DP-100^{Q&As}

DP-100^{Q&As}

Designing and Implementing a Data Science Solution on Azure

Pass Microsoft DP-100 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

https://www.pass4itsure.com/dp-100.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Microsoft
Official Exam Center

⚙ **Instant Download** After Purchase

⚙ **100% Money Back** Guarantee

⚙ **365 Days** Free Update

⚙ **800,000+** Satisfied Customers

QUESTION 1

An organization creates and deploys a multi-class image classification deep learning model that uses a set of labeled photographs.

The software engineering team reports there is a heavy inferencing load for the prediction web services during the summer. The production web service for the model fails to meet demand despite having a fully- utilized compute cluster where

the web service is deployed.

You need to improve performance of the image classification web service with minimal downtime and minimal administrative effort.

What should you advise the IT Operations team to do?

A. Create a new compute cluster by using larger VM sizes for the nodes, redeploy the web service to that cluster, and update the DNS registration for the service endpoint to point to the new cluster.

B. Increase the node count of the compute cluster where the web service is deployed.

C. Increase the minimum node count of the compute cluster where the web service is deployed.

D. Increase the VM size of nodes in the compute cluster where the web service is deployed.

Correct Answer: B

The Azure Machine Learning SDK does not provide support scaling an AKS cluster. To scale the nodes in the cluster, use the UI for your AKS cluster in the Azure Machine Learning studio. You can only change the node count, not the VM size of the cluster.

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

QUESTION 2

You use the Azure Machine Learning Python SDK to define a pipeline that consists of multiple steps.

When you run the pipeline, you observe that some steps do not run. The cached output from a previous run is used instead.

You need to ensure that every step in the pipeline is run, even if the parameters and contents of the source directory have not changed since the previous run.

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. Use a PipelineData object that references a datastore other than the default datastore.

B. Set the regenerate_outputs property of the pipeline to True.

C. Set the allow_reuse property of each step in the pipeline to False.

D. Restart the compute cluster where the pipeline experiment is configured to run.

E. Set the outputs property of each step in the pipeline to True.

Correct Answer: BC

B: If regenerate_outputs is set to True, a new submit will always force generation of all step outputs, and disallow data reuse for any step of this run. Once this run is complete, however, subsequent runs may reuse the results of this run.

C: Keep the following in mind when working with pipeline steps, input/output data, and step reuse.

If data used in a step is in a datastore and allow_reuse is True, then changes to the data change won\\'t be detected. If the data is uploaded as part of the snapshot (under the step\\'s source_directory), though this is not recommended, then the

hash will change and will trigger a rerun.

Reference:

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

https://github.com/Azure/MachineLearningNotebooks/blob/master/how-to-use-azureml/machine-learning-pipelines/intro-to-pipelines/aml-pipelines-getting-started.ipynb

QUESTION 3

You plan to use automated machine learning by using Azure Machine Learning Python SDK v2 to train a regression model. You have data that has features with missing values, and categorical features with few distinct values.

You need to control whether automated machine learning automatically imputes missing values and encode categorical features as part of the training task.

Which enum of the automl package should you use?

A. ForecastHorizonMode

B. RegressionModels

C. FeaturizationMode

D. RegressionPrimaryMetrics

Correct Answer: C

AutoMLConfig Class

Represents configuration for submitting an automated ML experiment in Azure Machine Learning.

This configuration object contains and persists the parameters for configuring the experiment run, as well as the training data to be used at run time.

Constructor

Featurization str or FeaturizationConfig

\\'auto\\' / \\'off\\' / FeaturizationConfig Indicator for whether featurization step should be done automatically or not, or whether customized featurization should be used.

Column type is automatically detected. Based on the detected column type preprocessing/featurization is done as follows:

Categorical: Target encoding, one hot encoding, drop high cardinality categories, impute missing values.

*-> Numeric: Impute missing values, cluster distance, weight of evidence.

DateTime: Several features such as day, seconds, minutes, hours etc.

Text: Bag of words, pre-trained Word embedding, text target encoding.

