Microsoft.DP-100.v2024-12-18.q106

Exam Code:	DP-100
Exam Name:	Designing and Implementing a Data Science Solution on Azure
Certification Provider:	Microsoft
Free Question Number:	106
Version:	v2024-12-18
# of views:	2268
# of Questions views:	1060
https://www.exam-tests.com/DP-100-exam/Microsoft.DP-100.v2024-12-18.q106.html	

NEW QUESTION: 1
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.
After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.
You train a classification model by using a logistic regression algorithm.
You must be able to explain the model's predictions by calculating the importance of each feature, both as an overall global relative importance value and as a measure of local importance for a specific set of predictions.
You need to create an explainer that you can use to retrieve the required global and local feature importance values.
Solution: Create a TabularExplainer.
Does the solution meet the goal?
A. Yes
B. No
Answer: (SHOW ANSWER)
Explanation
Instead use Permutation Feature Importance Explainer (PFI).
Note 1:
Note 2: Permutation Feature Importance Explainer (PFI): Permutation Feature Importance is a technique used to explain classification and regression models. At a high level, the way it works is by randomly shuffling data one feature at a time for the entire dataset and calculating how much the performance metric of interest changes. The larger the change, the more important that feature is. PFI can explain the overall behavior of any underlying model but does not explain individual predictions.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-machine-learning-interpretability

NEW QUESTION: 2
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:
You plan to use this configuration to run a script that trains a random forest model and then tests it with validation dat a. The label values for the validation data are stored in a variable named y_test variable, and the predicted probabilities from the model are stored in a variable named y_predicted.
You need to add logging to the script to allow Hyperdrive to optimize hyperparameters for the AUC metric. Solution: Run the following code:
Does the solution meet the goal?
A. Yes
B. No
Answer: (SHOW ANSWER)
Explanation
Use a solution with logging.info(message) instead.
Note: Python printing/logging example:
logging.info(message)
Destination: Driver logs, Azure Machine Learning designer
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-debug-pipelines

NEW QUESTION: 3
You deploy a model in Azure Container Instance.
You must use the Azure Machine Learning SDK to call the model API.
You need to invoke the deployed model using native SDK classes and methods.
How should you complete the command? To answer, select the appropriate options in the answer areas.
NOTE: Each correct selection is worth one point.
Answer:
Reference:
https://docs.microsoft.com/bs-latn-ba/azure/machine-learning/how-to-deploy-azure-container-instance
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-troubleshoot-deployment

NEW QUESTION: 4
You need to define an evaluation strategy for the crowd sentiment models.
Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.
Answer:
Explanation:
Scenario:
Experiments for local crowd sentiment models must combine local penalty detection data.
Crowd sentiment models must identify known sounds such as cheers and known catch phrases. Individual crowd sentiment models will detect similar sounds.

Note: Evaluate the changed in correlation between model error rate and centroid distance In machine learning, a nearest centroid classifier or nearest prototype classifier is a classification model that assigns to observations the label of the class of training samples whose mean (centroid) is closest to the observation.
References:
https://en.wikipedia.org/wiki/Nearest_centroid_classifier
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/sweep-clustering

NEW QUESTION: 5
You have machine learning models produce unfair predictions across sensitive features.
You must use a post-processing technique to apply a constraint to the models to mitigate their unfairness.
You need to select a post-processing technique and model type.
What should you use? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.
Answer:

NEW QUESTION: 6
You deploy a model in Azure Container Instance.
You must use the Azure Machine Learning SDK to call the model API.
You need to invoke the deployed model using native SDK classes and methods.
How should you complete the command? To answer, select the appropriate options in the answer areas.
NOTE: Each correct selection is worth one point.
Answer:
Reference:
https://docs.microsoft.com/bs-latn-ba/azure/machine-learning/how-to-deploy-azure-container-instance
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-troubleshoot-deployment

