

DP-100 Dumps

Designing and Implementing a Data Science Solution on Azure

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



NEW QUESTION 1

- (Exam Topic 3)

You have an Azure Machine Learning workspace that contains a training cluster and an inference cluster. You plan to create a classification model by using the Azure Machine Learning designer.

You need to ensure that client applications can submit data as HTTP requests and receive predictions as responses.

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.

The screenshot shows the 'Actions' list on the left and the 'Answer area' on the right. The 'Actions' list contains five items:

- Create a real-time inference pipeline and run the pipeline on the compute cluster.
- Create a batch inference pipeline and run the pipeline on the compute cluster.
- Deploy a service to the compute cluster.
- Create a pipeline that trains a classification model and run the pipeline on the compute cluster.
- Deploy a service to the inference cluster.

The 'Answer area' is currently empty, with up and down arrows for reordering.

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

The screenshot shows the 'Actions' list on the left and the 'Answer area' on the right. The 'Answer area' contains three actions in the correct sequence, indicated by red dashed boxes and arrows:

- Create a pipeline that trains a classification model and run the pipeline on the compute cluster.
- Create a batch inference pipeline and run the pipeline on the compute cluster.
- Deploy a service to the inference cluster.

The 'Actions' list on the left remains the same as in the previous screenshot.

NEW QUESTION 2

- (Exam Topic 3)

HOTSPOT

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.	<input type="radio"/>	<input type="radio"/>
The estimator will mount the local data-folder folder and make it available to the script through a parameter.	<input type="radio"/>	<input type="radio"/>
The train.py script file will be created if it does not exist.	<input type="radio"/>	<input type="radio"/>
The estimator can run Scikit-learn experiments.	<input type="radio"/>	<input type="radio"/>

- A. Mastered
B. Not Mastered

Answer: A

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 3

- (Exam Topic 3)

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

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

You create a model to forecast weather conditions based on historical data.

You need to create a pipeline that runs a processing script to load data from a datastore and pass the processed data to a machine learning model training script.

Solution: Run the following code:

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

Does the solution meet the goal?

- A. Yes
B. No

Answer: B

Explanation:

train_step is missing. Reference:

<https://docs.microsoft.com/en-us/python/api/azureml-pipeline-core/azureml.pipeline.core.pipelinedata?view=azu>

NEW QUESTION 4

- (Exam Topic 3)

HOTSPOT

You register the following versions of a model.

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

You use the Azure ML Python SDK to run a training experiment. You use a variable named run to reference the experiment run.

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

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

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

NOTE: Each correct selection is worth one point.

	Yes	No
The code will cause a previous version of the saved model to be overwritten.	<input type="radio"/>	<input type="radio"/>
The version number will now be 4.	<input type="radio"/>	<input type="radio"/>
The latest version of the stored model will have a property of value: 87.43.	<input type="radio"/>	<input type="radio"/>

- A. Mastered
B. Not Mastered

Answer: A

Explanation:

Reference:

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

NEW QUESTION 5

- (Exam Topic 3)

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

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

An IT department creates the following Azure resource groups and resources:

Resource group	Resources
ml_resources	<ul style="list-style-type: none"> 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: <ul style="list-style-type: none"> 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. Run the training script as an experiment on the aks-cluster compute target.

Does the solution meet the goal?

- A. Yes
B. No

Answer: B

Explanation:

Need to attach the mlvm virtual machine as a compute target in the Azure Machine Learning workspace.

Reference:

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

NEW QUESTION 6

- (Exam Topic 3)

You train a machine learning model.

You must deploy the model as a real-time inference service for testing. The service requires low CPU utilization and less than 48 MB of RAM. The compute target for the deployed service must initialize automatically while minimizing cost and administrative overhead.

Which compute target should you use?

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

Answer: C

<https://docs.microsoft.com/id-id/azure/machine-learning/how-to-deploy-and-where>

- A. Mastered
B. Not Mastered

Answer: A

Explanation:

Graphical user interface, text, application Description automatically generated

Box 1: No

Environment is a required parameter. The environment to use for the run. If no environment is specified, azureml.core.runconfig.DEFAULT_CPU_IMAGE will be used as the Docker image for the run.

The following example shows how to instantiate a new environment. from azureml.core import Environment

myenv = Environment(name="myenv") Box 2: Yes

Parameter compute_target: The compute target where training will happen. This can either be a ComputeTarget object, the name of an existing ComputeTarget, or the string "local". If no compute target is specified, your local machine will be used.

Box 3: Yes

Parameter source_directory. A local directory containing code files needed for a run. Parameter script. The file path relative to the source_directory of the script to be run. Reference:

<https://docs.microsoft.com/en-us/python/api/azureml-core/azureml.core.scriptrunconfig> <https://docs.microsoft.com/en-us/python/api/azureml-core/azureml.core.environment.environment>

NEW QUESTION 9

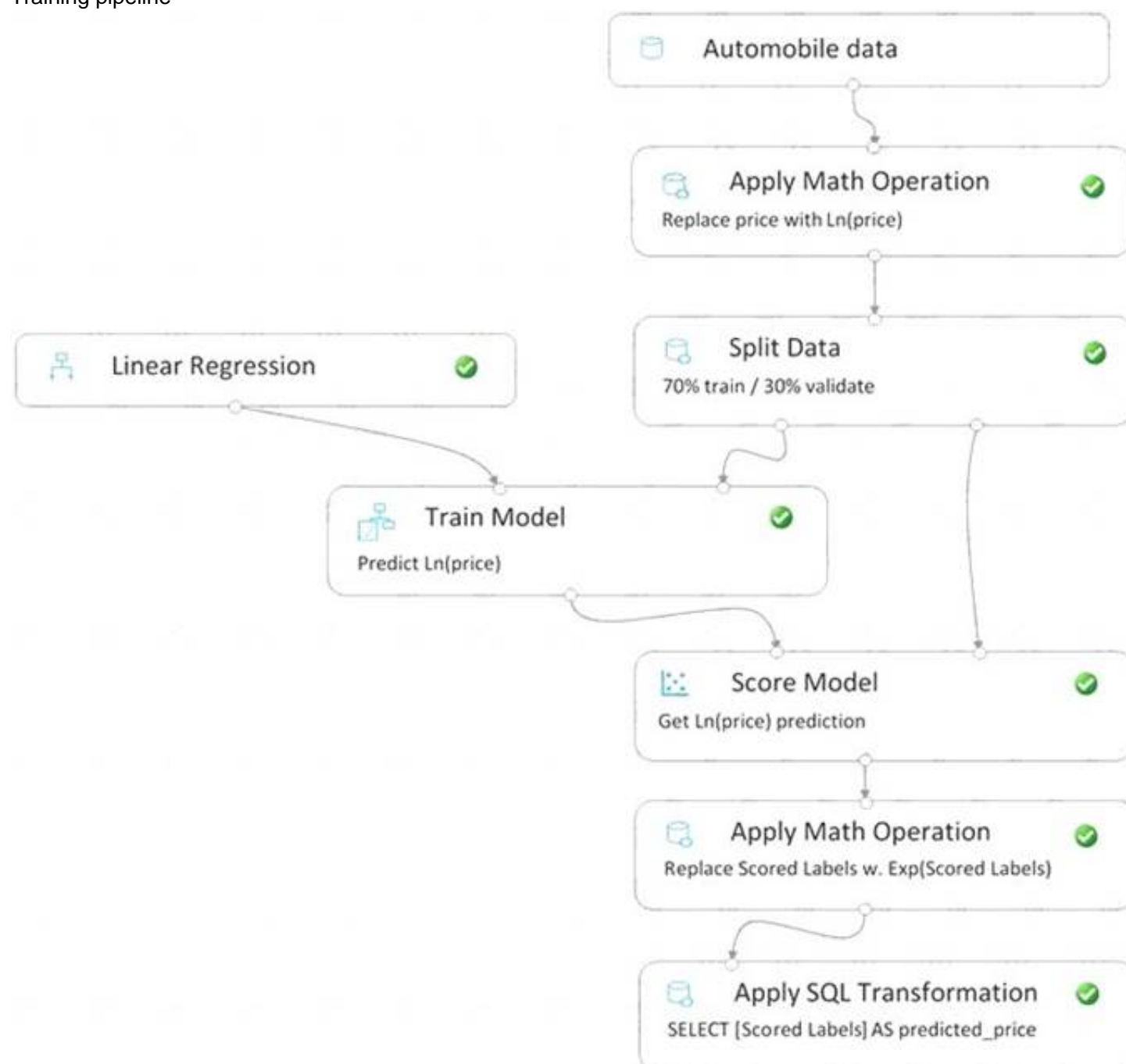
- (Exam Topic 3)

You create a pipeline in designer to train a model that predicts automobile prices.

Because of non-linear relationships in the data, the pipeline calculates the natural log (Ln) of the prices in the training data, trains a model to predict this natural log of price value, and then calculates the exponential of the scored label to get the predicted price.

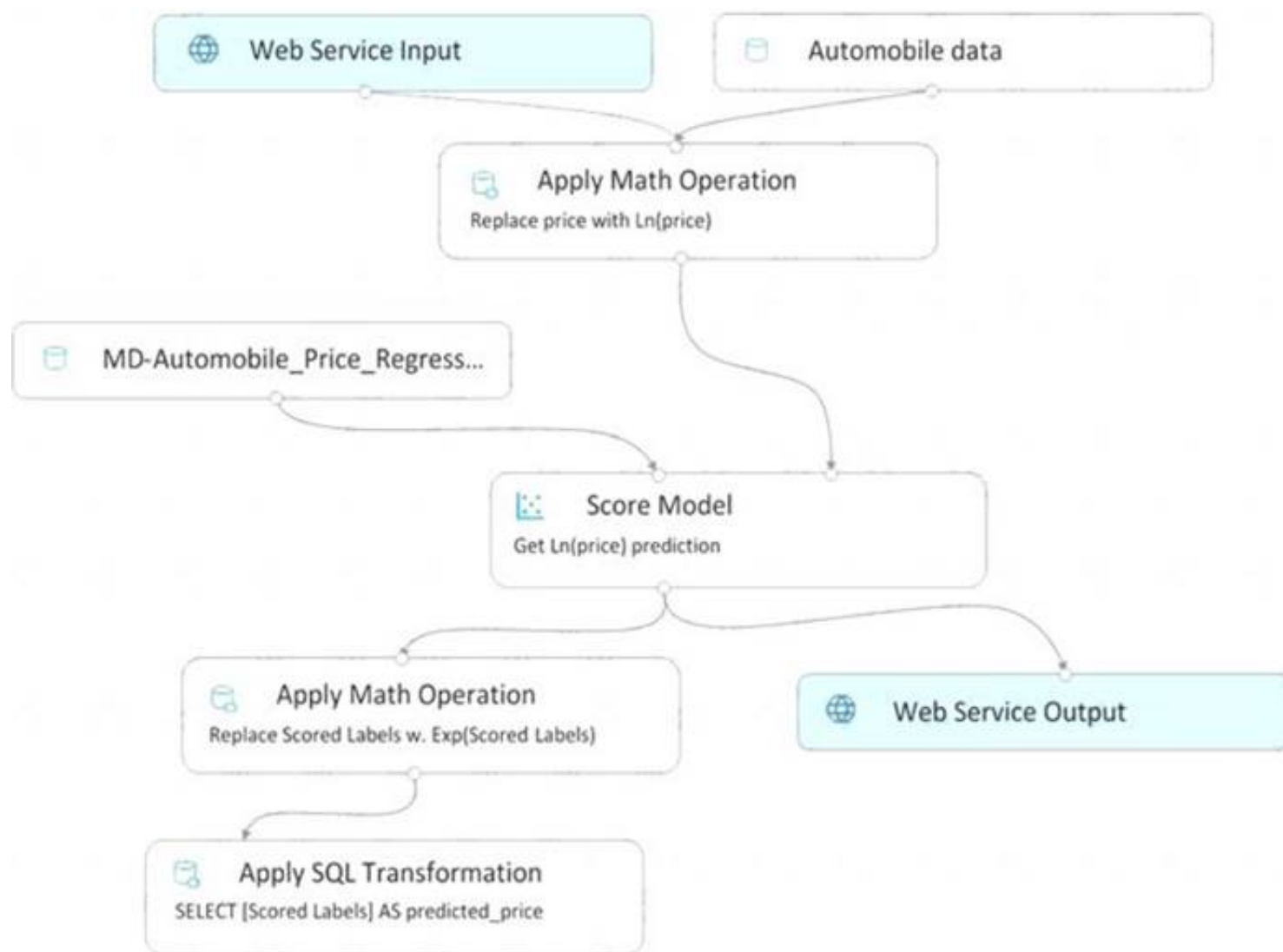
The training pipeline is shown in the exhibit. (Click the Training pipeline tab.)

