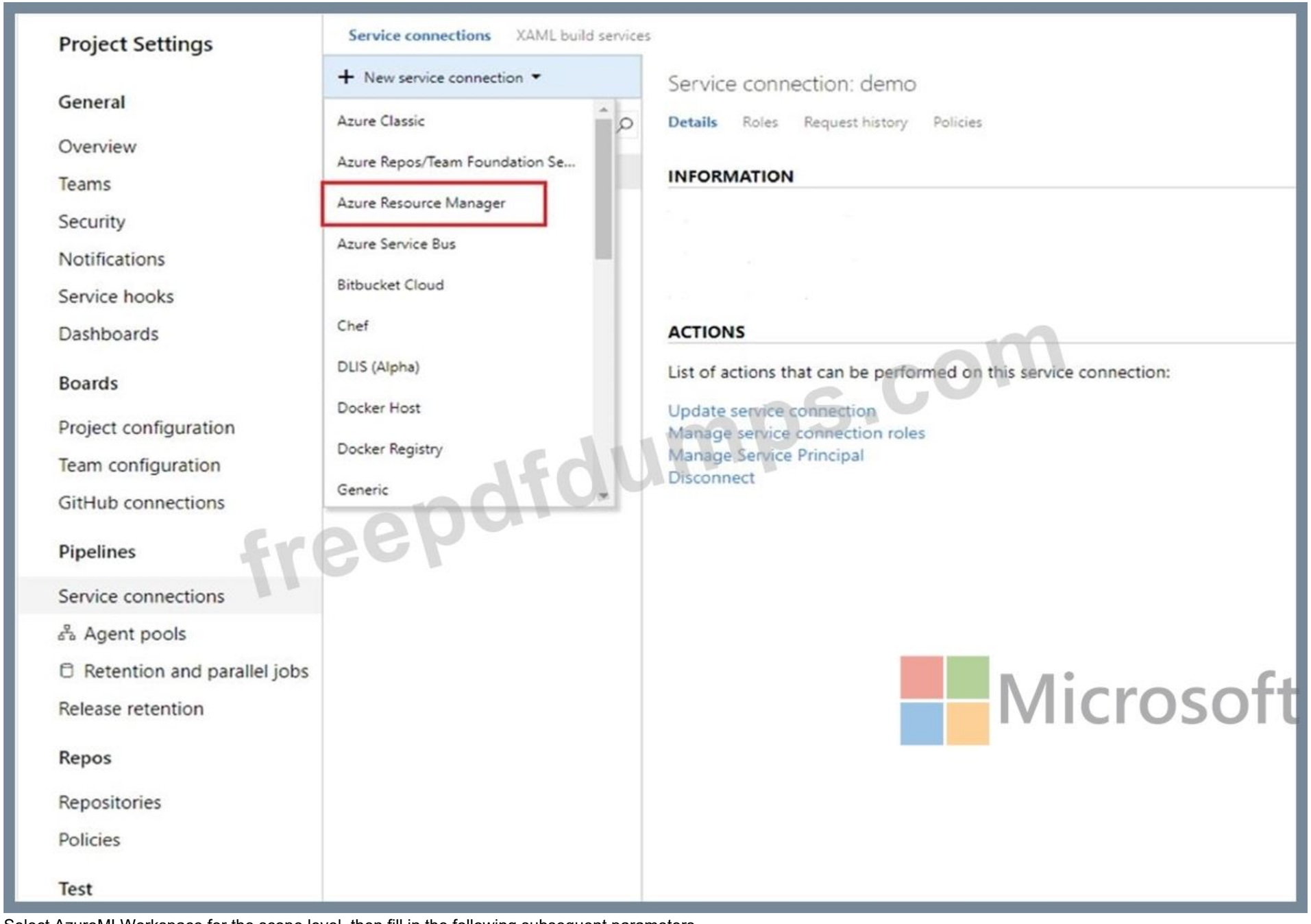
Go to Pipelines, and then select New pipeline.

Step 3: Install the Machine Learning extension for Azure Pipelines

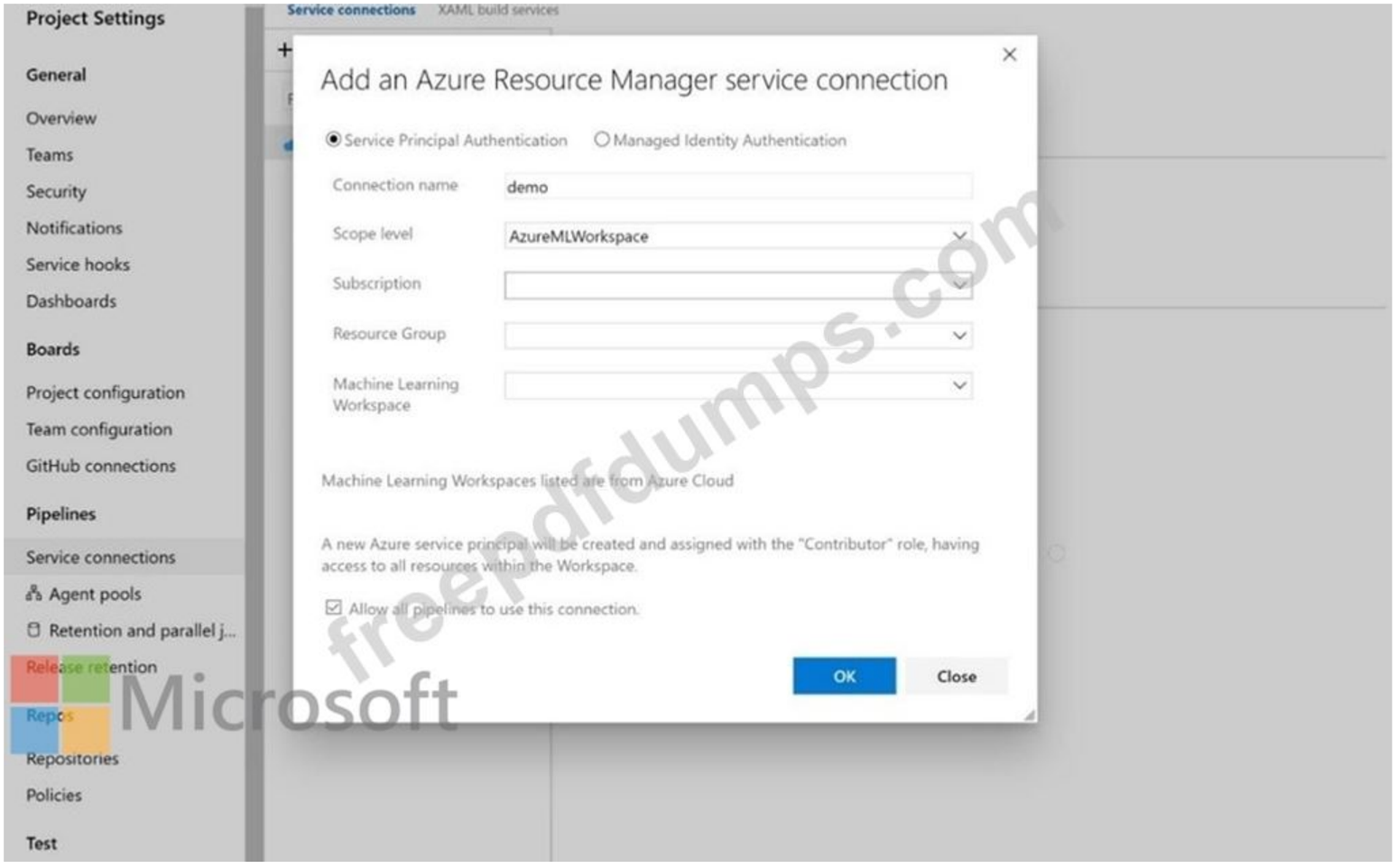
You must install and configure the Azure CLI and ML extension.

