> **Vendor: Microsoft**

> **Exam Code: DP-100**

> **Exam Name: Designing and Implementing a Data Science Solution on Azure**

> **New Questions (May/2020)**

Visit PassLeader and Download Full Version DP-100 Exam Dumps

NEW QUESTION 162
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. Choose three.)

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:
Not C and E: The UI is included the Enterprise edition only.
https://azure.microsoft.com/en-us/pricing/details/machine-learning/

NEW QUESTION 163
You create a deep learning model for image recognition on Azure Machine Learning service using GPU- based training. You must deploy the model to a context that allows for real-time GPU-based inferencing. You need to configure compute resources for model inferencing. Which compute type should you use?

A. Azure Container Instance
B. Azure Kubernetes Service
C. Field Programmable Gate Array
D. Machine Learning Compute

Answer: B
Explanation:
You can use Azure Machine Learning to deploy a GPU-enabled model as a web service. Deploying a model on Azure Kubernetes Service (AKS) is one option. The AKS cluster provides a GPU resource that is used by the model for inference. Inference, or model scoring, is the phase where the deployed model is used to make predictions. Using GPUs instead of CPUs offers performance advantages on highly parallelizable computation.
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-deploy-inferencing-gpus

NEW QUESTION 164
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. Choose two.)

A. run(mini_batch)
B. main()
C. batch()
D. init()
E. score(mini_batch)

Answer: AD
Explanation:
https://github.com/Azure/MachineLearningNotebooks/tree/master/how-to-use-azureml/machine-learning-pipelines/parallel-run

NEW QUESTION 165
You create a multi-class image classification deep learning model. You train the model by using PyTorch version 1.2. You need to ensure that the correct version of PyTorch can be identified for the inferencing environment when the model is deployed. What should you do?

A. Save the model locally as a.pt file, and deploy the model as a local web service.
B. Deploy the model on computer that is configured to use the default Azure Machine Learning conda environment.
C. Register the model with a .pt file extension and the default version property.
D. Register the model, specifying the model_framework and model_framework_version properties.

Answer: D
Explanation:
framework_version: The PyTorch version to be used for executing training code.
https://docs.microsoft.com/en-us/python/api/azureml-train-core/azureml.train.dnn.pytorch?view=azure-ml-py

NEW QUESTION 166
You are a lead data scientist for a project that tracks the health and migration of birds. You create a multi- class image classification deep learning model that uses a set of labeled bird photographs collected by experts. You have 100,000 photographs of birds. All photographs use the JPG format and are stored in an Azure blob container in an Azure subscription. You need to access the bird photograph files in the Azure blob container from the Azure Machine Learning service workspace that will be used for deep learning model training. You must minimize data movement. What should you do?

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

Answer: D

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

NEW QUESTION 167
You are creating a new Azure Machine Learning pipeline using the designer. The pipeline must train a model using data in a comma-separated values (CSV) file that is published on a website. You have not created a dataset for this file. You need to ingest the data from the CSV file into the designer pipeline using the minimal administrative effort. Which module should you add to the pipeline in Designer?

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

Answer: D
Explanation:
The preferred way to provide data to a pipeline is a Dataset object. The Dataset object points to data that lives in or is accessible from a datastore or at a Web URL. The Dataset class is abstract, so you will create an instance of either a FileDataset (referring to one or more files) or a TabularDataset that's created by from one or more files with delimited columns of data. https://docs.microsoft.com/en-us/azure/machine-learning/how-to-create-your-first-pipeline

NEW QUESTION 168
You have a comma-separated values (CSV) file containing data from which you want to train a classification model. You are using the Automated Machine Learning interface in Azure Machine Learning studio to train the classification model. You set the task type to Classification. You need to ensure that the Automated Machine Learning process evaluates only linear models. What should you do?

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

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

NEW QUESTION 169
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:
When you compute model explanations and visualize them, you're not limited to an existing model explanation for an automated ML model. You can also get an 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.
Incorrect:
Not A: In the context of machine learning, data drift is the change in model input data that leads to model performance degradation. It is one of the top reasons where model accuracy degrades over time, thus monitoring data drift helps detect model performance issues.
Not B: A confusion matrix is used to describe the performance of a classification model. Each row displays the instances of the true, or actual class in your dataset, and each column represents the instances of the class that was predicted by the model.
Not C: Hyperparameters are adjustable parameters you choose for model training that guide the training process. The HyperDrive package helps you automate choosing these parameters.
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-machine-learning-interpretability
-automl

NEW QUESTION 170
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.
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-train-ml-models

NEW QUESTION 171
You are creating a classification model for a banking company to identify possible instances of credit card fraud. You plan to create the model in Azure Machine Learning by using automated machine learning. The training dataset that you are using is highly unbalanced. You need to evaluate the classification model. Which primary metric should you use?