Training pipeline



You create a real-time inference pipeline from the training pipeline, as shown in the exhibit. (Click the Real-time pipeline tab.)

Real-time pipeline



You need to modify the inference pipeline to ensure that the web service returns the exponential of the scored label as the predicted automobile price and that client applications are not required to include a price value in the input values.

Which three modifications must you make to the inference pipeline? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

- A. Connect the output of the Apply SQL Transformation to the Web Service Output module.
- B. Replace the Web Service Input module with a data input that does not include the price column.
- C. Add a Select Columns module before the Score Model module to select all columns other than price.
- D. Replace the training dataset module with a data input that does not include the price column.
- E. Remove the Apply Math Operation module that replaces price with its natural log from the data flow.
- F. Remove the Apply SQL Transformation module from the data flow.

Answer: ACE

NEW QUESTION 10

- (Exam Topic 3)

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

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

You 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: Calculate the column median value and use the median value as the replacement for any missing value in the column.

Does the solution meet the goal?

- A. Yes
- B. No

Answer: B

Explanation:

Use the Multiple Imputation by Chained Equations (MICE) method. References: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3074241/>
<https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/clean-missing-data>

NEW QUESTION 10

- (Exam Topic 3)

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

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

You are creating a new experiment in Azure Machine Learning Studio.

One class has a much smaller number of observations than the other classes in the training set. You need to select an appropriate data sampling strategy to compensate for the class imbalance. Solution: You use the Stratified split for the sampling mode.

Does the solution meet the goal?

- A. Yes
- B. No

Answer: B

Explanation:

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

Note: SMOTE is used to increase the number of underrepresented cases in a dataset used for machine learning. SMOTE is a better way of increasing the number of rare cases than simply duplicating existing cases.

References:

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

NEW QUESTION 15

- (Exam Topic 3)


You are developing a deep learning model by using TensorFlow. You plan to run the model training workload on an Azure Machine Learning Compute Instance.

You must use CUDA-based model training. You need to provision the Compute Instance.

Which two virtual machines sizes can you use? To answer, select the appropriate virtual machine sizes in the answer area.

NOTE: Each correct selection is worth one point.

Virtual machine size

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

A. Mastered

B. Not Mastered

Answer: A

Explanation:

CUDA is a parallel computing platform and programming model developed by Nvidia for general computing on its own GPUs (graphics processing units). CUDA enables developers to speed up compute-intensive applications by harnessing the power of GPUs for the parallelizable part of the computation.

Reference:

<https://www.infoworld.com/article/3299703/what-is-cuda-parallel-programming-for-gpus.html>

NEW QUESTION 18

- (Exam Topic 3)

You train a model and register it in your Azure Machine Learning workspace. You are ready to deploy the model as a real-time web service.

You deploy the model to an Azure Kubernetes Service (AKS) inference cluster, but the deployment fails because an error occurs when the service runs the entry script that is associated with the model deployment.

You need to debug the error by iteratively modifying the code and reloading the service, without requiring a re-deployment of the service for each code update.

What should you do?

A. Register a new version of the model and update the entry script to load the new version of the model from its registered path.

B. Modify the AKS service deployment configuration to enable application insights and re-deploy to AKS.

C. Create an Azure Container Instances (ACI) web service deployment configuration and deploy the model on ACI.

D. Add a breakpoint to the first line of the entry script and redeploy the service to AKS.

E. Create a local web service deployment configuration and deploy the model to a local Docker container.

Answer: C

Explanation:

How to work around or solve common Docker deployment errors with Azure Container Instances (ACI) and Azure Kubernetes Service (AKS) using Azure Machine Learning.

The recommended and the most up to date approach for model deployment is via the Model.deploy() API using an Environment object as an input parameter. In this case our service will create a base docker image for you during deployment stage and mount the required models all in one call. The basic deployment tasks are:

* 1. Register the model in the workspace model registry.

* 2. Define Inference Configuration:

* a. Create an Environment object based on the dependencies you specify in the environment yaml file or use one of our procured environments.

* b. Create an inference configuration (InferenceConfig object) based on the environment and the scoring script.

* 3. Deploy the model to Azure Container Instance (ACI) service or to Azure Kubernetes Service (AKS).

NEW QUESTION 19

- (Exam Topic 3)

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

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

You are 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: Relative Squared Error, Coefficient of Determination, Accuracy, Precision, Recall, F1 score, and AUC.

Does the solution meet the goal?

- A. Yes
- B. No

Answer: B

Explanation:

Relative Squared Error, Coefficient of Determination are good metrics to evaluate the linear regression model, but the others are metrics for classification models.

References:

<https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/evaluate-model>

NEW QUESTION 22

- (Exam Topic 3)

You use the designer to create a training pipeline for a classification model. The pipeline uses a dataset that includes the features and labels required for model training.

You create a real-time inference pipeline from the training pipeline. You observe that the schema for the generated web service input is based on the dataset and includes the label column that the model predicts. Client applications that use the service must not be required to submit this value.

You need to modify the inference pipeline to meet the requirement. What should you do?

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

Answer: A

Explanation:

By default, the Web Service Input will expect the same data schema as the module output data which connects to the same downstream port as it. You can remove the target variable column in the inference pipeline using Select Columns in Dataset module. Make sure that the output of Select Columns in Dataset removing target variable column is connected to the same port as the output of the Web Service Input module.

Reference:

<https://docs.microsoft.com/en-us/azure/machine-learning/tutorial-designer-automobile-price-deploy>

NEW QUESTION 24

- (Exam Topic 3)

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

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

You have a Python script named train.py in a local folder named scripts. The script trains a regression model by using scikit-learn. The script includes code to load a training data file which is also located in the scripts folder.

You must run the script as an Azure ML experiment on a compute cluster named aml-compute.

You need to configure the run to ensure that the environment includes the required packages for model training. You have instantiated a variable named aml-compute that references the target compute cluster.

Solution: Run the following code:

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

Does the solution meet the goal?

- A. Yes
- B. No

Answer: B

Explanation:

The scikit-learn estimator provides a simple way of launching a scikit-learn training job on a compute target. It is implemented through the SKLearn class, which can be used to support single-node CPU training.

Example:

```
from azureml.train.sklearn import SKLearn
}
estimator = SKLearn(source_directory=project_folder, compute_target=compute_target, entry_script='train_iris.py'
)
```

Reference:

<https://docs.microsoft.com/en-us/azure/machine-learning/how-to-train-scikit-learn>

NEW QUESTION 27

- (Exam Topic 3)

You are developing a machine learning, experiment by using Azure. The following images show the input and output of a machine learning experiment:

rows	columns		
4	3		
view as	Server ID	Risk Level	Cost
	N102696	High	4500
	N102874	Low	5000
	N107027	Medium	4000
	N106548	High	4800

Input

rows	columns				
4	5				
view as	Server ID	Risk Level-High	Risk Level-Low	Risk Level-Medium	Cost
	N102696	1	0	0	4500
	N102874	0	1	0	5000
	N107027	0	0	1	4000
	N106548	1	0	0	4800

Output

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.

You need to perform the data transformation applied to the Risk Level column. Which module should you use?

What is the expected input column type for this transformation?

Apply Filter
Build Counting Transform
Convert to Indicator Values

Categorical
Numerical
String

- A. Mastered
B. Not Mastered

Answer: A

Explanation:

You need to perform the data transformation applied to the Risk Level column. Which module should you use?

What is the expected input column type for this transformation?

Apply Filter
Build Counting Transform
Convert to Indicator Values

Categorical
Numerical
String

NEW QUESTION 29

- (Exam Topic 3)

You retrain an existing model.

You need to register the new version of a model while keeping the current version of the model in the registry.

What should you do?

- A. Register a model with a different name from the existing model and a custom property named 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.
D. Do not register the new model.
E. Delete the existing model and register the new one with the same name.

Answer: B

Explanation:

Model version: A version of a registered model. When a new model is added to the Model Registry, it is added as Version 1. Each model registered to the same model name increments the version number.

Reference:

<https://docs.microsoft.com/en-us/azure/databricks/applications/mlflow/model-registry>

NEW QUESTION 32

- (Exam Topic 3)

You create a new Azure subscription. No resources are provisioned in the subscription. You need to create an Azure Machine Learning workspace.

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

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

Answer: BCD

Explanation:

B: You can use an Azure Resource Manager template to create a workspace for Azure Machine Learning. Example:

```
{"type": "Microsoft.MachineLearningServices/workspaces",
```

...

C: You can create a workspace for Azure Machine Learning with Azure CLI Install the machine learning extension.

Create a resource group: `az group create --name <resource-group-name> --location <location>`

To create a new workspace where the services are automatically created, use the following command: `az ml workspace create -w <workspace-name> -g <resource-group-name>`

D: You can create and manage Azure Machine Learning workspaces in the Azure portal.

- Sign in to the Azure portal by using the credentials for your Azure subscription.
- In the upper-left corner of Azure portal, select + Create a resource.
- Use the search bar to find Machine Learning.
- Select Machine Learning.
- In the Machine Learning pane, select Create to begin.

[Home](#) > [New](#) > [Machine Learning](#) >

Machine Learning

Create a machine learning workspace

[Basics](#) [Networking](#) [Advanced](#) [Tags](#) [Review + create](#)

Project details

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

Subscription * ⓘ

Resource group * ⓘ [Create new](#)

Workspace details

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

Workspace name * ⓘ

Region * ⓘ

Workspace edition * ⓘ

- Basic
- Enterprise

 For your convenience, these resources are available in this workspace: [Application Insights](#), [Azure Key Vault](#)

Reference:

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

NEW QUESTION 33

- (Exam Topic 3)

You are running a training experiment on remote compute in Azure Machine Learning.

The experiment is configured to use a conda environment that includes the mlflow and azureml-contrib-run packages.