Step 4: Create a service connection

How to set up your service connection



Select AzureMLWorkspace for the scope level, then fill in the following subsequent parameters.



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: 186

You have an Azure Machine Learning workspace. You plan to tune model hyperparameters by using a sweep job.

You need to find a sampling method that supports early termination of low-performance jobs and continuous hyperpara meters.

Solution: Use the Sobol sampling method over the hyperpara meter space.

Does the solution meet the goal?

A. No

B. Yes

Answer: A (LEAVE A REPLY)

NEW QUESTION: 187

You use Azure Machine Learning Designer to load the following datasets into an experiment:

Dataset1

Age	Length	Width
3	22	13
7	11	96
18	32	85

Dataset2

Age	Length	Width
11	101	65
6	98	23
33	22	54
17	52	12

You use Azure Machine Learning Designer to load the following datasets into an experiment:

You need to create a dataset that has the same columns and header row as the input datasets and contains all rows from both input datasets.

Solution: Use the Join Data component.

Does the solution meet the goal?

A. No

B. Yes

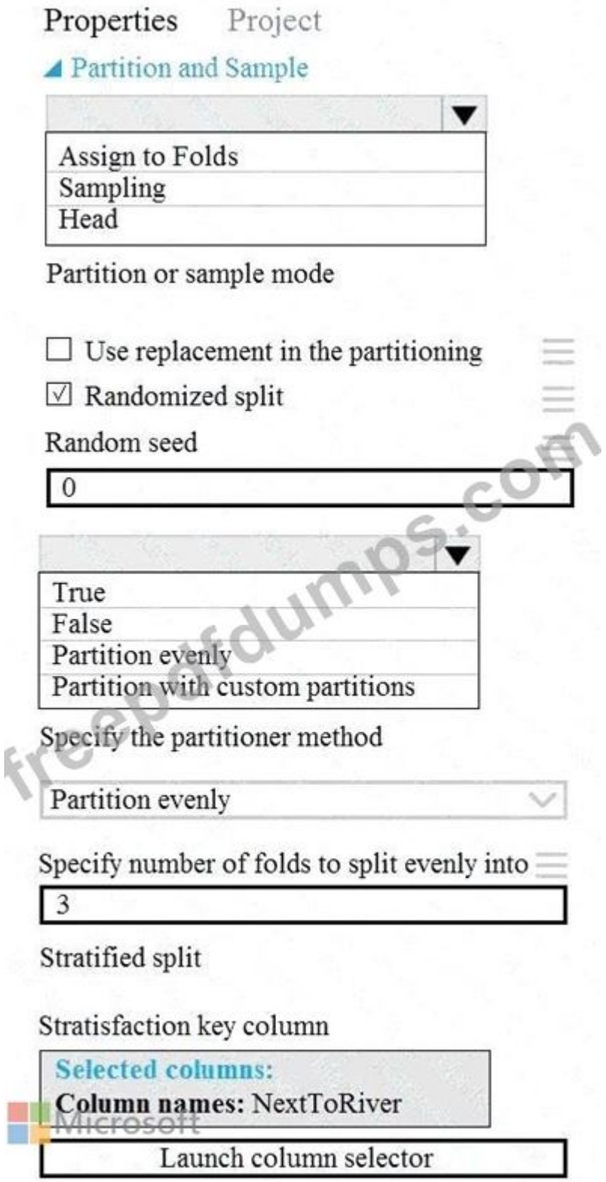
Answer: A (LEAVE A REPLY)

NEW QUESTION: 188

You need to identify the methods for dividing the data according to the testing requirements.

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

NOTE: Each correct selection is worth one point.



Answer:

Properties Project

Partition and Sample

Assign to Folds
Sampling
Head

Partition or sample mode

☐ Use replacement in the partitioning
☑ Randomized split

Random seed

0

True
False

Partition evenly
Partition with custom partitions

Specify the partitioner method

Partition evenly

Specify number of folds to split evenly into

3

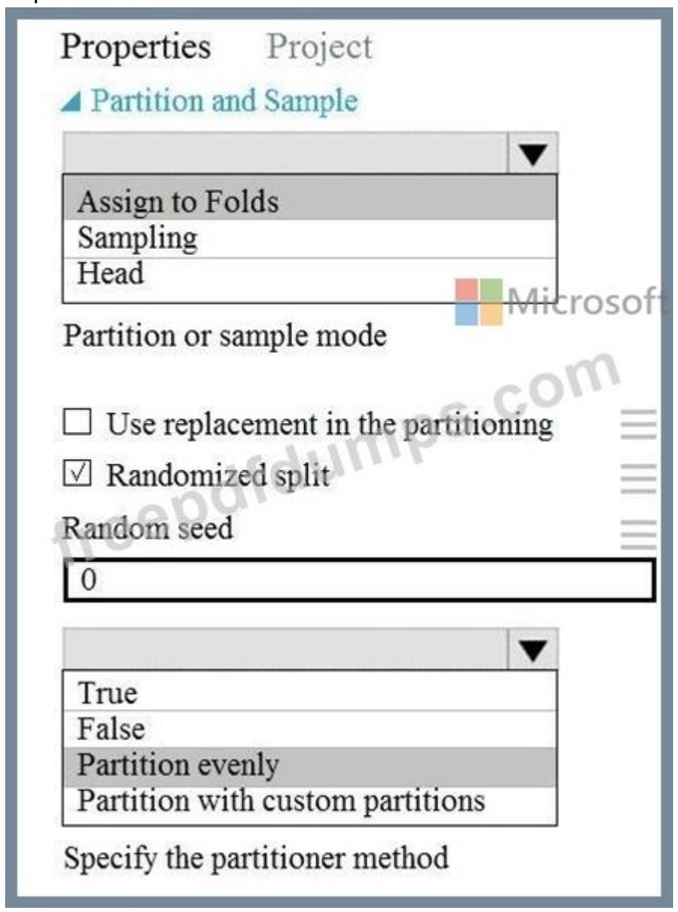
Stratified split

Stratisfaction key column

Selected columns:
Column names: NextToRiver

Launch column selector

Explanation:



Partition evenly

Specify number of folds to split evenly into

3

Stratified split

Stratisfaction key column

Selected columns:
Column names: NextToRiver

Launch column selector

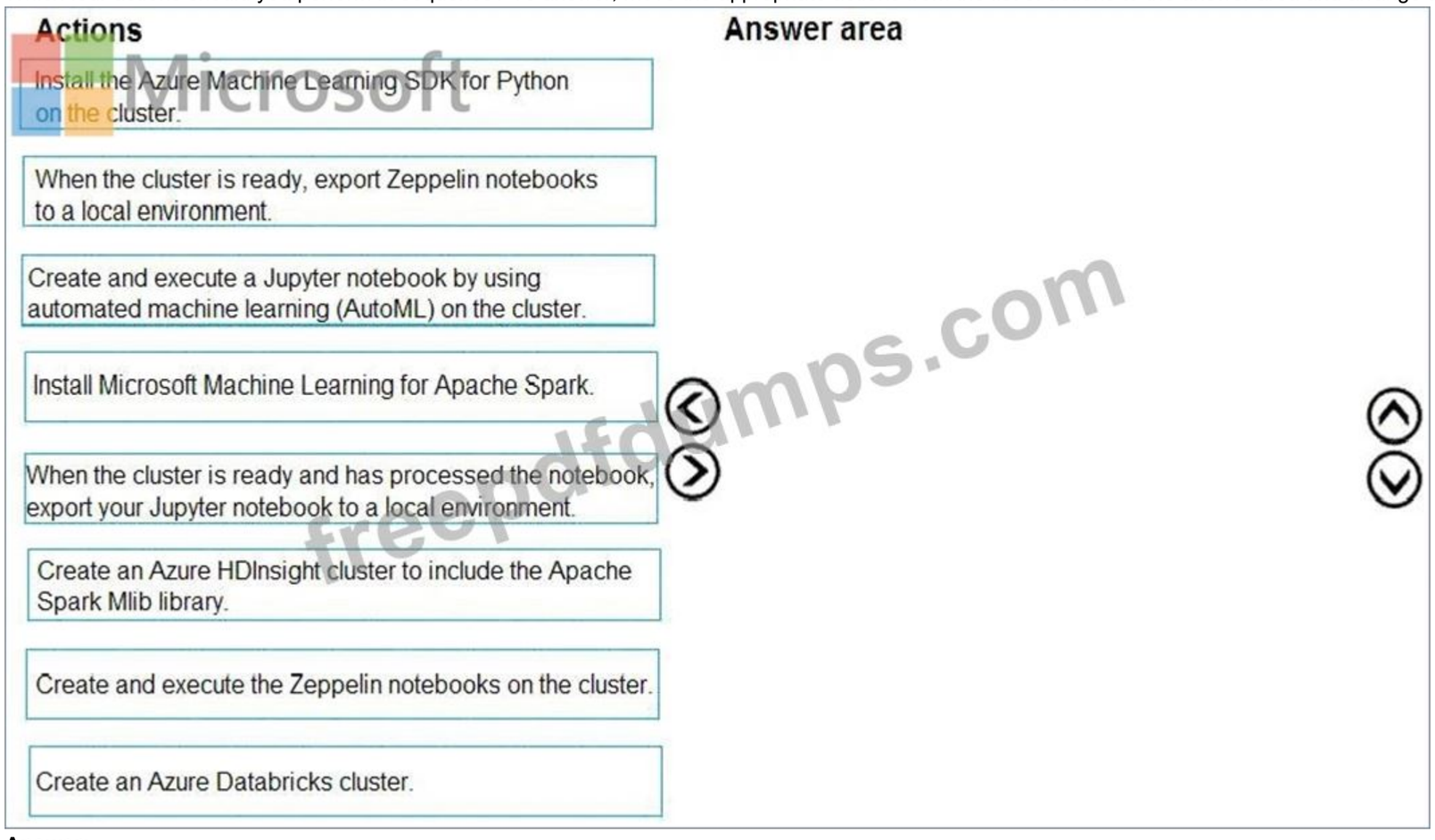
Scenario: Testing
You must produce multiple partitions of a dataset based on sampling using the Partition and Sample module in Azure Machine Learning Studio.
Box 1: Assign to folds
Use Assign to folds option when you want to divide the dataset into subsets of the data. This option is also useful when you want to create a custom number of folds for cross-validation, or to split rows into several groups.
Not Head: Use Head mode to get only the first n rows. This option is useful if you want to test a pipeline on a small number of rows, and don ' t need the data to be balanced or sampled in any way.
Not Sampling: The Sampling option supports simple random sampling or stratified random sampling. This is useful if you want to create a smaller representative sample dataset for testing.
Box 2: Partition evenly
Specify the partitioner method: Indicate how you want data to be apportioned to each partition, using these options:
Partition evenly: Use this option to place an equal number of rows in each partition. To specify the number of output partitions, type a whole number in the Specify number of folds to split evenly into text box.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/algorithm-module-reference/partition-and-sample

NEW QUESTION: 189
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 (LEAVE A REPLY)
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

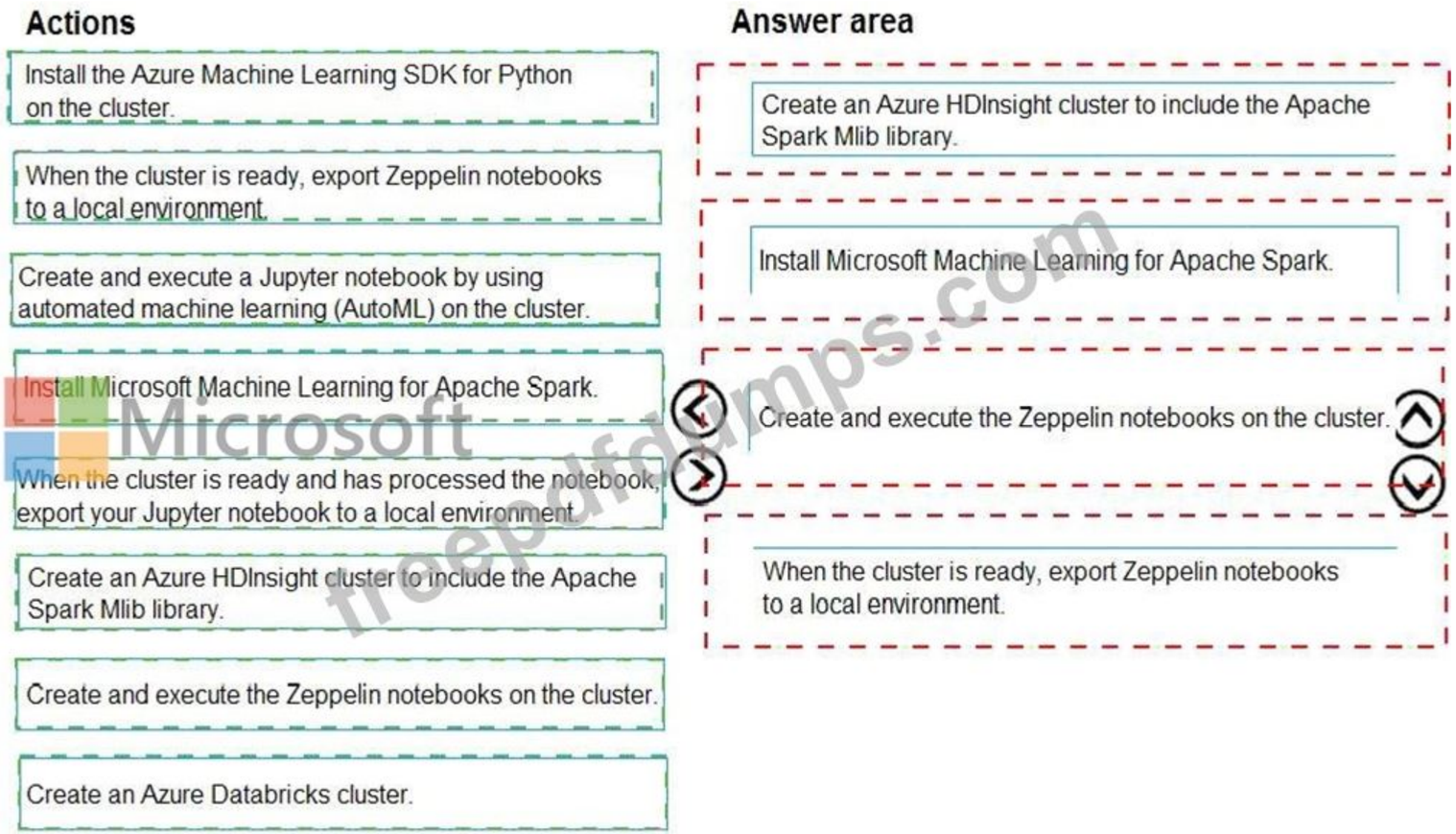
NEW QUESTION: 190
You plan to run a Python script as an Azure Machine Learning experiment.
The script must read files from a hierarchy of folders. The files will be passed to the script as a dataset argument.
You must specify an appropriate mode for the dataset argument.
Which two modes can you use? Each correct answer presents a complete solution.
NOTE: Each correct selection is worth one point.
A. to_pandas_dataframe ()
B. as_download()
C. as_upload()
D. as mount ()
Answer: B (LEAVE A REPLY)
Reference:
https://docs.microsoft.com/en-us/python/api/azureml-core/azureml.data.filedataset?view=azure-ml-py