NEW QUESTION: 7
You are using C-Support Vector classification to do a multi-class classification with an unbalanced training dataset. The C-Support Vector classification using Python code shown below:
You need to evaluate the C-Support Vector classification code.
Which evaluation statement should you use? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.
Answer:
Explanation
Box 1: Automatically adjust weights inversely proportional to class frequencies in the input data The "balanced" mode uses the values of y to automatically adjust weights inversely proportional to class frequencies in the input data as n_samples / (n_classes * np.bincount(y)).
Box 2: Penalty parameter
Parameter: C : float, optional (default=1.0)
Penalty parameter C of the error term.
References:

https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html

NEW QUESTION: 8
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:
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 (LEAVE A REPLY)
Need to attach the mlvm virtual machine as a compute target in the Azure Machine Learning workspace.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/concept-compute-target

NEW QUESTION: 9
You have a multi-class image classification deep learning model that uses a set of labeled photographs. You create the following code to select hyperparameter values when training the model.
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:
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-tune-hyperparameters

NEW QUESTION: 10
You plan to preprocess text from CSV files. You load the Azure Machine Learning Studio default stop words list.
You need to configure the Preprocess Text module to meet the following requirements:
Ensure that multiple related words from a single canonical form.
Remove pipe characters from text.
Remove words to optimize information retrieval.
Which three options should you select? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.
Answer:
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/preprocess-text

NEW QUESTION: 11
You are developing a machine learning solution by using the Azure Machine Learning designer.
You need to create a web service that applications can use to submit data feature values and retrieve a predicted label.
Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.
Answer:
1 - Create and run a training pipeling
2 - Deploy a service to an inference cluster.
3 - Create and run a real-time inference pipeline.

NEW QUESTION: 12
space 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:
Reference:
https://docs.microsoft.com/en-us/python/api/azureml-train-core/azureml.train.dnn

NEW QUESTION: 13
You collect data from a nearby weather station. You have a pandas dataframe named weather_df that includes the following data:
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.
Answer:
Explanation
Box 1: forcasting
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: 14
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.
Answer:
Explanation
Box 1: Attached compute
Box 2: Inference cluster
Box 3: Training cluster
Box 4: Attached compute

NEW QUESTION: 15
You are using a Git repository to track work in an Azure Machine Learning workspace.
You need to authenticate a Git account by using SSH.
Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.
Answer:
Explanation
Graphical user interface, text, application, chat or text message Description automatically generated
Authenticate your Git Account with SSH:
Step 1: Generating a public/private key pair
Generate a new SSH key
1. Open the terminal window in the Azure Machine Learning Notebook Tab.
2. Paste the text below, substituting in your email address.
ssh-keygen -t rsa -b 4096 -C "your_email@example.com"

This creates a new ssh key, using the provided email as a label.
> Generating public/private rsa key pair.
Step 2: Add the public key to the Git Account
In your terminal window, copy the contents of your public key file.
Step 3: Clone the Git repository by using an SSH repository URL
1. Copy the SSH Git clone URL from the Git repo.
2. Paste the url into the git clone command below, to use your SSH Git repo URL. This will look something like:
git clone git@example.com:GitUser/azureml-example.git
Cloning into 'azureml-example'.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/concept-train-model-git-integration

NEW QUESTION: 16
You need to use the Python language to build a sampling strategy for the global penalty detection models.
How should you complete the code segment? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.
Answer:
Explanation
Box 1: import pytorch as deeplearninglib
Box 2: ..DistributedSampler(Sampler)..
DistributedSampler(Sampler):
Sampler that restricts data loading to a subset of the dataset.
It is especially useful in conjunction with class:`torch.nn.parallel.DistributedDataParallel`. In such case, each process can pass a DistributedSampler instance as a DataLoader sampler, and load a subset of the original dataset that is exclusive to it.
Scenario: Sampling must guarantee mutual and collective exclusively between local and global segmentation models that share the same features.
Box 3: optimizer = deeplearninglib.train. GradientDescentOptimizer(learning_rate=0.10)

NEW QUESTION: 17
A set of CSV files contains sales records. All the CSV files have the same data schema.
Each CSV file contains the sales record for a particular month and has the filename sales.csv. Each file in stored in a folder that indicates the month and year when the data was recorded. The folders are in an Azure blob container for which a datastore has been defined in an Azure Machine Learning workspace. The folders are organized in a parent folder named sales to create the following hierarchical structure:
At the end of each month, a new folder with that month's sales file is added to the sales folder.
You plan to use the sales data to train a machine learning model based on the following requirements:

You must define a dataset that loads all of the sales data to date into a structure that can be easily converted to a dataframe.
You must be able to create experiments that use only data that was created before a specific previous month, ignoring any data that was added after that month.
You must register the minimum number of datasets possible.
You need to register the sales data as a dataset in Azure Machine Learning service workspace.
What should you do?
A. Create a tabular dataset that references the datastore and explicitly specifies each 'sales/mm-yyyy/ sales.csv' file every month. Register the dataset with the name sales_dataset each month, replacing the existing dataset and specifying a tag named indicating the month and year it was registered. Use this dataset for all experiments.
B. Create a tabular dataset that references the datastore and specifies the path 'sales/*/sales.csv', register the dataset with the name sales_dataset and a tag named month indicating the month and year it was registered, and use this dataset for all experiments.
C. Create a new tabular dataset that references the datastore and explicitly specifies each 'sales/mm-yyyy/ sales.csv' file every month. Register the dataset with the name sales_dataset_MM-YYYY each month with appropriate MM and YYYY values for the month and year. Use the appropriate month-specific dataset for experiments.
D. Create a tabular dataset that references the datastore and explicitly specifies each 'sales/mm-yyyy/ sales.csv' file. Register the dataset with the name sales_dataset each month as a new version and with a tag named month indicating the month and year it was registered. Use this dataset for all experiments, identifying the version to be used based on the month tag as necessary.
Answer: B (LEAVE A REPLY)
Explanation
Specify the path.
Example:
The following code gets the workspace existing workspace and the desired datastore by name. And then passes the datastore and file locations to the path parameter to create a new TabularDataset, weather_ds.

```
from azureml.core import Workspace, Datastore, Dataset
datastore_name = 'your datastore name'
# get existing workspace
workspace = Workspace.from_config()
# retrieve an existing datastore in the workspace by name
datastore = Datastore.get(workspace, datastore_name)
# create a TabularDataset from 3 file paths in datastore
datastore_paths = [(datastore, 'weather/2018/11.csv'),
(datastore, 'weather/2018/12.csv'),
(datastore, 'weather/2019/*.csv')]
weather_ds = Dataset.Tabular.from_delimited_files(path=datastore_paths)
```

NEW QUESTION: 18
You manage an Azure Machine Learning workspace named workspace1by using the Python SDK v2.
You must register datastores in workspace 1 for Azure Blot storage and Azure Fetes storage to meet the following requirements.
* Azure Active Directory (Azure AD) authentication must be used for access to storage when possible.
* Credentials and secrets steed in workspace1 must be valid lot a specified time period when accessing Azure Files storage.
You need to configure a security access method used to register the Azure Blob and azure files storage in workspace1.

Which security access method should you configure? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.
Answer:

NEW QUESTION: 19
You train a model by using Azure Machine Learning. You use Azure Blob Storage to store production data.
The model must be re-trained when new data is uploaded to Azure Blob Storage. You need to minimize development and coding.
You need to configure Azure services to develop a re-training solution.
Which Azure services should you use? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.
Answer:

NEW QUESTION: 20
Your Azure Machine Learning workspace has a dataset named real_estate_dat a. A sample of the data in the dataset follows.
You want to use automated machine learning to find the best regression model for predicting the price column.
You need to configure an automated machine learning experiment using the Azure Machine Learning SDK.
How should you complete the code? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.
Answer:
Reference:
https://docs.microsoft.com/en-us/python/api/azureml-train-automl-client/azureml.train.automl.automlconfig.automlconfig?view=azure-ml-py

NEW QUESTION: 21
You publish a batch inferencing pipeline that will be used by a business application.
The application developers need to know which information should be submitted to and returned by the REST interface for the published pipeline.
You need to identify the information required in the REST request and returned as a response from the published pipeline.
Which values should you use in the REST request and to expect in the response? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.
Answer:
Explanation:
Box 1: JSON containing an OAuth bearer token
Specify your authentication header in the request.
To run the pipeline from the REST endpoint, you need an OAuth2 Bearer-type authentication header.
Box 2: JSON containing the experiment name
Add a JSON payload object that has the experiment name.
Example:
rest_endpoint = published_pipeline.endpoint
response = requests.post(rest_endpoint,
headers=auth_header,
json={"ExperimentName": "batch_scoring",