You must use MLflow as the logging package for tracking metrics generated in the experiment. You need to complete the script for the experiment.

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

NOTE: Each correct selection is worth one point.


```
import numpy as np
# Import library to log metrics

from azureml.core import Run
import mlflow
import logging

# Start logging for this run

run = Run.get_context()
mlflow.start_run()
logger = logging.getLogger('Run')
reg_rate = 0.01
# Log the reg_rate metric

run.log('reg_rate', np.float(reg_rate))
mlflow.log_metric('reg_rate', np.float(reg_rate))
logger.info(np.float(reg_rate))

# Stop logging for this run

run.complete()
mlflow.end_run()
logger.setLevel(logging.INFO)
```

- A. Mastered
B. Not Mastered

Answer: A

Explanation:

Box 1: import mlflow

Import the mlflow and Workspace classes to access MLflow's tracking URI and configure your workspace. Box 2: mlflow.start_run()

Set the MLflow experiment name with set_experiment() and start your training run with start_run(). Box 3: mlflow.log_metric('..')

Use log_metric() to activate the MLflow logging API and begin logging your training run metrics. Box 4: mlflow.end_run()

Close the run: run.endRun() Reference:

<https://docs.microsoft.com/en-us/azure/machine-learning/how-to-use-mlflow>

NEW QUESTION 35

- (Exam Topic 3)

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

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

Answer: C

Explanation:

A Deep Learning Virtual Machine is a pre-configured environment for deep learning using GPU instances.

References:

<https://azuremarketplace.microsoft.com/en-au/marketplace/apps/microsoft-ads.dsvm-deep-learning>

NEW QUESTION 36

- (Exam Topic 3)

You are hired as a data scientist at a winery. The previous data scientist used Azure Machine Learning. You need to review the models and explain how each model makes decisions.

Which explainer modules should you use? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Model type	Explainer
A random forest model for predicting the alcohol content in wine given a set of covariates	<div>▼</div> <div>Tabular HAN Text Image</div>
A natural language processing model for analyzing field reports	<div>▼</div> <div>Tree HAN Text Image</div>
An image classifier that determines the quality of the grape based upon its physical characteristics.	<div>▼</div> <div>Kernel HAN Text Image</div>

- A. Mastered
B. Not Mastered

Answer: A

Explanation:

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

NEW QUESTION 37

- (Exam Topic 3)
You create an Azure Machine Learning workspace.
You must create a custom role named DataScientist that meets the following requirements:

- Role members must not be able to delete the workspace.
- Role members must not be able to create, update, or delete compute resource in the workspace.
- Role members must not be able to add new users to the workspace.

You need to create a JSON file for the DataScientist role in the Azure Machine Learning workspace. The custom role must enforce the restrictions specified by the IT Operations team.
Which JSON code segment should you use?

A)

```
{
  "Name": "DataScientist",
  "IsCustom": true,
  "Description": "Project Data Scientist role",
  "Actions": ["*"],
  "NotActions": [
    "Microsoft.MachineLearningServices/workspaces/*/delete",
    "Microsoft.MachineLearningServices/workspaces/computes/*/write",
    "Microsoft.MachineLearningServices/workspaces/computes/*/delete",
    "Microsoft.Authorization/*/write"
  ],
  "AssignableScopes": [
    "/subscriptions/<id>/resourceGroups/ml-rg/providers/Microsoft.MachineLearningServices/workspaces/ml-ws"
  ]
}
```

B)

```
{
  "Name": "DataScientist",
  "IsCustom": true,
  "Description": "Project Data Scientist role",
  "Actions": ["*"],
  "NotActions": [],
  "AssignableScopes": [
    "/subscriptions/<id>/resourceGroups/ml-rg/providers/Microsoft.MachineLearningServices/workspaces/ml-ws"
  ]
}
```

C)

```
{
  "Name": "DataScientist",
  "IsCustom": true,
  "Description": "Project Data Scientist role",
  "Actions": ["Microsoft.MachineLearningServices/workspaces/*/delete",
    "Microsoft.MachineLearningServices/workspaces/computes/*/write",
    "Microsoft.MachineLearningServices/workspaces/computes/*/delete",
    "Microsoft.Authorization/*/write"
  ],
  "NotActions": [],
  "AssignableScopes": [
    "/subscriptions/<id>/resourceGroups/ml-rg/providers/Microsoft.MachineLearningServices/workspaces/ml-ws"
  ]
}
```

D)

```
{
  "Name": "DataScientist",
  "IsCustom": true,
  "Description": "Project Data Scientist role",
  "Actions": [],
  "NotActions": ["*"],
  "AssignableScopes": [
    "/subscriptions/<id>/resourceGroups/ml-rg/providers/Microsoft.MachineLearningServices/workspaces/ml-ws"
  ]
}
```

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

Answer: A**Explanation:**

The following custom role can do everything in the workspace except for the following actions:

- It can't create or update a compute resource.
- It can't delete a compute resource.
- It can't add, delete, or alter role assignments.
- It can't delete the workspace.

To create a custom role, first construct a role definition JSON file that specifies the permission and scope for the role. The following example defines a custom role named "Data Scientist Custom" scoped at a specific workspace level:

data_scientist_custom_role.json :

```
{
  "Name": "Data Scientist Custom", "IsCustom": true,
  "Description": "Can run experiment but can't create or delete compute.", "Actions": ["*"],
  "NotActions": [
    "Microsoft.MachineLearningServices/workspaces/*/delete", "Microsoft.MachineLearningServices/workspaces/write",
    "Microsoft.MachineLearningServices/workspaces/computes/*/write", "Microsoft.MachineLearningServices/workspaces/computes/*/delete",
    "Microsoft.Authorization/*/write"
  ],
  "AssignableScopes": [ "/subscriptions/<subscription_id>/resourceGroups/<resource_group_name>/providers/Microsoft.MachineLearni"
]
}
```

Reference:

<https://docs.microsoft.com/en-us/azure/machine-learning/how-to-assign-roles>

NEW QUESTION 40

- (Exam Topic 3) You are solving a classification task. The dataset is imbalanced.

You need to select an Azure Machine Learning Studio module to improve the classification accuracy. Which module should you use?

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

Answer: C**Explanation:**

Use the SMOTE module in Azure Machine Learning Studio (classic) to increase the number of underrepresented cases in a dataset used for machine learning. SMOTE is a better way of increasing the number of rare cases than simply duplicating existing cases.

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

Reference:

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

NEW QUESTION 44

- (Exam Topic 3)

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 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)

```
}
)
```

- A. Mastered
- B. Not Mastered

Answer: A

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 48

- (Exam Topic 3)

You use the Azure Machine Learning service to create a tabular dataset named training.data. You plan to use this dataset in a training script.

You create a variable that references the dataset using the following code: training_ds = workspace.datasets.get("training_data")

You define an estimator to run the script.

You need to set the correct property of the estimator to ensure that your script can access the training.data dataset

Which property should you set?

A)

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

B)

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

C)

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

D)

```
source_directory = training_ds
```

- A. Option A
- B. Option B

- C. Option C
D. Option D

Answer: A

Explanation:

Example:

Get the training dataset

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

Create an estimator that uses the remote compute hyper_estimator = SKLearn(source_directory=experiment_folder, inputs=[diabetes_ds.as_named_input('diabetes')], # Pass the dataset as an input compute_target = cpu_cluster, conda_packages=['pandas','ipykernel','matplotlib'], pip_packages=['azureml-sdk','argparse','pyarrow'], entry_script='diabetes_training.py')

Reference:

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

NEW QUESTION 49

- (Exam Topic 3)

You are evaluating a completed binary classification machine learning model. You need to use the precision as the valuation metric. Which visualization should you use?

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

Answer: A

Explanation:

References:

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

NEW QUESTION 53

- (Exam Topic 3)

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

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

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

You need to use automated machine learning to generate the best model for predicting the sales price based on the features of the house.

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

Setting	Value
Prediction task	<div> <div></div> <div> Classification Forecasting Regression Outlier </div> </div>
Target column	<div> <div></div> <div> Price Bedrooms Size HasGarage HomeType </div> </div>

- A. Mastered
B. Not Mastered

Answer: A

Explanation:

Box 1: Regression

Regression is a supervised machine learning technique used to predict numeric values. Box 2: Price

Reference:

<https://docs.microsoft.com/en-us/learn/modules/create-regression-model-azure-machine-learning-designer>

NEW QUESTION 54

- (Exam Topic 3)

You are preparing to use the Azure ML SDK to run an experiment and need to create compute. You run the following code:

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

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

NOTE: Each correct selection is worth one point.

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

- A. Mastered
B. Not Mastered

Answer: A

Explanation:

Box 1: No

If a training cluster already exists it will be used. Box 2: Yes

The wait_for_completion method waits for the current provisioning operation to finish on the cluster. Box 3: Yes

Low Priority VMs use Azure's excess capacity and are thus cheaper but risk your run being pre-empted.

Box 4: No

Need to use training_compute.delete() to deprovision and delete the AmlCompute target. Reference:

<https://notebooks.azure.com/azureml/projects/azureml-getting-started/html/how-to-use-azureml/training/train-on> <https://docs.microsoft.com/en-us/python/api/azureml-core/azureml.core.compute.computetarget>

NEW QUESTION 59

- (Exam Topic 3)

You create a multi-class image classification deep learning model that uses a set of labeled images. You create a script file named train.py that uses the PyTorch 1.3 framework to train the model.

You must run the script by using an estimator. The code must not require any additional Python libraries to be installed in the environment for the estimator. The time required for model training must be minimized.

You need to define the estimator that will be used to run the script. Which estimator type should you use?

- A. TensorFlow
B. PyTorch
C. SKLearn
D. Estimator

Answer: B

Explanation:

For PyTorch, TensorFlow and Chainer tasks, Azure Machine Learning provides respective PyTorch, TensorFlow, and Chainer estimators to simplify using these frameworks.

Reference:

<https://docs.microsoft.com/en-us/azure/machine-learning/how-to-train-ml-models>

NEW QUESTION 64

- (Exam Topic 3)

You create a batch inference pipeline by using the Azure ML SDK. You run the pipeline by using the following code:

```
from azureml.pipeline.core import Pipeline
from azureml.core.experiment import Experiment
pipeline = Pipeline(workspace=ws, steps=[parallelrun_step]) pipeline_run = Experiment(ws, 'batch_pipeline').submit(pipeline) You need to monitor the progress of
```


the pipeline execution.

What are two possible ways to achieve this goal? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

A. Run the following code in a notebook:

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

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

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

D. Run the following code in a notebook:

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

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

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

A. Option A

B. Option B

C. Option C

D. Option D

E. Option E

Answer: DE

Explanation:

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

➤ Azure Machine Learning Studio.

➤ Console output from the PipelineRun object. `from azureml.widgets import RunDetails RunDetails(pipeline_run).show()`

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

<https://docs.microsoft.com/en-us/azure/machine-learning/how-to-use-parallel-run-step#monitor-the-parallel-run>

NEW QUESTION 66

- (Exam Topic 3)

You are building a binary classification model by using a supplied training set. The training set is imbalanced between two classes.

You need to resolve the data imbalance.