NEW QUESTION: 191

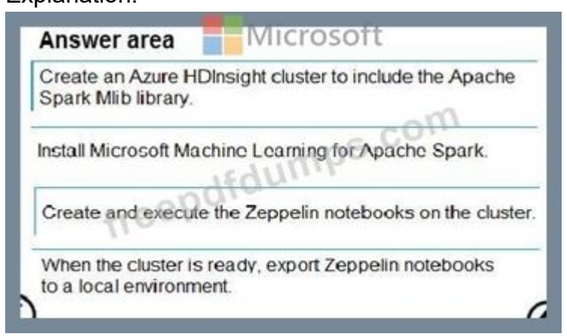
You are building an intelligent solution using machine learning models.
The environment must support the following requirements:
Data scientists must build notebooks in a cloud environment
Data scientists must use automatic feature engineering and model building in machine learning pipelines.
Notebooks must be deployed to retrain using Spark instances with dynamic worker allocation.
Notebooks must be exportable to be version controlled locally.
You need to create the environment.
Which four actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.



Answer:



Explanation:



Step 1: Create an Azure HDInsight cluster to include the Apache Spark Mlib library Step 2: Install Microsot Machine Learning for Apache Spark You install AzureML on your Azure HDInsight cluster.

Microsoft Machine Learning for Apache Spark (MMLSpark) provides a number of deep learning and data science tools for Apache Spark, including seamless integration of Spark Machine Learning pipelines with Microsoft Cognitive Toolkit (CNTK) and OpenCV, enabling you to quickly create powerful, highly-scalable predictive and analytical models for large image and text datasets.

Step 3: Create and execute the Zeppelin notebooks on the cluster

Step 4: When the cluster is ready, export Zeppelin notebooks to a local environment.

Notebooks must be exportable to be version controlled locally.

References:

https://docs.microsoft.com/en-us/azure/hdinsight/spark/apache-spark-zeppelin-notebook
https://azuremlbuild.blob.core.windows.net/pysparkapi/intro.html

NEW QUESTION: 192
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 version with 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 model. Do not register the new model.
D. Delete the existing model and register the new one with the same name.
Answer: B (LEAVE A REPLY)
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

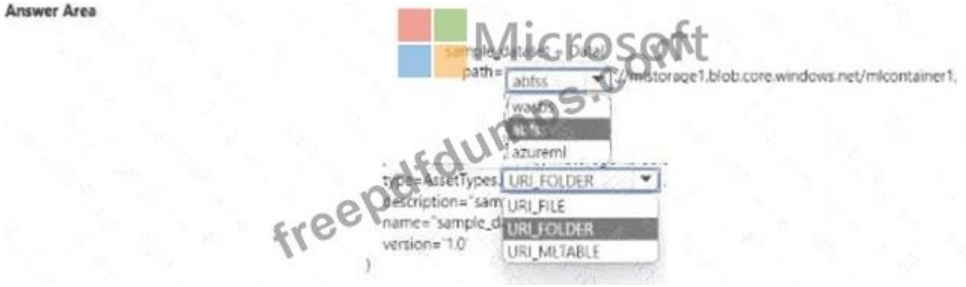
NEW QUESTION: 193
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.
After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.
An IT department creates the following Azure resource groups and resources:

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

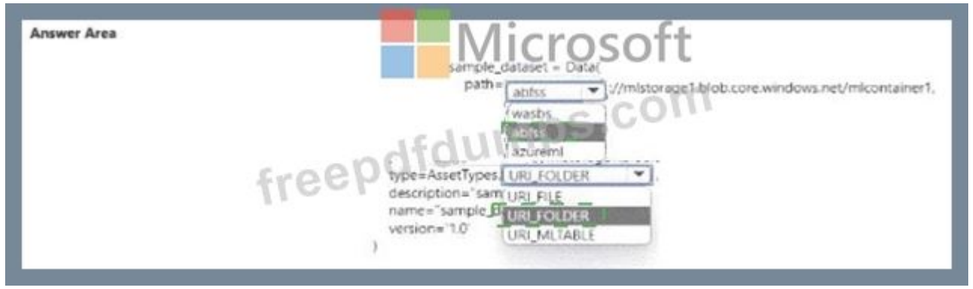
The IT department creates an Azure Kubernetes Service (AKS)-based inference compute target named aks- cluster in the Azure Machine Learning workspace.
You have a Microsoft Surface Book computer with a GPU. Python 3.6 and Visual Studio Code are installed.
You need to run a script that trains a deep neural network (DNN) model and logs the loss and accuracy metrics.
Solution: Attach the mlvm virtual machine as a compute target in the Azure Machine Learning workspace.
Install the Azure ML SDK on the Surface Book and run Python code to connect to the workspace. Run the training script as an experiment on the mlvm remote compute resource.
A. Yes
B. No
Answer: A (LEAVE A REPLY)
Use the VM as a compute target.
Note: A compute target is a designated compute resource/environment where you run your training script or host your service deployment. This location may be your local machine or a cloud-based compute resource.

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

NEW QUESTION: 194
You manage an Azure Machine Learning won pace named workspace 1 by using the Python SDK v2. You create a Gene-al Purpose v2 Azure storage account named mlstorage1. The storage account includes a pulley accessible container name micOTtalnerl. The container stores 10 blobs with files in the CSV format.
You must develop Python SDK v2 code to create a data asset referencing all blobs in the container named mtcontamer1.
You need to complete the Python SDK v2 code.
How should you complete the code? To answer select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.