"ParameterAssignments": {"process_count_per_node": 6}})
run_id = response.json()["Id"]
Box 3: JSON containing the run ID
Make the request to trigger the run. Include code to access the Id key from the response dictionary to get the value of the run ID.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/tutorial-pipeline-batch-scoring-classification

NEW QUESTION: 22
You arc I mating a deep learning model to identify cats and dogs. You have 25,000 color images.
You must meet the following requirements:
* Reduce the number of training epochs.
* Reduce the size of the neural network.
* Reduce over-fitting of the neural network.
You need to select the image modification values.
Which value should you use? To answer, select the appropriate Options in the answer area.
NOTE: Each correct selection is worth one point.
Answer:

NEW QUESTION: 23
You create a binary classification model using Azure Machine Learning Studio.
You must use a Receiver Operating Characteristic (RO C) curve and an F1 score to evaluate the model.
You need to create the required business metrics.
How should you complete the experiment? To answer, select the appropriate options in the dialog box in the answer area.
NOTE: Each correct selection is worth one point.
Answer:

NEW QUESTION: 24
You create an experiment in Azure Machine Learning Studio- You add a training dataset that contains 10.000 rows. The first 9.000 rows represent class 0 (90 percent). The first 1.000 rows represent class 1 (10 percent).
The training set is unbalanced between two Classes. You must increase the number of training examples for class 1 to 4,000 by using data rows. You add the Synthetic Minority Oversampling Technique (SMOTE) module to the experiment.
You need to configure the module.
Which values should you use? To answer, select the appropriate options in the dialog box in the answer area.
NOTE: Each correct selection is worth one point.
Answer:

NEW QUESTION: 25
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. Replace the Web Service Input module with a data input that does not include the price column.
B. Remove the Apply SQL Transformation module from the data flow.
C. Add a Select Columns module before the Score Model module to select all columns other than price.
D. Connect the output of the Apply SQL Transformation to the Web Service Output module.
E. Replace the training dataset module with a data input that does not include the price column.
F. Remove the Apply Math Operation module that replaces price with its natural log from the data flow.
Answer: C,D,F (LEAVE A REPLY)

NEW QUESTION: 26
You have an existing GitHub repository containing Azure Machine Learning project files.
You need to clone the repository to your Azure Machine Learning shared workspace file system.
Which four actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.
NOTE: More than one order of answer choices is correct. You will receive credit for any of the correct orders you select.
Answer:
1 - Fro the terminal window in the Azure Machine Learning interface, run the git clone command.
2 - From the terminal window in the Azure Machine Learning interface, run the cat~/.ssh/id_rsa.pub command.
3 - From the terminal window in the Azure Machine Learning interface, run the ssh-keygen command.
4 - Add a public key to the GitHub account.

NEW QUESTION: 27
You use the Azure Machine Learning SDK to run a training experiment that trains a classification model and calculates its accuracy metric.
The model will be retrained each month as new data is available.
You must register the model for use in a batch inference pipeline.
You need to register the model and ensure that the models created by subsequent retraining experiments are registered only if their accuracy is higher than the currently registered model.
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. Specify a different name for the model each time you register it.
B. Register the model with the same name each time regardless of accuracy, and always use the latest version of the model in the batch inferencing pipeline.
C. Specify the model framework version when registering the model, and only register subsequent models if this value is higher.
D. Specify a property named accuracy with the accuracy metric as a value when registering the model, and only register subsequent models if their accuracy is higher than the accuracy property value of the currently registered model.

E. Specify a tag named accuracy with the accuracy metric as a value when registering the model, and only register subsequent models if their accuracy is higher than the accuracy tag value of the currently registered model.

Answer: C,E (LEAVE A REPLY)

Explanation

E: Using tags, you can track useful information such as the name and version of the machine learning library used to train the model. Note that tags must be alphanumeric.

Reference:

https://notebooks.azure.com/xavierheriat/projects/azureml-getting-started/html/how-to-use-azureml/deployment/

NEW QUESTION: 28

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

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

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

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

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

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

Answer: (SHOW ANSWER)

Explanation

Log a numerical or string value to the run with the given name using log(name, value, description=''). Logging a metric to a run causes that metric to be stored in the run record in the experiment. You can log the same metric multiple times within a run, the result being considered a vector of that metric.