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

A. Penalize the classification

B. Resample the data set using under sampling or oversampling

C. Generate synthetic samples in the minority class.

D. Use accuracy as the evaluation metric of the model.

E. Normalize the training feature set.

Answer: ABD

Explanation:

References:

<https://machinelearningmastery.com/tactics-to-combat-imbalanced-classes-in-your-machine-learning-dataset/>

NEW QUESTION 67

- (Exam Topic 3)

You are a data scientist working for a hotel booking website company. You use the Azure Machine Learning service to train a model that identifies fraudulent transactions.

You must deploy the model as an Azure Machine Learning real-time web service using the `Model.deploy` method in the Azure Machine Learning SDK. The deployed web service must return real-time predictions of fraud based on transaction data input.

You need to create the script that is specified as the `entry_script` parameter for the `InferenceConfig` class used to deploy the model.

What should the entry script do?

A. Start a node on the inference cluster where the web service is deployed.

B. Register the model with appropriate tags and properties.

C. Create a Conda environment for the web service compute and install the necessary Python packages.

D. Load the model and use it to predict labels from input data.

E. Specify the number of cores and the amount of memory required for the inference compute.

Answer: D

Explanation:

The entry script receives data submitted to a deployed web service and passes it to the model. It then takes the response returned by the model and returns that to the client. The script is specific to your model. It must understand the data that the model expects and returns.

The two things you need to accomplish in your entry script are: Loading your model (using a function called `init()`)

Running your model on input data (using a function called `run()`) Reference:

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

NEW QUESTION 69

- (Exam Topic 3)

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. ScriptRunConfig
- C. Workspace
- D. Experiment

Answer: A

Explanation:

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

NEW QUESTION 73

- (Exam Topic 3)

You create a datastore named training_data that references a blob container in an Azure Storage account. The blob container contains a folder named csv_files in which multiple comma-separated values (CSV) files are stored.

You have a script named train.py in a local folder named ./script that you plan to run as an experiment using an estimator. The script includes the following code to read data from the csv_files folder:

```
import os
import argparse
import pandas as pd

from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from azureml.core import Run

run = Run.get_context()
parser = argparse.ArgumentParser()
parser.add_argument('--data-folder', type=str, dest='data_folder', help='data reference')
args = parser.parse_args()

data_folder = args.data_folder
csv_files = os.listdir(data_folder)
training_data = pd.concat((pd.read_csv(os.path.join(data_folder, csv_file)) for csv_file in csv_files))
```

Code goes on to split the training data and train a logistic regression model

You have the following script.

```
from azureml.core import Workspace, Datastore, Experiment
from azureml.train.sklearn import SKLearn
```

```
ws = Workspace.from_config()
exp = Experiment(workspace=ws, name='csv_training')
ds = Datastore.get(ws, datastore_name='training_data')
data_ref = ds.path('csv_files')
```

Code to define estimator goes here

```
run = exp.submit(config=estimator)
run.wait_for_completion(show_output=True)
```

You need to configure the estimator for the experiment so that the script can read the data from a data reference named data_ref that references the csv_files folder in the training_data datastore.

Which code should you use to configure the estimator?

```
A. estimator = SKLearn(source_directory='./script',
    inputs=[data_ref.as_named_input('data-folder').to_pandas_dataframe()],
    compute_target='local',
    entry_script='train.py')

B. script_params = {
    '--data-folder': data_ref.as_mount()
}
estimator = SKLearn(source_directory='./script',
    script_params=script_params,
    compute_target='local',
    entry_script='train.py')

C. estimator = SKLearn(source_directory='./script',
    inputs=[data_ref.as_named_input('data-folder').as_mount()],
    compute_target='local',
    entry_script='train.py')

D. script_params = {
    '--data-folder': data_ref.as_download(path_on_compute='csv_files')
}
estimator = SKLearn(source_directory='./script',
    script_params=script_params,
    compute_target='local',
    entry_script='train.py')

E. estimator = SKLearn(source_directory='./script',
    inputs=[data_ref.as_named_input('data-folder').as_download(path_on_compute='csv_files')],
    compute_target='local',
    entry_script='train.py')
```

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

Answer: B

Explanation:

Besides passing the dataset through the inputs parameter in the estimator, you can also pass the dataset through script_params and get the data path (mounting point) in your training script via arguments. This way, you can keep your training script independent of azureml-sdk. In other words, you will be able use the same training script for local debugging and remote training on any cloud platform.

Example:

```
from azureml.train.sklearn import SKLearn
script_params = {
    # mount the dataset on the remote compute and pass the mounted path as an argument to the training script '--data-folder':
    mnist_ds.as_named_input('mnist').as_mount(),
    '--regularization': 0.5
}
est = SKLearn(source_directory=script_folder, script_params=script_params, compute_target=compute_target, environment_definition=env,
    entry_script='train_mnist.py')
# Run the experiment
run = experiment.submit(est)
run.wait_for_completion(show_output=True)
Reference:
https://docs.microsoft.com/es-es/azure/machine-learning/how-to-train-with-datasets
```

NEW QUESTION 77

- (Exam Topic 3)

You use Azure Machine Learning to train and register a model.

You must deploy the model into production as a real-time web service to an inference cluster named service-compute that the IT department has created in the Azure Machine Learning workspace.

Client applications consuming the deployed web service must be authenticated based on their Azure Active Directory service principal.

You need to write a script that uses the Azure Machine Learning SDK to deploy the model. The necessary modules have been imported.

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

NOTE: Each correct selection is worth one point.


```
# Assume the necessary modules have been imported
deploy_target = 
    AksCompute
    AmlCompute
    RemoteCompute
    BatchCompute
 (ws, "service-compute")

deployment_config = 
    AksWebservice
    AciWebservice
    LocalWebService
 .deploy_configuration(cpu_cores=1, memory_gb=1,


    token_auth_enabled=True
    token_auth_enabled=False
    auth_enabled=True
    auth_enabled=False
 )

service = Model.deploy(ws, "ml-service",
    [model], inference_config, deployment_config, deploy_target)
service.wait_for_deployment(show_output = True)
```

- A. Mastered
B. Not Mastered

Answer: A

Explanation:

Box 1: AksCompute Example:

aks_target = AksCompute(ws,"myaks")

If deploying to a cluster configured for dev/test, ensure that it was created with enough
cores and memory to handle this deployment configuration. Note that memory is also used by
things such as dependencies and AML components.

deployment_config = AksWebservice.deploy_configuration(cpu_cores = 1, memory_gb = 1)

service = Model.deploy(ws, "myservice", [model], inference_config, deployment_config, aks_target)

Box 2: AksWebservice

Box 3: token_auth_enabled=Yes

Whether or not token auth is enabled for the Webservice.

Note: A Service principal defined in Azure Active Directory (Azure AD) can act as a principal on which authentication and authorization policies can be enforced in Azure Databricks.

The Azure Active Directory Authentication Library (ADAL) can be used to programmatically get an Azure AD access token for a user.

Reference:

<https://docs.microsoft.com/en-us/azure/machine-learning/how-to-deploy-azure-kubernetes-service> <https://docs.microsoft.com/en-us/azure/databricks/dev-tools/api/latest/aad/service-prin-aad-token>

NEW QUESTION 79

- (Exam Topic 3)

A coworker registers a datastore in a Machine Learning services workspace by using the following code:

```
Datastore.register_azure_blob_container(workspace=ws,
    datastore_name='demo_datastore',
    container_name='demo_datacontainer',
    account_name='demo_account',
    account_key='0A0A0A-0A0A00A-0A00A0A0A0A0A',
    create_if_not_exists=True)
```

You need to write code to access the datastore from a notebook.

Answer Area

```
import azureml.core
from azureml.core import Workspace, Datastore
ws = Workspace.from_config()

datastore = 
    Workspace
    Datastore
    Experiment
    Run
 .get( 
    ws
    run
    experiment
    log
 , ' 
    demo_datastore
    demo_datacontainer
    demo_account
    Datastore
 )
```

- A. Mastered
B. Not Mastered

Answer: A

Explanation:

Box 1: DataStore

To get a specific datastore registered in the current workspace, use the get() static method on the Datastore class:

Get a named datastore from the current workspace

datastore = Datastore.get(ws, datastore_name='your datastore name') Box 2: ws
Box 3: demo_datastore Reference:
<https://docs.microsoft.com/en-us/azure/machine-learning/how-to-access-data>

NEW QUESTION 83

- (Exam Topic 3)

You are building a machine learning model for translating English language textual content into French language textual content.

You need to build and train the machine learning model to learn the sequence of the textual content. Which type of neural network should you use?

- A. Multilayer Perceptions (MLPs)
- B. Convolutional Neural Networks (CNNs)
- C. Recurrent Neural Networks (RNNs)
- D. Generative Adversarial Networks (GANs)

Answer: C

Explanation:

To translate a corpus of English text to French, we need to build a recurrent neural network (RNN).

Note: RNNs are designed to take sequences of text as inputs or return sequences of text as outputs, or both. They're called recurrent because the network's hidden layers have a loop in which the output and cell state from each time step become inputs at the next time step. This recurrence serves as a form of memory. It allows contextual information to flow through the network so that relevant outputs from previous time steps can be applied to network operations at the current time step.

References:

<https://towardsdatascience.com/language-translation-with-rnns-d84d43b40571>

NEW QUESTION 84

- (Exam Topic 3)

You use the following code to run a script as an experiment in Azure Machine Learning:

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

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

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

Answer: B

Explanation:

You can list all of the files that are associated with this run record by called run.get_file_names() Reference:

<https://docs.microsoft.com/en-us/azure/machine-learning/how-to-track-experiments>

NEW QUESTION 88

- (Exam Topic 3)

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.

Actions	Answer Area
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.	

- A. Mastered
B. Not Mastered

Answer: A

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 92

- (Exam Topic 3)

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
C. configure the Link Azure ML workspace option.
D. From ML-workspace
E. create an attached compute.
F. From ML-workspace
G. create a compute cluster.

Answer: B

Explanation:

Reference:

<https://docs.microsoft.com/en-us/azure/machine-learning/how-to-use-mlflow-azure-databricks>

NEW QUESTION 95

- (Exam Topic 3)

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

Explanation:

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.parall> <https://docs.microsoft.com/en->

us/python/api/azureml-core/azureml.core.compute.amlcompute(class)

NEW QUESTION 97

- (Exam Topic 3)

You are performing clustering by using the K-means algorithm. You need to define the possible termination conditions.

Which three conditions can you use? Each correct answer presents a complete solution. NOTE: Each correct selection is worth one point.

- A. A fixed number of iterations is executed.
- B. The residual sum of squares (RSS) rises above a threshold.
- C. The sum of distances between centroids reaches a maximum.
- D. The residual sum of squares (RSS) falls below a threshold.
- E. Centroids do not change between iterations.

Answer: ADE

Explanation:

References:

<https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/k-means-clustering> <https://nlp.stanford.edu/IR-book/html/htmledition/k-means-1.html>

NEW QUESTION 101

- (Exam Topic 3)

You plan to deliver a hands-on workshop to several students. The workshop will focus on creating data visualizations using Python. Each student will use a device that has internet access.

Student devices are not configured for Python development. Students do not have administrator access to install software on their devices. Azure subscriptions are not available for students.

You need to ensure that students can run Python-based data visualization code. Which Azure tool should you use?