Answer:



Explanation:

NEW QUESTION: 195
You create a workspace by using Azure Machine Learning Studio.
You must run a Python SDK v2 notebook in the workspace by using Azure Machine Learning Studio. You must preserve the current values of variables set in the notebook for the current instance.
You need to maintain the state of the notebook.
What should you do?
A. Change the compute.
B. Change the current kernel

C. Stop the compute.
D. Stop the current kernel.
Answer: B (LEAVE A REPLY)

NEW QUESTION: 196
You have a multi-class image classification deep learning model that uses a set of labeled photographs. You create the following code to select hyperparameter values when training the model.

```
from azureml.train.hyperdrive import BayesianParameterSampling
param_sampling = BayesianParametersSampling ({
 "learning_rate": uniform(0.01, 0.1),
 "batch_size": choice(16, 32, 64, 128)}
)
```

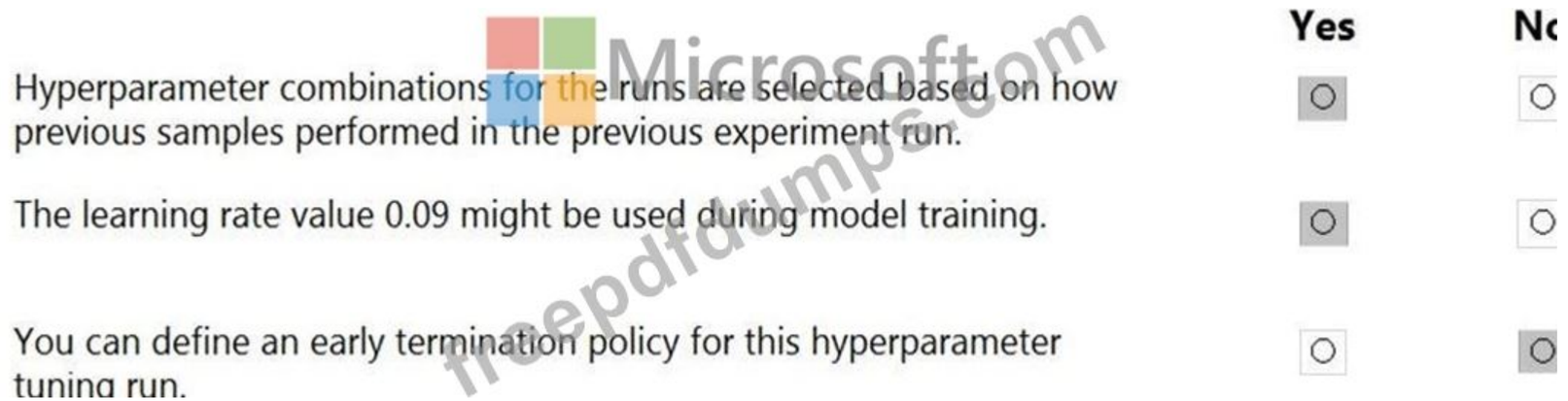
For each of the following statements, select Yes if the statement is true. Otherwise, select No.
NOTE: Each correct selection is worth one point.

	Yes	No
Hyperparameter combinations for the runs are selected based on how previous samples performed in the previous experiment run.	○	○
The learning rate value 0.09 might be used during model training.	○	○
You can define an early termination policy for this hyperparameter tuning run.	○	○

Answer:

	Yes	No
Hyperparameter combinations for the runs are selected based on how previous samples performed in the previous experiment run.	○	○
The learning rate value 0.09 might be used during model training.	○	○
You can define an early termination policy for this hyperparameter tuning run.	○	○

Explanation:



Box 1: Yes

Hyperparameters are adjustable parameters you choose to train a model that govern the training process itself.

Azure Machine Learning allows you to automate hyperparameter exploration in an efficient manner, saving you significant time and resources. You specify the range of hyperparameter values and a maximum number of training runs. The system then automatically launches multiple simultaneous runs with different parameter configurations and finds the configuration that results in the best performance, measured by the metric you choose. Poorly performing training runs are automatically early terminated, reducing wastage of compute resources. These resources are instead used to explore other hyperparameter configurations.

Box 2: Yes

uniform(low, high) - Returns a value uniformly distributed between low and high Box 3: No Bayesian sampling does not currently support any early termination policy.

Reference:

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

NEW QUESTION: 197

You manage an Azure Machine Learning workspace. You build a model for which you must configure a Responsible AI dashboard. Based on what you learn from the dashboard, you must perform the following activities:

* Determine what must be done to get a desirable outcome from the model.
* Identify the features that have the most direct effect on your outcome of interest.

You need to select the components to use for the Responsible AI dashboard configuration. Which two components should you add? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

A. error analysis
B. causal
C. explanation
D. counterfactuals
Answer: B,D (LEAVE A REPLY)

NEW QUESTION: 198

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: (SHOW ANSWER)
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: 199
You manage an Azure Machine Learning workspace that includes a batch endpoint. You plan to deploy a model to the batch endpoint. You need to configure compute for the deployment. Which compute should you use?
A. Kubernetes cluster
B. Remote VM
C. Azure Batch
D. Azure Databricks
Answer: D (LEAVE A REPLY)

NEW QUESTION: 200
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.

Microsoft Azure

Azure Databricks

Home

Workspace

Recents

Data

Clusters

Jobs

Models

Search

Create Cluster

New Cluster

Cancel | Create Cluster

2-8 Workers: 28.0-112.0 GB Memory, 8-32 Cores, 1.5-6 DBU
1 Driver: 14.0 GB Memory, 4 Cores, 0.75 DBU

Cluster Name

mysparkcluster

Cluster Mode

Standard

Pool

None

Databricks Runtime Version Learn more

Runtime: 6.4 (Scala 2.11, Spark 2.4.5)

New This Runtime version supports only Python 3.

Autopilot Options

☑ Enable autoscaling

☑ Terminate after 120 minutes of inactivity

Worker Type

Standard_DS3_v2 14.0 GB Memory, 4 Cores, 0.75 DBU

Min Workers: 2

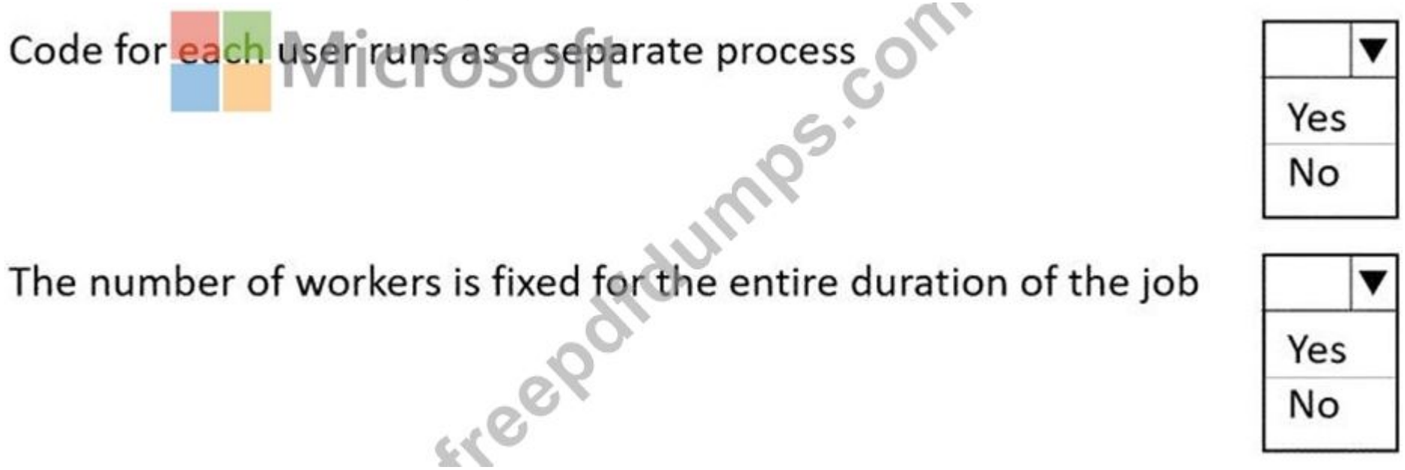
Max Workers: 8

Driver Type

Same as worker 14.0 GB Memory, 4 Cores, 0.75 DBU

▸ Advanced Options

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.