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

Reference:

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

NEW QUESTION: 29

You publish a batch inferencing pipeline that will be used by a business application.

The application developers need to know which information should be submitted to and returned by the REST interface for the published pipeline.

You need to identify the information required in the REST request and returned as a response from the published pipeline.

Which values should you use in the REST request and to expect in the response? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer:

Explanation

Box 1: JSON containing an OAuth bearer token

Specify your authentication header in the request.

To run the pipeline from the REST endpoint, you need an OAuth2 Bearer-type authentication header.

Box 2: JSON containing the experiment name

Add a JSON payload object that has the experiment name.

Example:

```
rest_endpoint = published_pipeline.endpoint
response = requests.post(rest_endpoint,
```

headers=auth_header,
json={"ExperimentName": "batch_scoring",
"ParameterAssignments": {"process_count_per_node": 6}})
run_id = response.json()["Id"]
Box 3: JSON containing the run ID
Make the request to trigger the run. Include code to access the Id key from the response dictionary to get the value of the run ID.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/tutorial-pipeline-batch-scoring-classification

NEW QUESTION: 30
You train and register a model by using the Azure Machine Learning SDK on a local workstation. Python 3.6 and Visual Studio Code are installed on the workstation.
When you try to deploy the model into production as an Azure Kubernetes Service (AKS)-based web service, you experience an error in the scoring script that causes deployment to fail.
You need to debug the service on the local workstation before deploying the service to production.
Which four actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.
Answer:
Explanation
Graphical user interface, text, application, email Description automatically generated
Step 1: Install Docker on the workstation
Prerequisites include having a working Docker installation on your local system.
Build or download the dockerfile to the compute node.
Step 2: Create an AksWebservice deployment configuration and deploy the model to it To deploy a model to Azure Kubernetes Service, create a deployment configuration that describes the compute resources needed.
# 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) service.wait_for_deployment(show_output = True) print(service.state) print(service.get_logs()) Step 3: Create a LocalWebservice deployment configuration for the service and deploy the model to it To deploy locally, modify your code to use LocalWebservice.deploy_configuration() to create a deployment configuration. Then use Model.deploy() to deploy the service.
Step 4: Debug and modify the scoring script as necessary. Use the reload() method of the service after each modification.
During local testing, you may need to update the score.py file to add logging or attempt to resolve any problems that you've discovered. To reload changes to the score.py file, use reload(). For example, the following code reloads the script for the service, and then sends data to it.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-deploy-azure-kubernetes-service
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-troubleshoot-deployment-local