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

Answer: C

Explanation:

References:

<https://notebooks.azure.com/>

NEW QUESTION 104

- (Exam Topic 3)

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

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

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

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

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

Solution: Replace each missing value using the Multiple Imputation by Chained Equations (MICE) method. Does the solution meet the goal?

- A. Yes
- B. NO

Answer: A

Explanation:

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

Note: Multivariate imputation by chained equations (MICE), sometimes called “fully conditional specification” or “sequential regression multiple imputation” has emerged in the statistical literature as one principled method of addressing missing data. Creating multiple imputations, as opposed to single imputations, accounts for the statistical uncertainty in the imputations. In addition, the chained equations approach is very flexible and can handle variables of varying types (e.g., continuous or binary) as well as complexities such as bounds or survey skip patterns.

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

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

NEW QUESTION 107

- (Exam Topic 3)

You plan to run a Python script as an Azure Machine Learning experiment.

The script must read files from a hierarchy of folders. The files will be passed to the script as a dataset argument.

You must specify an appropriate mode for the dataset argument.

Which two modes can you use? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

- A. to_pandas_dataframe ()
- B. as_download()
- C. as_upload()
- D. as_mount ()

Answer: B

Explanation:

Reference:

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

NEW QUESTION 109

- (Exam Topic 3)

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

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

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

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

NOTE: Each correct selection is worth one point.

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

▼

-0.106276

0.149676

0.859122

1

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

▼

a positive linear relationship

a negative linear relationship

no linear relationship

A. Mastered

B. Not Mastered

Answer: A

Explanation:

Box 1: 0.859122

Box 2: a positively linear relationship

+1 indicates a strong positive linear relationship

-1 indicates a strong negative linear correlation

0 denotes no linear relationship between the two variables. References:

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

NEW QUESTION 112

- (Exam Topic 3)

You are using a decision tree algorithm. You have trained a model that generalizes well at a tree depth equal to 10.

You need to select the bias and variance properties of the model with varying tree depth values.

Which properties should you select for each tree depth? To answer, select the appropriate options in the answer area.

Answer Area

Tree Depth	Bias	Variance
5	<div>▼</div> <div>High</div> <div>Low</div> <div>Identical</div>	<div>▼</div> <div>High</div> <div>Low</div> <div>Identical</div>
15	<div>▼</div> <div>High</div> <div>Low</div> <div>Identical</div>	<div>▼</div> <div>High</div> <div>Low</div> <div>Identical</div>

A. Mastered

B. Not Mastered

Answer: A

Explanation:

In decision trees, the depth of the tree determines the variance. A complicated decision tree (e.g. deep) has low bias and high variance.

Note: In statistics and machine learning, the bias–variance tradeoff is the property of a set of predictive models whereby models with a lower bias in parameter estimation have a higher variance of the parameter estimates across samples, and vice versa. Increasing the bias will decrease the variance. Increasing the variance will decrease the bias.

References:

<https://machinelearningmastery.com/gentle-introduction-to-the-bias-variance-trade-off-in-machine-learning/>

NEW QUESTION 116

- (Exam Topic 3)

You use an Azure Machine Learning workspace.

You have a trained model that must be deployed as a web service. Users must authenticate by using Azure Active Directory.

What should you do?

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

Answer: A

Explanation:

To control token authentication, use the token_auth_enabled parameter when you create or update a deployment

Token authentication is disabled by default when you deploy to Azure Kubernetes Service.

Note: The model deployments created by Azure Machine Learning can be configured to use one of two authentication methods:

key-based: A static key is used to authenticate to the web service.

token-based: A temporary token must be obtained from the Azure Machine Learning workspace (using Azure Active Directory) and used to authenticate to the web service.

Reference:

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

NEW QUESTION 118

- (Exam Topic 3)

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

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

You are creating a new experiment in Azure Machine Learning Studio.

One class has a much smaller number of observations than the other classes in the training

You need to select an appropriate data sampling strategy to compensate for the class imbalance. Solution: You use the Synthetic Minority Oversampling Technique (SMOTE) sampling mode. Does the solution meet the goal?

- A. Yes
- B. No

Answer: A

Explanation:

SMOTE is used to increase the number of underrepresented cases in a dataset used for machine learning. SMOTE is a better way of increasing the number of rare cases than simply duplicating existing cases.

References:

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

NEW QUESTION 120

- (Exam Topic 3)

You are building an intelligent solution using machine learning models. The environment must support the following requirements:

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

You need to create the environment.

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

Actions	Answer area
Install the Azure Machine Learning SDK for Python on the cluster.	
When the cluster is ready, export Zeppelin notebooks to a local environment.	
Create and execute a Jupyter notebook by using automated machine learning (AutoML) on the cluster.	
Install Microsoft Machine Learning for Apache Spark.	⬅
When the cluster is ready and has processed the notebook, export your Jupyter notebook to a local environment.	⬆
Create an Azure HDInsight cluster to include the Apache Spark Mlib library.	
Create and execute the Zeppelin notebooks on the cluster.	
Create an Azure Databricks cluster.	

- A. Mastered
B. Not Mastered

Answer: A

Explanation:

Step 1: Create an Azure HDInsight cluster to include the Apache Spark Mlib library

Step 2: Install Microsot Machine Learning for Apache Spark You install AzureML on your Azure HDInsight cluster.

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

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

Step 4: When the cluster is ready, export Zeppelin notebooks to a local environment. Notebooks must be exportable to be version controlled locally.

References:

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

NEW QUESTION 122

- (Exam Topic 3)

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
<div><div></div><div>There will be 50 runs for this hyperparameter tuning experiment.</div></div>	<input type="radio"/>	<input type="radio"/>
<div><div></div><div>You can use the policy parameter in the HyperDriveConfig class to specify a security policy.</div></div>	<input type="radio"/>	<input type="radio"/>
<div><div></div><div>The experiment will create a run for every possible value for the learning rate parameter between 0.05 and 0.15.</div></div>	<input type="radio"/>	<input type="radio"/>

- A. Mastered
B. Not Mastered

Answer: A

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 125

- (Exam Topic 3)

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 Data Science Virtual Machine (DSVM) Windows edition.
B. Create a Geo AI Data Science Virtual Machine (Geo-DSVM) Windows edition.
C. Create a Deep Learning Virtual Machine (DLVM) Linux edition.
D. Create a Deep Learning Virtual Machine (DLVM) Windows edition.
E. Create a Data Science Virtual Machine (DSVM) Linux edition.

Answer: E

NEW QUESTION 130

- (Exam Topic 3)

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

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

You create a model to forecast weather conditions based on historical data.

You need to create a pipeline that runs a processing script to load data from a datastore and pass the processed data to a machine learning model training script.

Solution: Run the following code:

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

Does the solution meet the goal?

- A. Yes
B. No

Answer: B

Explanation:

Note: Data used in pipeline can be produced by one step and consumed in another step by providing a PipelineData object as an output of one step and an input of one or more subsequent steps.

Compare with this example, the pipeline train step depends on the process_step_output output of the pipeline process step:

```
from azureml.pipeline.core import Pipeline, PipelineData
from azureml.pipeline.steps import PythonScriptStep
```

```
datastore = ws.get_default_datastore()
```

```
process_step_output = PipelineData("processed_data", datastore=datastore)
process_step = PythonScriptStep(script_name="process.py",
```

```
arguments=["--data_for_train", process_step_output], outputs=[process_step_output], compute_target=aml_compute, source_directory=process_directory)
```

```
train_step = PythonScriptStep(script_name="train.py", arguments=["--data_for_train", process_step_output], inputs=[process_step_output],
```

```
compute_target=aml_compute, source_directory=train_directory)
```

```
pipeline = Pipeline(workspace=ws, steps=[process_step, train_step])
```

Reference:
<https://docs.microsoft.com/en-us/python/api/azureml-pipeline-core/azureml.pipeline.core.pipelinedata?view=azu>

NEW QUESTION 134

- (Exam Topic 3)

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()
```

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

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 137

- (Exam Topic 3)

You plan to provision an Azure Machine Learning Basic edition workspace for a data science project. You need to identify the tasks you will be able to perform in

the workspace.

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

NOTE: Each correct selection is worth one point. D

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

Answer: ABD

Explanation:

Reference:

<https://azure.microsoft.com/en-us/pricing/details/machine-learning/>

NEW QUESTION 138

- (Exam Topic 3)

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

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

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

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

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

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

Answer Area

```
import pandas as pd
salesData = pd.melt(
```

dataFrame

pandas

salesData

year

```
, id_vars='
```

shop

year

value

Shop X, Shop Y, Shop Z

```
', value_vars='
```

'shop'

'year'

['year']

['2017', '2018']

```
)
```

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: dataFrame

Syntax: pandas.melt(frame, id_vars=None, value_vars=None, var_name=None, value_name='value', col_level=None)[source]

Where frame is a DataFrame Box 2: shop

Paramter id_vars id_vars : tuple, list, or ndarray, optional Column(s) to use as identifier variables.

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

value_vars : tuple, list, or ndarray, optional

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

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

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

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

pd.melt(df, id_vars=['A'], value_vars=['B', 'C']) A variable value

0 a B 1

1 b B 3

2 c B 5

3 a C 2

4 b C 4

5 c C 6

References:

<https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.melt.html>

NEW QUESTION 139

- (Exam Topic 3)

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the

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

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

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:

```
from sklearn.metrics import roc_auc_score
import logging
# code to train model omitted
auc = roc_auc_score(y_test, y_predicted)
logging.info("AUC: " + str(auc))
```

Does the solution meet the goal?

- A. Yes
- B. No

Answer: A

Explanation:

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 144

- (Exam Topic 3)

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.

- A. Mastered
- B. Not Mastered

Answer: A

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 149

- (Exam Topic 3)

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

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

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

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

Option	Value
Technique	<div> <div>Grid search</div> <div>Outlier detection</div> <div>Dimensionality reduction</div> </div>
Constraint	<div> <div>Demographic parity</div> <div>False-positive rate parity</div> </div>

- A. Mastered
B. Not Mastered

Answer: A

Explanation:

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

Box 1: Grid Search

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

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

- Reduction: These algorithms take a standard black-box machine learning estimator (e.g., a LightGBM model) and generate a set of retrained models using a sequence of re-weighted training datasets.
- Post-processing: These algorithms take an existing classifier and the sensitive feature as input.

Box 2: Demographic parity

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

Reference:

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

NEW QUESTION 154

- (Exam Topic 3)

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

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

You need to add code to run the steps.

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

NOTE: Each correct selection is worth one point.

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

Answer: CD

Explanation:

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

Build the pipeline. Example:

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

Submit the pipeline to be run

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

Reference:

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

NEW QUESTION 158

- (Exam Topic 3)

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

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

An IT department creates the following Azure resource groups and resources:

Resource group	Resources
ml_resources	<ul style="list-style-type: none">• 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: <ul style="list-style-type: none">• 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

Explanation:

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 162

- (Exam Topic 3)

You must store data in Azure Blob Storage to support Azure Machine Learning. You need to transfer the data into Azure Blob Storage.

What are three possible ways to achieve the goal? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

- A. Bulk Insert SQL Query
B. AzCopy
C. Python script
D. Azure Storage Explorer
E. Bulk Copy Program (BCP)

Answer: BCD

Explanation:

You can move data to and from Azure Blob storage using different technologies: Azure Storage-Explorer

AzCopy Python SSIS

References:

<https://docs.microsoft.com/en-us/azure/machine-learning/team-data-science-process/move-azure-blob>

NEW QUESTION 167

- (Exam Topic 3)

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

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

You are creating a new experiment in Azure Machine Learning Studio.

One class has a much smaller number of observations than the other classes in the training set. You need to select an appropriate data sampling strategy to compensate for the class imbalance. Solution: You use the Scale and Reduce sampling mode.

Does the solution meet the goal?

- A. Yes
B. No

Answer: B

Explanation:

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

Note: SMOTE is used to increase the number of underrepresented cases in a dataset used for machine learning. SMOTE is a better way of increasing the number of rare cases than simply duplicating existing cases.

References:

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

NEW QUESTION 169

- (Exam Topic 3)

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

You need to select a data cleaning method. Which method should you use?

- A. Synthetic Minority
- B. Replace using Probabilistic PAC
- C. Replace using MICE
- D. Normalization

Answer: B

NEW QUESTION 170

- (Exam Topic 3)

You are evaluating a Python NumPy array that contains six data points defined as follows: data = [10, 20, 30, 40, 50, 60]

You must generate the following output by using the k-fold algorithm implantation in the Python Scikit-learn machine learning library:

train: [10 40 50 60], test: [20 30]

train: [20 30 40 60], test: [10 50]

train: [10 20 30 50], test: [40 60]

You need to implement a cross-validation to generate the output.

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

NOTE: Each correct selection is worth one point.