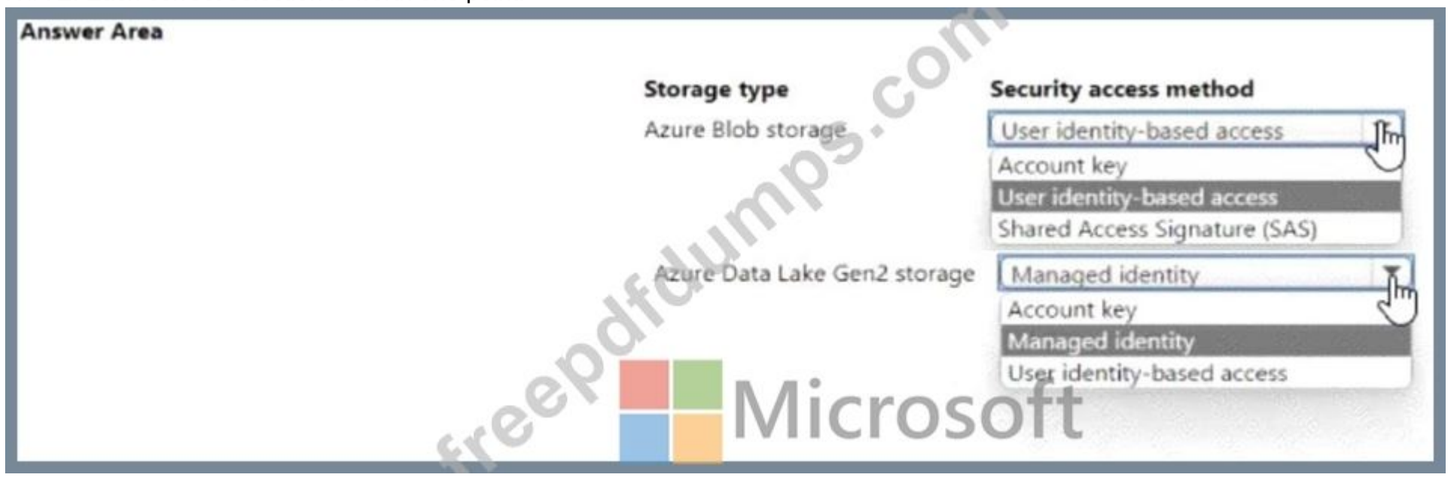


Answer:

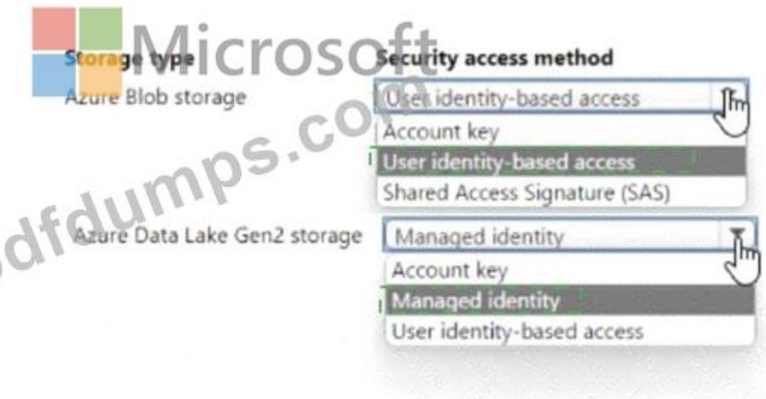
Explanation:
Box 1: No
Running user code in separate processes is not possible in Scala.
Box 2: No
Autoscaling is enabled. Minimum 2 workers, Maximum 8 workers.
Reference:
https://docs.databricks.com/clusters/configure.html

NEW QUESTION: 201
You manage an Azure Machine Learning workspace named workspace1 by using the Python SDK v2.
You must register datastores in workspace1 for Azure Blob and Azure Data Lake Gen2 storage to meet the following requirements:
* Data scientists accessing the datastore must have the same level of access.
* Access must be restricted to specified containers or folders.
You need to configure a security access method used to register the Azure Blob and Azure Data lake Gen?
storage in workspace1. Which security access method should you configure? To answer, select the appropriate options in the answers area.

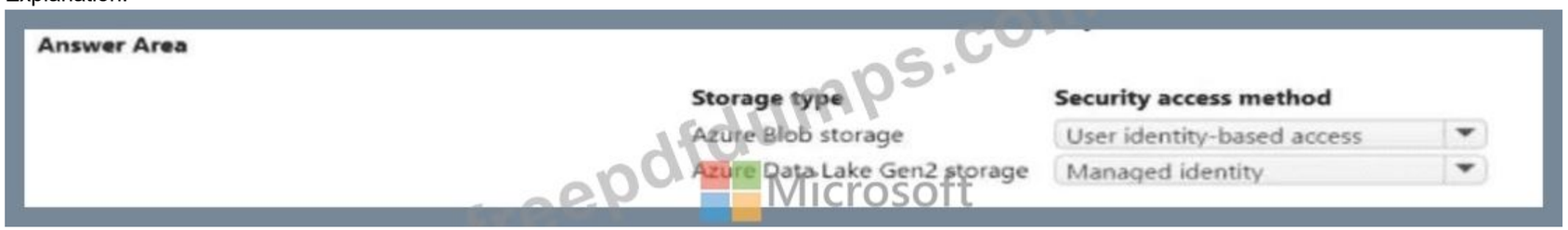
NOTE: Each correct selection is worth one point.



Answer:



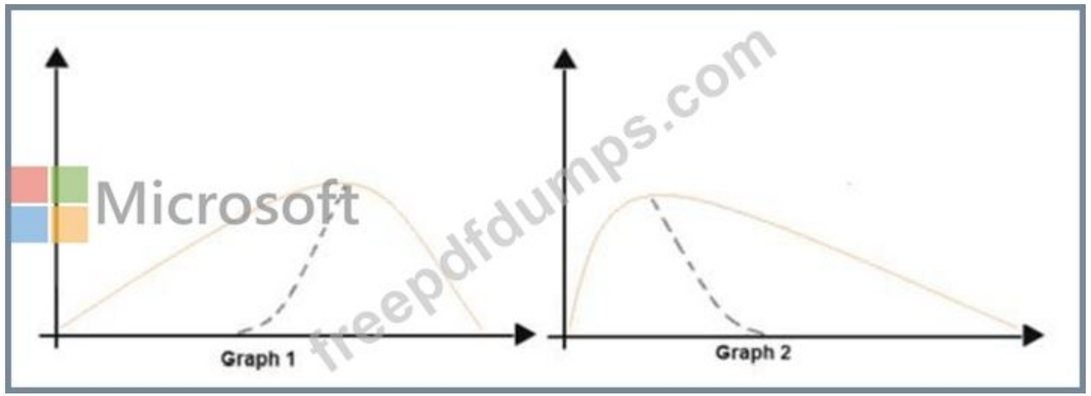
Explanation:



NEW QUESTION: 202

You are analyzing the asymmetry in a statistical distribution.

The following image contains two density curves that show the probability distribution of two datasets.



Use the drop-down menus to select the answer choice that answers each question based on the information presented in the graphic.