Reference:

https://learn.microsoft.com/en-us/python/api/azureml-train-automl-client/azureml.train.automl.automlconfig.automlconfig
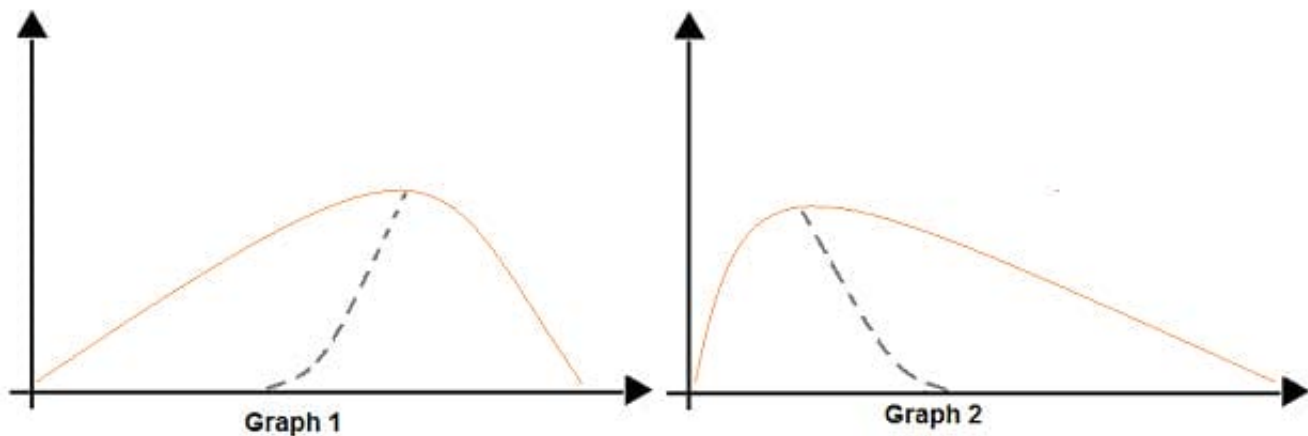https://docs.microsoft.com/en-us/azure/machine-learning/how-to-use-automated-ml-for-ml-models

QUESTION 4

HOTSPOT

You are analyzing the asymmetry in a statistical distribution.

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

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

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area

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

Correct Answer:

Answer Area

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

Box 1: Positive skew

Positive skew values means the distribution is skewed to the right.

Box 2: Negative skew

Negative skewness values mean the distribution is skewed to the left.

References:

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

QUESTION 5

DRAG DROP

You have a dataset that contains over 150 features. You use the dataset to train a Support Vector Machine (SVM) binary classifier.

You need to use the Permutation Feature Importance module in Azure Machine Learning Studio to compute a set of feature importance scores for the dataset.

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

Select and Place:

Actions	Answer Area
Add a Split Data module to create training and test datasets.	
Set the Metric for measuring performance property to Classification – Accuracy and then run the experiment	
Add a Permutation Feature importance module and connect the trained model and test dataset.	
Add a dataset to the experiment.	
Add a Two-class Support Vector Machine module to initialize the SVM classifier.	

Correct Answer:

Answer Area	Actions
	Add a Two-class Support Vector Machine module to initialize the SVM classifier.
	Add a dataset to the experiment.
	Add a Split Data module to create training and test datasets.
	Add a Permutation Feature importance module and connect the trained model and test dataset.
	Set the Metric for measuring performance property to Classification – Accuracy and then run the experiment

Step 1: Add a Two-Class Support Vector Machine module to initialize the SVM classifier.

Step 2: Add a dataset to the experiment

Step 3: Add a Split Data module to create training and test dataset.

To generate a set of feature scores requires that you have an already trained model, as well as a test dataset.

Step 4: Add a Permutation Feature Importance module and connect to the trained model and test dataset.

Step 5: Set the Metric for measuring performance property to Classification - Accuracy and then run the experiment.

Reference:

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

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

Latest DP-100 Dumps   DP-100 PDF Dumps   DP-100 Practice Test