```
from numpy import array
from sklearn.model_selection import K-Means
k-fold
CrossValidation
ModelSelection

data = array([10, 20, 30, 40, 50, 60])
kfold = Kfold(n_splits=1
2
3
6, shuffle = True, random_state=1)

for train, test in kFold, split(data
k-fold
array
train, test):

print('train: %s, test: %5' % (data[train], data[test]))
```

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: k-fold

Box 2: 3

K-F olds cross-validator provides train/test indices to split data in train/test sets. Split dataset into k consecutive folds (without shuffling by default).

The parameter n_splits (int, default=3) is the number of folds. Must be at least 2. Box 3: data

Example: Example:

```
>>>
>>> from sklearn.model_selection import KFold
>>> X = np.array([[1, 2], [3, 4], [1, 2], [3, 4]])
>>> y = np.array([1, 2, 3, 4])
>>> kf = KFold(n_splits=2)
>>> kf.get_n_splits(X) 2
>>> print(kf)
KFold(n_splits=2, random_state=None, shuffle=False)
>>> for train_index, test_index in kf.split(X): print("TRAIN:", train_index, "TEST:", test_index) X_train, X_test = X[train_index], X[test_index] y_train, y_test =
y[train_index], y[test_index] TRAIN: [2 3] TEST: [0 1]
TRAIN: [0 1] TEST: [2 3]
```

References:

https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.KFold.html

NEW QUESTION 175

- (Exam Topic 3)

You are performing sentiment analysis using a CSV file that includes 12,000 customer reviews written in a short sentence format. You add the CSV file to Azure Machine Learning Studio and configure it as the starting point dataset of an experiment. You add the Extract N-Gram Features from Text module to the experiment to extract key phrases from the customer review column in the dataset.

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

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

NOTE: Each correct selection is worth one point.

Properties

Project

Extract N-Gram Features from Text

Text column

Selected columns

Column type: String Feature

Launch column selector

Vocabulary mode

▼

Create

ReadOnly

Update

Merge

N-Grams size

▼

3

4

4,000

12,000

0

Weighting function

▼

Minimum word length

3

Maximum word length

25

Minimum n-gram document absolu...

5

Maximum n-gram document ratio

1

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Vocabulary mode: Create

For Vocabulary mode, select Create to indicate that you are creating a new list of n-gram features. N-Grams size: 3

For N-Grams size, type a number that indicates the maximum size of the n-grams to extract and store. For example, if you type 3, unigrams, bigrams, and trigrams will be created.

Weighting function: Leave blank

The option, Weighting function, is required only if you merge or update vocabularies. It specifies how terms in the two vocabularies and their scores should be weighted against each other.

References:

<https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/extract-n-gram-features-from>

NEW QUESTION 179

- (Exam Topic 3)

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: AD

Explanation:

Reference:

<https://github.com/Azure/MachineLearningNotebooks/tree/master/how-to-use-azureml/machine-learningpipeline>

NEW QUESTION 181

- (Exam Topic 3)

HOTSPOT

You collect data from a nearby weather station. You have a pandas dataframe named weather_df that includes the following data:

Temperature	Observation_time	Humidity	Pressure	Visibility	Days_since_last observation
74	2019/10/2 00:00	0.62	29.87	3	0.5
89	2019/10/2 12:00	0.70	28.88	10	0.5
72	2019/10/3 00:00	0.64	30.00	8	0.5
80	2019/10/3 12:00	0.66	29.75	7	0.5

The data is collected every 12 hours: noon and midnight.

You plan to use automated machine learning to create a time-series model that predicts temperature over the next seven days. For the initial round of training, you want to train a maximum of 50 different models.

You must use the Azure Machine Learning SDK to run an automated machine learning experiment to train these models.

You need to configure the automated machine learning run.

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

NOTE: Each correct selection is worth one point.

```

automl_config = AutoMLConfig(task="
                                regression
                                forecasting
                                classification
                                deep learning

                                training_data=weather_df,
                                label_column_name="
                                humidity
                                pressure
                                visibility
                                temperature
                                days_since_last
                                observation_time

                                time_column_name="
                                humidity
                                pressure
                                visibility
                                temperature
                                days_since_last
                                observation_time

                                max_horizon=
                                2
                                6
                                7
                                12
                                14
                                50

                                iterations=
                                2
                                6
                                7
                                12
                                14
                                50

                                iteration_timeout_minutes=5,
                                primary_metric="r2_score")

```

- A. Mastered
B. Not Mastered

Answer: A

Explanation:

Box 1: forecasting

Task: The type of task to run. Values can be 'classification', 'regression', or 'forecasting' depending on the type of automated ML problem to solve.

Box 2: temperature

The training data to be used within the experiment. It should contain both training features and a label column (optionally a sample weights column).

Box 3: observation_time

time_column_name: The name of the time column. This parameter is required when forecasting to specify the datetime column in the input data used for building the time series and inferring its frequency. This setting is being deprecated. Please use forecasting_parameters instead.

Box 4: 7

"predicts temperature over the next seven days"

max_horizon: The desired maximum forecast horizon in units of time-series frequency. The default value is 1. Units are based on the time interval of your training data, e.g., monthly, weekly that the forecaster should

predict out. When task type is forecasting, this parameter is required.

Box 5: 50

"For the initial round of training, you want to train a maximum of 50 different models."

Iterations: The total number of different algorithm and parameter combinations to test during an automated ML experiment. Reference:

<https://docs.microsoft.com/en-us/python/api/azureml-train-automl-client/azureml.train.automl.automlconfig.auto>

NEW QUESTION 183

- (Exam Topic 3)

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

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

You 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,f2i

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.File.from_files(paths)
```

Does the solution meet the goal?

- A. Yes
B. No

Answer: B

Explanation:

Use two file paths.

Use Dataset.Tabular_from_delimited, instead of Dataset.File.from_files as the data isn't cleansed. Reference:

<https://docs.microsoft.com/en-us/azure/machine-learning/how-to-create-register-datasets>

NEW QUESTION 185

- (Exam Topic 3)

You plan to use the Hyperdrive feature of Azure Machine Learning to determine the optimal hyperparameter values when training a model.

You must use Hyperdrive to try combinations of the following hyperparameter values:

- learning_rate: any value between 0.001 and 0.1
- batch_size: 16, 32, or 64

You need to configure the search space for the Hyperdrive experiment.

Which two parameter expressions should you use? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

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

- D. a choice expression for batch_size
E. a uniform expression for batch_size

Answer: BD

Explanation:

B: Continuous hyperparameters are specified as a distribution over a continuous range of values. Supported distributions include:

- uniform(low, high) - Returns a value uniformly distributed between low and high

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

- one or more comma-separated values
 - a range object
 - any arbitrary list object
- Reference:

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

NEW QUESTION 187

- (Exam Topic 3)

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

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

You plan to use a Python script to run an Azure Machine Learning experiment. The script creates a reference to the experiment run context, loads data from a file, identifies the set of unique values for the label column, and completes the experiment run:

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

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

The experiment must record the unique labels in the data as metrics for the run that can be reviewed later.

You must add code to the script to record the unique label values as run metrics at the point indicated by the comment.

Solution: Replace the comment with the following code: for label_val in label_vals:

run.log('Label Values', label_val) Does the solution meet the goal?

- A. Yes
B. No

Answer: A

Explanation:

The run_log function is used 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 189

- (Exam Topic 3)

You have an Azure Machine Learning workspace named workspace1 that is accessible from a public endpoint. The workspace contains an Azure Blob storage datastore named store1 that represents a blob container in an Azure storage account named account1. You configure workspace1 and account1 to be accessible by using private endpoints in the same virtual network.

You must be able to access the contents of store1 by using the Azure Machine Learning SDK for Python. You must be able to preview the contents of store1 by using Azure Machine Learning studio.

You need to configure store1.

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

NOTE: Each correct selection is worth one point.

Requirement	Action
Access the contents of store1 by using the Azure Machine Learning SDK for Python.	<div><div></div><div><div>Set store1 as the default datastore.</div><div>Disable data validation for store1.</div><div>Update authentication for store1.</div><div>Regenerate the keys of account1.</div></div></div>
Preview the contents of store1 by using Azure Machine Learning studio.	<div><div></div><div><div>Set store1 as the default datastore.</div><div>Disable data validation for store1.</div><div>Update authentication for store1.</div><div>Regenerate the keys of account1.</div></div></div>

- A. Mastered
B. Not Mastered

Answer: A

Explanation:

Text, table Description automatically generated

Box 1: Regenerate the keys of account1.

Azure Blob Storage support authentication through Account key or SAS token.

To authenticate your access to the underlying storage service, you can provide either your account key, shared access signatures (SAS) tokens, or service principal

Box 2: Update the authentication for store1.

For Azure Machine Learning studio users, several features rely on the ability to read data from a dataset; such as dataset previews, profiles and automated machine learning. For these features to work with storage behind virtual networks, use a workspace managed identity in the studio to allow Azure Machine Learning to access the storage account from outside the virtual network.

Note: Some of the studio's features are disabled by default in a virtual network. To re-enable these features, you must enable managed identity for storage accounts you intend to use in the studio.

The following operations are disabled by default in a virtual network:

➤ Preview data in the studio.

Reference:

<https://docs.microsoft.com/en-us/azure/machine-learning/how-to-access-data>

NEW QUESTION 192

- (Exam Topic 3)

You plan to use the Hyperdrive feature of Azure Machine Learning to determine the optimal hyperparameter values when training a model.

You must use Hyperdrive to try combinations of the following hyperparameter values. You must not apply an early termination policy.

learning_rate: any value between 0.001 and 0.1

• batch_size: 16, 32, or 64

You need to configure the sampling method for the Hyperdrive experiment

Which two sampling methods can you use? Each correct answer is a complete solution. NOTE: Each correct selection is worth one point.