A. normalized_mean_absolute_error
B. AUC_weighted
C. accuracy

D. normalized_root_mean_squared_error
E. spearman_correlation

Answer: B
Explanation:
AUC_weighted is a Classification metric. AUC is the Area under the Receiver Operating Characteristic Curve. Weighted is the arithmetic mean of the score for each class, weighted by the number of true instances in each class.
Incorrect:
Not A: normalized_mean_absolute_error is a regression metric, not a classification metric.
Not C: When comparing approaches to imbalanced classification problems, consider using metrics beyond accuracy such as recall, precision, and AUROC. It may be that switching the metric you optimize for during parameter selection or model selection is enough to provide desirable performance detecting the minority class.
Not D: normalized_root_mean_squared_error is a regression metric, not a classification metric.
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-understand-automated-ml

NEW QUESTION 172
Hotspot
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.)

Answer Area

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

Answer:

Answer Area

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

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

NEW QUESTION 173
Drag and Drop
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.

Answer:

Actions

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 a PythonScriptStep() to run image_fetcher.py on the cpu-compute compute target.

Configure a PythonScriptStep() to run bird_classifier_train.py on the cpu-compute compute target.

Configure a PythonScriptStep() to run bird_classifier_train.py on the gpu-compute compute target.

Answer Area

Configure a DataTransferStep() to fetch new image data from public web portal, running on the cpu-compute compute target.

Configure a PythonScriptStep() to run image_resize.py on the cpu-compute compute target.

Configure an EstimatorStep() to run an estimator that runs the bird_classifier_train.py model training script on the cpu_compute compute target.

Explanation:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-train-pytorch

NEW QUESTION 174
Hotspot
You are a lead data scientist for a project that tracks the health and migration of birds. You create a multi- image classification deep learning model that uses a set of labeled bird photos collected by experts. You plan to use the model to develop a cross-platform mobile app that predicts the species of bird captured by app users. You must test and deploy the trained model as a web service. The deployed model must meet the following requirements:
- An authenticated connection must not be required for testing.
- The deployed model must perform with low latency during inferencing.
- The REST endpoints must be scalable and should have a capacity to handle large number of requests when multiple end users are using the mobile application.
You need to verify that the web service returns predictions in the expected JSON format when a valid REST request is submitted. Which compute resources should you use? (To answer, select the appropriate options in the answer area.)

Answer Area

Context	Resource
Test	▼
	ds-workstation notebook VM
	aks-compute cluster
	cpu-compute cluster
	gpu-compute cluster
Production	▼
	ds-workstation notebook VM
	aks-compute cluster
	cpu-compute cluster
	gpu-compute cluster

Answer:

Answer Area

Context	Resource
Test	▼
	ds-workstation notebook VM
	aks-compute cluster
	cpu-compute cluster
	gpu-compute cluster
Production	▼
	ds-workstation notebook VM
	aks-compute cluster
	cpu-compute cluster
	gpu-compute cluster

Explanation:
Box 1: ds-workstation notebook VM. An authenticated connection must not be required for testing. On a Microsoft Azure virtual machine (VM), including a Data Science Virtual Machine (DSVM), you create local user accounts while provisioning the VM. Users then authenticate to the VM by using these credentials.
Box 2: gpu-compute cluster. Image classification is well suited for GPU compute clusters.
https://docs.microsoft.com/en-us/azure/machine-learning/data-science-virtual-machine/dsvm-common-identity
https://docs.microsoft.com/en-us/azure/architecture/reference-architectures/ai/training-deep-learning

NEW QUESTION 175
Drag and Drop
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.)

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

Answer:

Compute types

- Attached compute
- Inference cluster
- Training cluster

Answer Area

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

NEW QUESTION 176
......

Visit PassLeader and Download Full Version DP-100 Exam Dumps