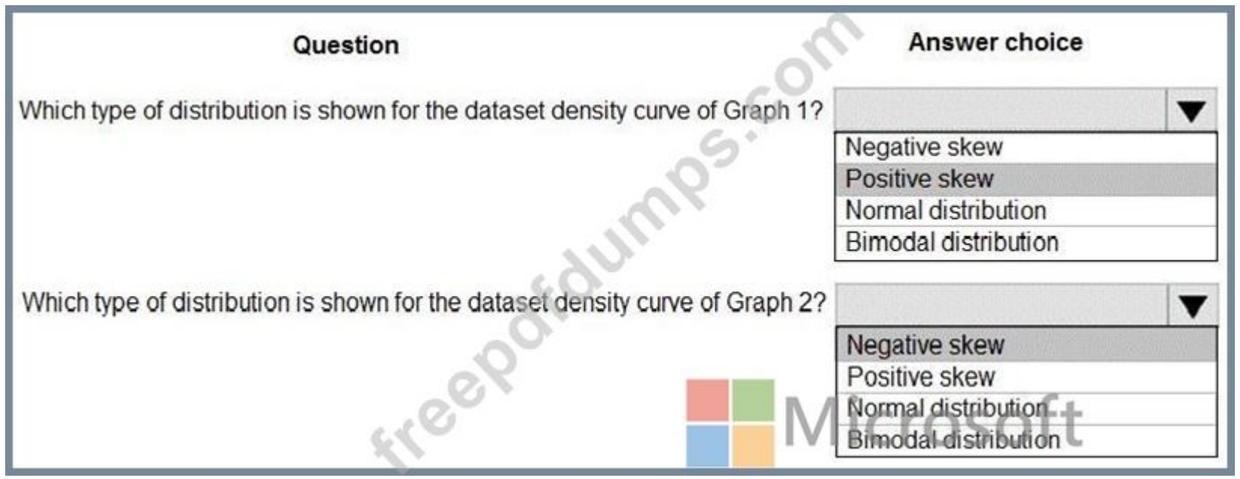
NOTE: Each correct selection is worth one point.

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

Answer:

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

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- statistics

NEW QUESTION: 203

You have an Azure Machine Learning workspace.

You have the following code:

```
command_job = command(
    code="./src",
    command="train.py",
    environment=env,
    compute="cpu-cluster",
)
```

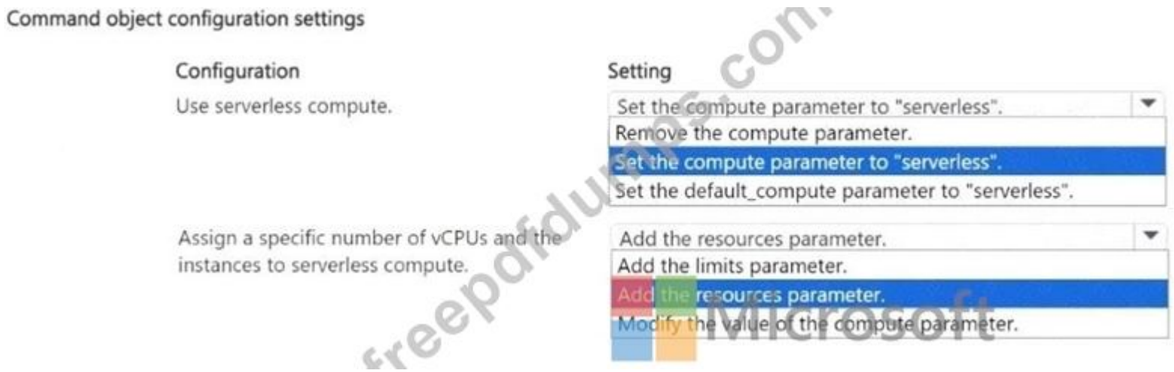
You plan to rely on serverless compute to train a model by using Azure Machine Learning Python SDK v2.

The serverless compute must use a designated number of nodes of a specific virtual machine type.

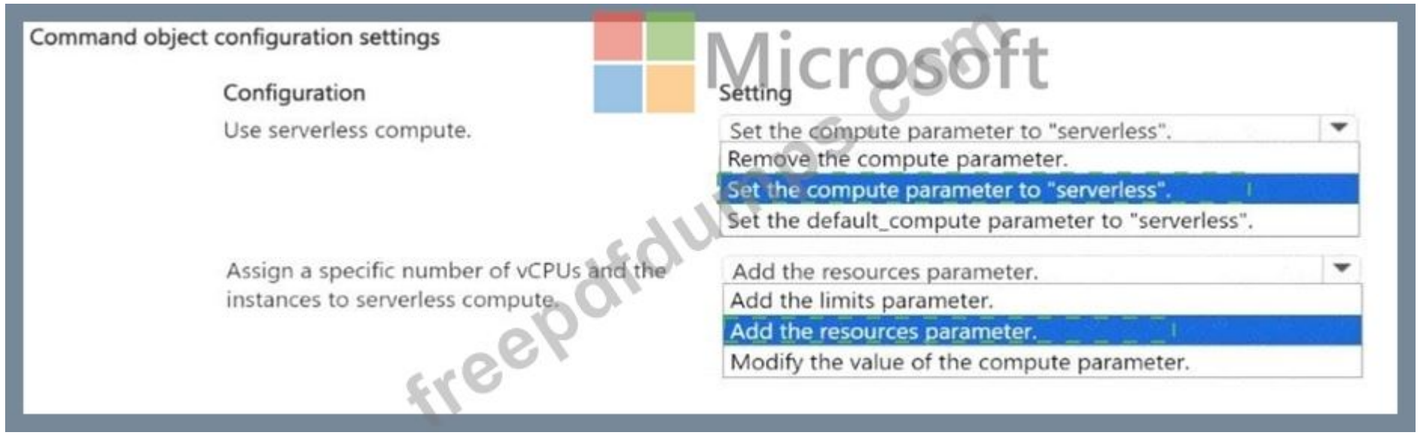
You need to modify the code to run the training job according to the plan.

How should you modify the command object? To answer, select the appropriate oations in the answer area.

NOTE: Each correct selection is worth one point.



Answer:



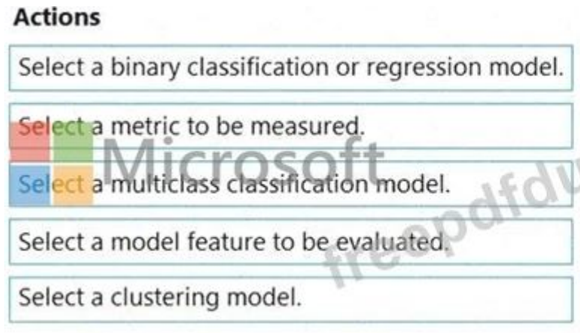
Explanation:

NEW QUESTION: 204

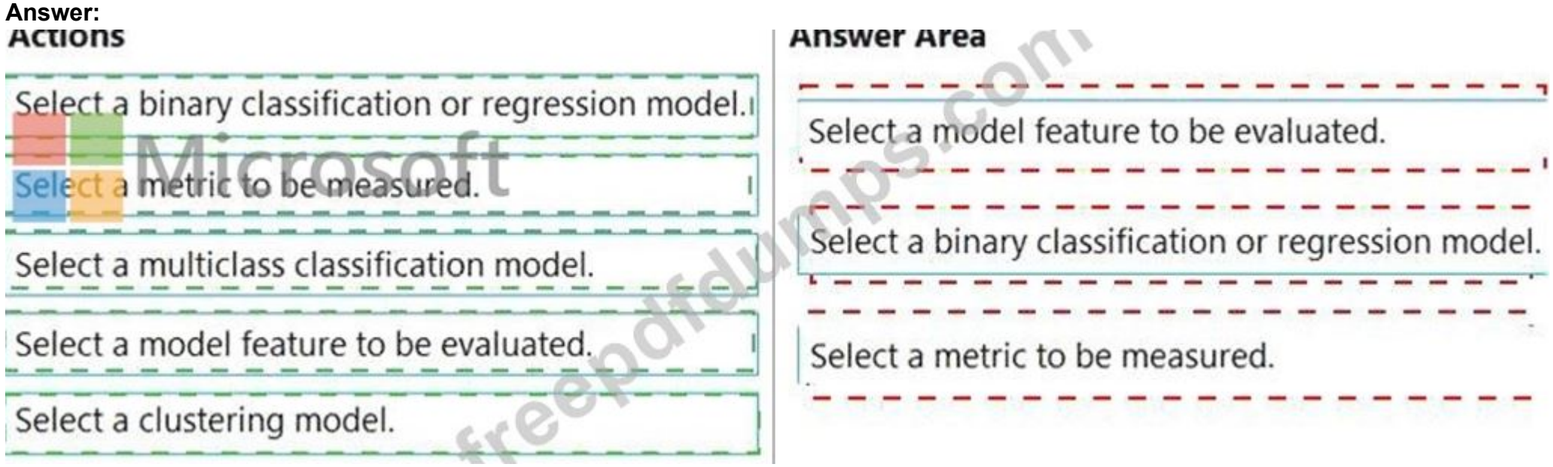
You have several machine learning models registered in an Azure Machine Learning workspace.

You must use the Fairlearn dashboard to assess fairness in a selected model.

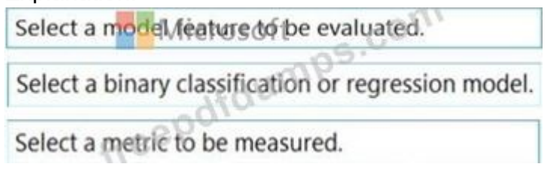
Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.



Answer:



Explanation:



Step 1: Select a model feature to be evaluated.

Step 2: Select a binary classification or regression model.

Register your models within Azure Machine Learning. For convenience, store the results in a dictionary, which maps the id of the registered model (a string in name:version format) to the predictor itself.

Example:

model_dict = {}

lr_reg_id = register_model(" fairness_logistic_regression " , lr_predictor) model_dict[lr_reg_id] = lr_predictor svm_reg_id = register_model(" fairness_svm " , svm_predictor) model_dict[svm_reg_id] = svm_predictor Step 3: Select a metric to be measured Precompute fairness metrics.

Create a dashboard dictionary using Fairlearn ' s metrics package.

Reference:

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

NEW QUESTION: 205

You are using hyperparameter tuning in Azure Machine Learning Python SDK v2 to train a model. You configure the hyperparameter tuning experiment by running the following code:

```
from azure.ai.ml.sweep import Normal, Uniform

command_job_for_sweep = command_job(
    learning_rate=Normal(10, 3),
    keep_probability=Uniform(0.05, 0.1),
    batch_size=Choice(values=[16, 32, 64, 128]),
    number_of_hidden_layers=Choice(range(3,5))
)
```

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

Answer Area

Statements	Yes	No
By defining sampling in this manner, every possible combination of the parameters will be tested.	○	○
Random values of the learning rate parameter will be selected from a normal distribution with a mean of 10 and a standard deviation of 3.	○	○
The keep_probability parameter value will always be either **0.05** or **0.1**.	○	○
Random values for the number_of_hidden_layers parameter will be selected from a normal distribution with a mean of 3 and a standard deviation of 5.	○	○

Answer:

Statements	Yes	No
By defining sampling in this manner, every possible combination of the parameters will be tested.	○	●
Random values of the learning rate parameter will be selected from a normal distribution with a mean of 10 and a standard deviation of 3.	●	○
The keep_probability parameter value will always be either **0.05** or **0.1**.	○	●
Random values for the number_of_hidden_layers parameter will be selected from a normal distribution with a mean of 3 and a standard deviation of 5.	○	●

NEW QUESTION: 206

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 (LEAVE A REPLY)

References:

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

NEW QUESTION: 207
You create an Azure Machine Learning workspace named ML-workspace. You also create an Azure Databricks workspace named DB-workspace. DB-workspace contains a cluster named DB-cluster.
You must use DB-cluster to run experiments from notebooks that you import into DB-workspace.
You need to use ML-workspace to track MLflow metrics and artifacts generated by experiments running on DB-cluster. The solution must minimize the need for custom code.
What should you do?
A. From DB-cluster, configure the Advanced Logging option.
B. From DB-workspace. configure the Link Azure ML workspace option.
C. From ML-workspace. create an attached compute.
D. From ML-workspace. create a compute cluster.
Answer: (SHOW ANSWER)
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-use-mlflow-azure-databricks

NEW QUESTION: 208
You manage an Azure AI Foundry project. You plan to create a vector index for a RAG solution. You need to build the index remotely by using a script.
Which two data sources can you use? Each correct answer presents a complete solution. Choose two. NOTE:
Each correct selection is worth one point.
A. GitHub repositories
B. Local files and foldets
C. OneLake in Microsoft Fabric
D. Azure AI Search
Answer: B,D (LEAVE A REPLY)

NEW QUESTION: 209
You plan to use a Deep Learning Virtual Machine (DLVM) to train deep learning models using Compute Unified Device Architecture (CUDA) computations.
You need to configure the DLVM to support CUDA.
What should you implement?
A. Intel Software Guard Extensions (Intel SGX) technology
B. Solid State Drives (SSD)
C. Graphic Processing Unit (GPU)
D. Computer Processing Unit (CPU) speed increase by using overcloking
E. High Random Access Memory (RAM) configuration
Answer: C (LEAVE A REPLY)
A Deep Learning Virtual Machine is a pre-configured environment for deep learning using GPU instances.
References:
https://azuremarketplace.microsoft.com/en-au/marketplace/apps/microsoft-ads.dsvm-deep-learning

NEW QUESTION: 210
You manage an Azure AI Foundry project. You build a multi-turn chatbot application.

You plan to filter your traces to identity issues while observing how the application is responding. The solution must not use an external knowledge base. You need to select an evaluation metric. Which built-in evaluator should you use?
A. GroundednessEvaluator
B. SeHHarmEvaluator
C. FlScoreEvaluator
D. IndirectAttackEvaluator
Answer: A (LEAVE A REPLY)

NEW QUESTION: 211
You manage an Azure Machine Learning workspace.
You experiment with an MLflow model that trains interactively by using a notebook in the workspace. You need to log dictionary type artifacts of the experiments in Azure Machine Learning by using MLflow. Which syntax should you use?
A. mlflow.log_metric("my_metric", my_dict)
B. mlflow.log_artifacts(my_dict>
C. mlflow.log_artifact(my_dict)
D. mlflow.log metrics(my diet)
Answer: A (LEAVE A REPLY)

NEW QUESTION: 212
You manage an Azure Machine Learning workspace.
You must log multiple metrics by using MLflow.
You need to maximize logging performance.
What are two possible ways to achieve this goal? Each correct answer presents a complete solution.
NOTE: Each correct selection is worth one point.
A. MLflowClient.log_batch
B. mlflowlog_metrics
C. mlflow.log. metric
D. mlflow.log_param
Answer: (SHOW ANSWER)