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

Answer: CD

Explanation:

C: Bayesian sampling is based on the Bayesian optimization algorithm and makes intelligent choices on the hyperparameter values to sample next. It picks the sample based on how the previous samples performed, such that the new sample improves the reported primary metric.

Bayesian sampling does not support any early termination policy Example:

```
from azureml.train.hyperdrive import BayesianParameterSampling
from azureml.train.hyperdrive import uniform, choice
param_sampling = BayesianParameterSampling( { "learning_rate": uniform(0.05, 0.1),
"batch_size": choice(16, 32, 64, 128)
})
```

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

Reference:

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

NEW QUESTION 197

- (Exam Topic 3)

You are evaluating a completed binary classification machine. You need to use the precision as the evaluation metric.

Which visualization should you use?

- A. scatter plot
- B. coefficient of determination
- C. Receiver Operating Characteristic (ROC) curve
- D. Gradient descent

Answer: C

Explanation:

Receiver operating characteristic (or ROC) is a plot of the correctly classified labels vs. the incorrectly classified labels for a particular model.

References:

<https://docs.microsoft.com/en-us/azure/machine-learning/how-to-understand-automated-ml#confusion-matrix>

NEW QUESTION 202

- (Exam Topic 3)

You are building an experiment using the Azure Machine Learning designer.

You split a dataset into training and testing sets. You select the Two-Class Boosted Decision Tree as the algorithm.

You need to determine the Area Under the Curve (AUC) of the model.

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

Modules

- Export Data
- Tune Model Hyperparameters
- Cross Validate Model
- Evaluate Model
- Score Model
- Train Model

Answer Area

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Step 1: Train Model

Two-Class Boosted Decision Tree

First, set up the boosted decision tree model.

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

* 2. Find the Train Model module, drag it onto the canvas, and then connect the output of the Two-Class Boosted Decision Tree module to the left input port of the Train Model module.

The Two-Class Boosted Decision Tree module initializes the generic model, and Train Model uses training data to train the model.

* 3. Connect the left output of the left Execute R Script module to the right input port of the Train Model

module (in this tutorial you used the data coming from the left side of the Split Data module for training). This portion of the experiment now looks something like this:



Step 2: Score Model

Score and evaluate the models

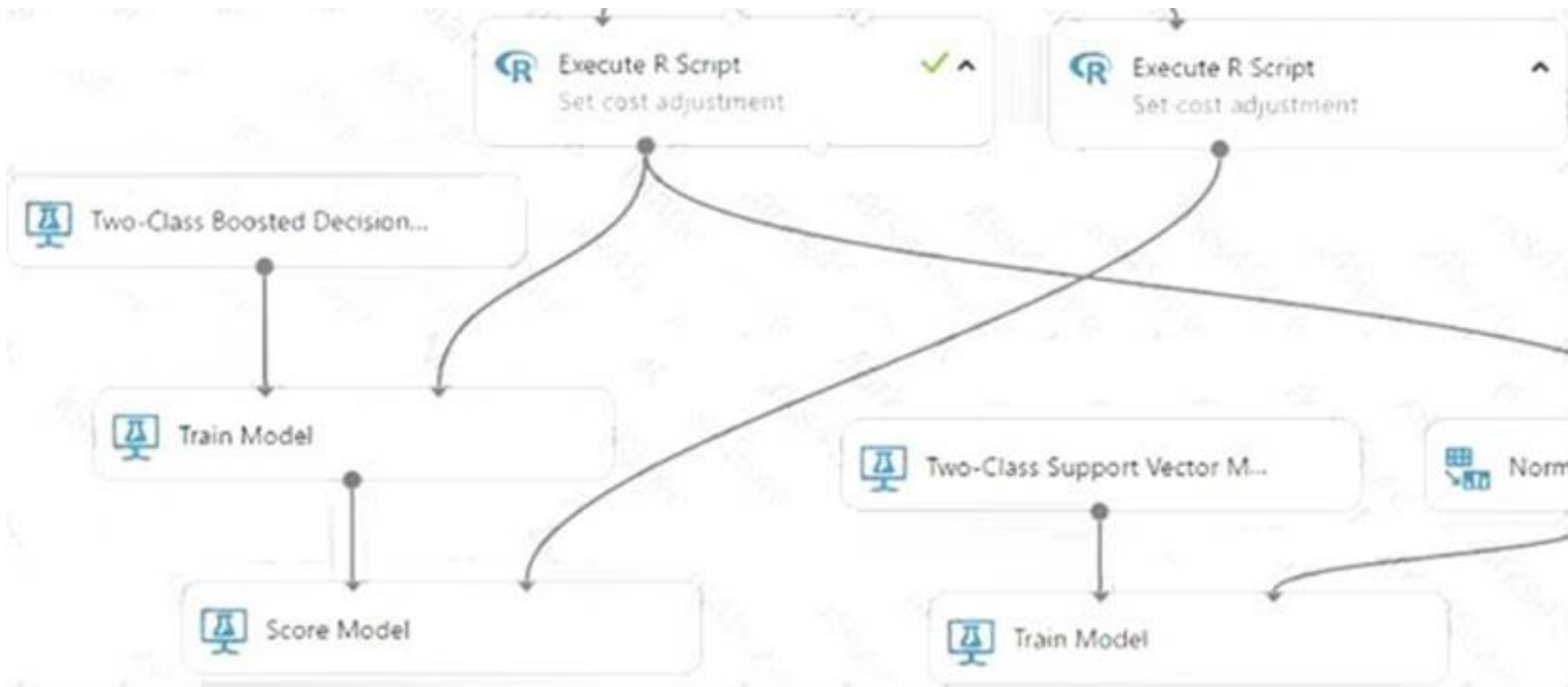
You use the testing data that was separated out by the Split Data module to score our trained models. You can then compare the results of the two models to see which generated better results.

Add the Score Model modules

* 1. Find the Score Model module and drag it onto the canvas.

* 2. Connect the Train Model module that's connected to the Two-Class Boosted Decision Tree module to the left input port of the Score Model module.

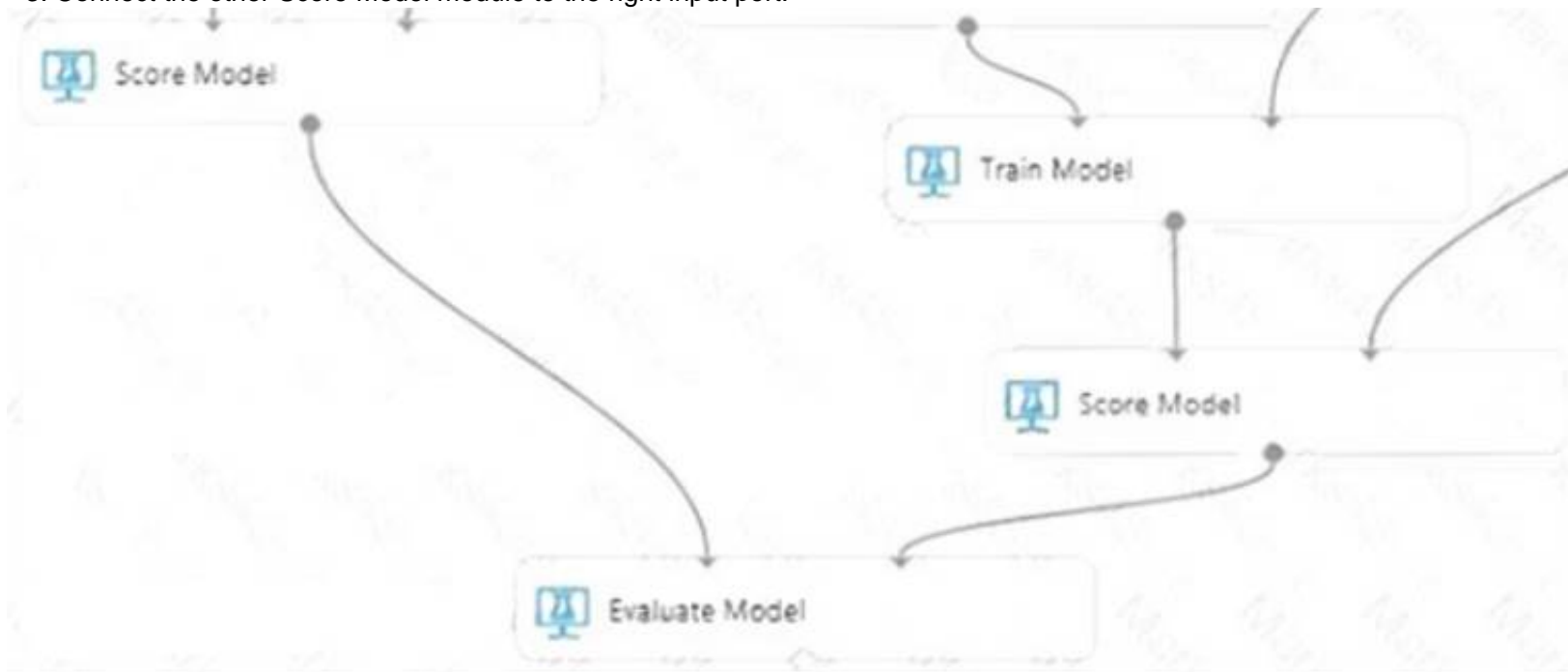
* 3. Connect the right Execute R Script module (our testing data) to the right input port of the Score Model module.



Step 3: Evaluate Model

To evaluate the two scoring results and compare them, you use an Evaluate Model module.

- * 1. Find the Evaluate Model module and drag it onto the canvas.
- * 2. Connect the output port of the Score Model module associated with the boosted decision tree model to the left input port of the Evaluate Model module.
- * 3. Connect the other Score Model module to the right input port.



NEW QUESTION 204

- (Exam Topic 3)

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

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

You train and register a machine learning model.

You plan to deploy the model as a real-time web service. Applications must use key-based authentication to use the model.

You need to deploy the web service.

Solution:

Create an AksWebservice instance.

Set the value of the auth_enabled property to False.

Set the value of the token_auth_enabled property to True.

Deploy the model to the service. Does the solution meet the goal?

A. Yes

B. No

Answer: B

Explanation:

Instead use only auth_enabled = TRUE Note: Key-based authentication.

Web services deployed on AKS have key-based auth enabled by default. ACI-deployed services have

key-based auth disabled by default, but you can enable it by setting auth_enabled = TRUE when creating the ACI web service. The following is an example of creating an ACI deployment configuration with key-based auth enabled.

deployment_config <- aci_webservice_deployment_config(cpu_cores = 1, memory_gb = 1,

auth_enabled = TRUE) Reference:

<https://azure.github.io/azureml-sdk-for-r/articles/deploying-models.html>

NEW QUESTION 209

- (Exam Topic 3)


```
source_directory=scripts_folder,  
entry_script="batch_pipeline.py",  
mini_batch_size="5",  
error_threshold=10,  
output_action="append_row",  
environment=batch_env,  
compute_target=compute_target,  
logging_level="DEBUG",  
node_count=4)
```

You need to obtain the output from the pipeline execution. Where will you find the output?

- A. the Activity Log in the Azure portal for the Machine Learning workspace
- B. a file named parallel_run_step.txt located in the output folder
- C. the digitidentification.py script
- D. the Inference Clusters tab in Machine Learning studio
- E. the debug log

Answer: B

Explanation:

output_action (str): How the output is to be organized. Currently supported values are 'append_row' and 'summary_only'.

➤ 'append_row' – All values output by run() method invocations will be aggregated into one unique file named parallel_run_step.txt that is created in the output location.

➤ 'summary_only' Reference:

<https://docs.microsoft.com/en-us/python/api/azureml-contrib-pipeline-steps/azureml.contrib.pipeline.steps.parall>

NEW QUESTION 213

- (Exam Topic 3)

You plan to use a Data Science Virtual Machine (DSVM) with the open source deep learning frameworks Caffe2 and Theano. You need to select a pre configured DSVM to support the framework.

What should you create?

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

Answer: E

NEW QUESTION 215

- (Exam Topic 3)

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

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

You are a data scientist using Azure Machine Learning Studio.

You need to normalize values to produce an output column into bins to predict a target column. Solution: Apply an Equal Width with Custom Start and Stop binning mode.

Does the solution meet the goal?

- A. Yes
- B. No

Answer: B

Explanation:

Use the Entropy MDL binning mode which has a target column. References:

<https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/group-data-into-bins>

NEW QUESTION 220

- (Exam Topic 3)

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

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

You are a data scientist using Azure Machine Learning Studio.

You need to normalize values to produce an output column into bins to predict a target column. Solution: Apply a Quantiles normalization with a QuantileIndex normalization.

Does the solution meet the GOAL?

- A. Yes
- B. No

Answer: B

Explanation:

Use the Entropy MDL binning mode which has a target column. References:

<https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/group-data-into-bins>

NEW QUESTION 224

- (Exam Topic 3)

You train a classification model by using a decision tree algorithm.

You create an estimator by running the following Python code. The variable `feature_names` is a list of all feature names, and `class_names` is a list of all class names.

```
from interpret.ext.blackbox import TabularExplainer
```

```
explainer = TabularExplainer(model,
                             x_train,
                             features=feature_names,
                             classes=class_names)
```

You need to explain the predictions made by the model for all classes by determining the importance of all features.

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 SHAP TreeExplainer will be used to interpret the model.	<input type="radio"/>	<input type="radio"/>
If you omit the features and classes parameters in the TabularExplainer instantiation, the explainer still works as expected.	<input type="radio"/>	<input type="radio"/>
You could interpret the model by using a MimicExplainer instead of a TabularExplainer.	<input type="radio"/>	<input type="radio"/>

A. Mastered

B. Not Mastered

Answer: A

Explanation:

For more detailed and visualizations more informative, you can choose to pass in feature names and output class names if doing classification.

Box 3: No

TabularExplainer automatically selects the most appropriate one for your use case, but you can call each of its three underlying explainers underneath (TreeExplainer, DeepExplainer, or KernelExplainer) directly.

Reference:

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

NEW QUESTION 229

- (Exam Topic 3)

You use the Azure Machine Learning designer to create and run a training pipeline. You then create a real-time inference pipeline.

You must deploy the real-time inference pipeline as a web service.

What must you do before you deploy the real-time inference pipeline?

A. Run the real-time inference pipeline.

B. Create a batch inference pipeline.

C. Clone the training pipeline.

D. Create an Azure Machine Learning compute cluster.

Answer: D

Explanation:

You need to create an inferencing cluster. Deploy the real-time endpoint

After your AKS service has finished provisioning, return to the real-time inferencing pipeline to complete deployment.

➤ Select Deploy above the canvas.

➤ Select Deploy new real-time endpoint.

➤ Select the AKS cluster you created.

➤ Select Deploy. Reference:

<https://docs.microsoft.com/en-us/azure/machine-learning/tutorial-designer-automobile-price-deploy>

NEW QUESTION 231

- (Exam Topic 3)

You create machine learning models by using Azure Machine Learning.

You plan to train and score models by using a variety of compute contexts. You also plan to create a new compute resource in Azure Machine Learning studio.

You need to select the appropriate compute types.

Which compute types should you select? To answer, drag the appropriate compute types to the correct requirements. Each compute type 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.

Compute types

- Attached compute
- Inference cluster
- Training cluster

Answer Area

Requirement	Compute type
Train models by using the Azure Machine Learning designer.	Compute type
Score new data through a trained model published as a real-time web service.	Compute type
Train models by using an Azure Databricks cluster.	Compute type
Deploy models by using the Azure Machine Learning designer.	Compute type

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: Attached compute

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

Box 2: Inference cluster Box 3: Training cluster Box 4: Attached compute

NEW QUESTION 232

- (Exam Topic 3)

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

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

You train and register a machine learning model.

You plan to deploy the model as a real-time web service. Applications must use key-based authentication to use the model.

You need to deploy the web service. Solution:

Create an AciWebService instance.

Set the value of the ssl_enabled property to True.

Deploy the model to the service. Does the solution meet the goal?

- A. Yes
- B. No

Answer: B

Explanation:

Instead use only auth_enabled = TRUE Note: Key-based authentication.

Web services deployed on AKS have key-based auth enabled by default. ACI-deployed services have

key-based auth disabled by default, but you can enable it by setting auth_enabled = TRUE when creating the ACI web service. The following is an example of creating an ACI deployment configuration with key-based auth enabled.

```
deployment_config <- aci_webService_deployment_config(cpu_cores = 1,
memory_gb = 1, auth_enabled = TRUE)
```

Reference:

<https://azure.github.io/azureml-sdk-for-r/articles/deploying-models.html>

NEW QUESTION 233

- (Exam Topic 3)

You are a data scientist working for a bank and have used Azure ML to train and register a machine learning model that predicts whether a customer is likely to repay a loan.

You want to understand how your model is making selections and must be sure that the model does not violate government regulations such as denying loans based on where an applicant lives.

You need to determine the extent to which each feature in the customer data is influencing predictions. What should you do?

- A. Enable data drift monitoring for the model and its training dataset.
- B. Score the model against some test data with known label values and use the results to calculate a confusion matrix.
- C. Use the Hyperdrive library to test the model with multiple hyperparameter values.
- D. Use the interpretability package to generate an explainer for the model.
- E. Add tags to the model registration indicating the names of the features in the training dataset.

Answer: D

Explanation:
for your model with different test data. The steps in this section show you how to compute and visualize engineered feature importance based on your test data.
Reference:
<https://docs.microsoft.com/en-us/azure/machine-learning/how-to-machine-learning-interpretability-automl>

NEW QUESTION 236

- (Exam Topic 3)
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

Explanation:
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 240

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

Answer Area

Parameter	Value
Estimator	<div>Estimator</div> <div>SkLearn</div> <div>PyTorch</div> <div>Tensorflow</div> <div>Chainer</div>
Training compute	<div>12 vCPU, 48 GB memory, 96 GB SSD</div> <div>12 vCPU, 112 GB memory, 680 GB SSD, 2 GPU, 24 GB GPU memory</div> <div>16 vCPU, 128 GB memory, 160 GB HDD, 80 GB NVME disk (4000 MBps)</div> <div>44 vCPU, 352 GB memory, 3.4 GHz CPU frequency all cores</div>

Click on the selection for training compute

- A. Mastered
- B. Not Mastered

Answer: A

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

NEW QUESTION 241

- (Exam Topic 3)
You create an Azure Machine Learning workspace named workspace1. You assign a custom role to a user of workspace1.
The custom role has the following JSON definition:

```
{
  "Name": "MyRole",
  "IsCustom": true,
  "Description": "New custom role description.",
  "Actions": ["*"],
  "NotActions": [
    "Microsoft.MachineLearningServices/workspaces/write",
    "Microsoft.MachineLearningServices/workspaces/computes/*/write",
    "Microsoft.MachineLearningServices/workspaces/computes/*/delete",
    "Microsoft.Authorization/*/write"
  ],
  "AssignableScopes": [
    "/subscriptions/<subscription_id>/resourceGroups/resourcegroup1/providers/
    Microsoft.MachineLearningServices/workspaces/workspace1"
  ]
}
```

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

Statements	Yes	No
The user can perform all actions in the workspace	<input type="radio"/>	<input type="radio"/>
The user can delete a compute resource in the workspace	<input type="radio"/>	<input type="radio"/>
The user can write metrics to the workspace	<input type="radio"/>	<input type="radio"/>

A.

Answer:

Explanation:

Graphical user interface, text, application, email Description automatically generated

Box 1: No

The actions listed in NotActions are prohibited.

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

Box 2: No

Deleting compute resources in the workspace is in the NotActions list. Box 3: Yes

Writing metrics is not listed in NotActions. Reference:

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

NEW QUESTION 244

- (Exam Topic 3)

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: AC

Explanation:

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-1838c203a43>

NEW QUESTION 249

- (Exam Topic 3)

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

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

You are using Azure Machine Learning Studio to perform feature engineering on a dataset. You need to normalize values to produce a feature column grouped

into bins.

Solution: Apply an Entropy Minimum Description Length (MDL) binning mode. Does the solution meet the goal?

- A. Yes
- B. No

Answer: A

Explanation:

Entropy MDL binning mode: This method requires that you select the column you want to predict and the column or columns that you want to group into bins. It then makes a pass over the data and attempts to determine the number of bins that minimizes the entropy. In other words, it chooses a number of bins that allows the data column to best predict the target column. It then returns the bin number associated with each row of your data in a column named <colname>quantized.

References:

<https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/group-data-into-bins>

NEW QUESTION 252

- (Exam Topic 3)

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

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

Answer: C

Explanation:

Azure Kubernetes Service (AKS) can be used real-time inference. Reference:

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

NEW QUESTION 255

- (Exam Topic 3)

You create a binary classification model to predict whether a person has a disease. You need to detect possible classification errors.

Which error type should you choose for each description? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Description	Error type
A person has a disease. The model classifies the case as having a disease.	<div>▼</div> <div>True Positives</div> <div>True Negatives</div> <div>False Positives</div> <div>False Negatives</div>
A person does not have a disease. The model classifies the case as having no disease.	<div>▼</div> <div>True Positives</div> <div>True Negatives</div> <div>False Positives</div> <div>False Negatives</div>
A person does not have a disease. The model classifies the case as having a disease.	<div>▼</div> <div>True Positives</div> <div>True Negatives</div> <div>False Positives</div> <div>False Negatives</div>
A person has a disease. The model classifies the case as having no disease.	<div>▼</div> <div>True Positives</div> <div>True Negatives</div> <div>False Positives</div> <div>False Negatives</div>

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: True Positive

A true positive is an outcome where the model correctly predicts the positive class Box 2: True Negative

A true negative is an outcome where the model correctly predicts the negative class. Box 3: False Positive

A false positive is an outcome where the model incorrectly predicts the positive class. Box 4: False Negative

A false negative is an outcome where the model incorrectly predicts the negative class. Note: Let's make the following definitions:

"Wolf" is a positive class. "No wolf" is a negative class.

We can summarize our "wolf-prediction" model using a 2x2 confusion matrix that depicts all four possible outcomes:

Reference:

<https://developers.google.com/machine-learning/crash-course/classification/true-false-positive-negative>

NEW QUESTION 260

.....

Thank You for Trying Our Product

* 100% Pass or Money Back

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

* One year free update

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

* Trusted by Millions

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

* Shop Securely

All transactions are protected by VeriSign!

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

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