NEW QUESTION: 31
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.
B.
C.
D.
Answer: (SHOW ANSWER)
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.MachineLearningServices/workspaces/<workspace_name>"
]
}
```

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

NEW QUESTION: 32
You are using the Hyperdrive feature in Azure Machine Learning to train a model.
You configure the Hyperdrive experiment by running the following code:
For each of the following statements, select Yes if the statement is true. Otherwise, select No.
NOTE: Each correct selection is worth one point.
Answer:
Explanation
Box 1: Yes
In random sampling, hyperparameter values are randomly selected from the defined search space. Random sampling allows the search space to include both discrete and continuous hyperparameters.
Box 2: Yes
learning_rate has a normal distribution with mean value 10 and a standard deviation of 3.
Box 3: No
keep_probability has a uniform distribution with a minimum value of 0.05 and a maximum value of 0.1.
Box 4: No
number_of_hidden_layers takes on one of the values [3, 4, 5].
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-tune-hyperparameters

NEW QUESTION: 33
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.
Answer:
Reference:
https://developers.google.com/machine-learning/crash-course/classification/true-false-positive-negative

NEW QUESTION: 34
You are creating a machine learning model that can predict the species of a penguin from its measurements.
You have a file that contains measurements for free species of penguin in comma delimited format.
The model must be optimized for area under the received operating characteristic curve performance metric averaged for each class.
You need to use the Automated Machine Learning user interface in Azure Machine Learning studio to run an experiment and find the best performing model.

Which five 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 collect order.

Answer:

Explanation

NEW QUESTION: 35

You need to produce a visualization for the diagnostic test evaluation according to the data visualization requirements.

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

Answer:

Explanation

Step 1: Sweep Clustering

Start by using the "Tune Model Hyperparameters" module to select the best sets of parameters for each of the models we're considering.

One of the interesting things about the "Tune Model Hyperparameters" module is that it not only outputs the results from the Tuning, it also outputs the Trained Model.

Step 2: Train Model

Step 3: Evaluate Model

Scenario: You need to provide the test results to the Fabrikam Residences team. You create data visualizations to aid in presenting the results.

You must produce a Receiver Operating Characteristic (ROC) curve to conduct a diagnostic test evaluation of the model. You need to select appropriate methods for producing the ROC curve in Azure Machine Learning Studio to compare the Two-Class Decision Forest and the Two-Class Decision Jungle modules with one another.

References:

http://breaking-bi.blogspot.com/2017/01/azure-machine-learning-model-evaluation.html

NEW QUESTION: 36

You arc I mating a deep learning model to identify cats and dogs. You have 25,000 color images.

You must meet the following requirements:

* Reduce the number of training epochs.
* Reduce the size of the neural network.
* Reduce over-fitting of the neural network.

You need to select the image modification values.

Which value should you use? To answer, select the appropriate Options in the answer area.

NOTE: Each correct selection is worth one point.

Answer:

NEW QUESTION: 37

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

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

You need to train the model.

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

Answer:

Explanation

NEW QUESTION: 38

You have a dataset that includes confidential data. You use the dataset to train a model.

You must use a differential privacy parameter to keep the data of individuals safe and private.

You need to reduce the effect of user data on aggregated results.

What should you do?

A. Decrease the value of the epsilon parameter to reduce the amount of noise added to the data

B. Increase the value of the epsilon parameter to decrease privacy and increase accuracy

C. Decrease the value of the epsilon parameter to increase privacy and reduce accuracy

D. Set the value of the epsilon parameter to 1 to ensure maximum privacy

Answer: C (LEAVE A REPLY)

Differential privacy tries to protect against the possibility that a user can produce an indefinite number of reports to eventually reveal sensitive data. A value known as epsilon measures how noisy, or private, a report is. Epsilon has an inverse relationship to noise or privacy. The lower the epsilon, the more noisy (and private) the data is.

Reference:

https://docs.microsoft.com/en-us/azure/machine-learning/concept-differential-privacy

NEW QUESTION: 39

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:

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

NOTE: Each correct selection is worth one point.

Answer:

Explanation

Box 1: 0.859122

Box 2: a positively linear relationship

+1 indicates a strong positive linear relationship

-1 indicates a strong negative linear correlation

0 denotes no linear relationship between the two variables.

References:

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

NEW QUESTION: 40

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:
The IT department creates an Azure Kubernetes Service (AKS)-based inference compute target named aks-cluster in the Azure Machine Learning workspace.
You have a Microsoft Surface Book computer with a GPU. Python 3.6 and Visual Studio Code are installed.
You need to run a script that trains a deep neural network (DNN) model and logs the loss and accuracy metrics.
Solution: Install the Azure ML SDK on the Surface Book. Run Python code to connect to the workspace and then run the training script as an experiment on local compute.
A. Yes
B. No
Answer: (SHOW ANSWER)
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: 41
You plan to run a Python script as an Azure Machine Learning experiment.
The script contains the following code:
```
import os, argparse, glob
from azureml.core import Run
parser = argparse.ArgumentParser()
parser.add_argument('--input-data',
type=str, dest='data_folder')
args = parser.parse_args()
data_path = args.data_folder
file_paths = glob.glob(data_path + "/*.jpg")
```
You must specify a file dataset as an input to the script. The dataset consists of multiple large image files and must be streamed directly from its source.
You need to write code to define a ScriptRunConfig object for the experiment and pass the ds dataset as an argument.
Which code segment should you use?
A. arguments = ['--input-data', ds.to_pandas_dataframe()]
B. arguments = ['--input-data', ds.as_mount()]
C. arguments = ['--data-data', ds]
D. arguments = ['--input-data', ds.as_download()]
Answer: A (LEAVE A REPLY)
Explanation
If you have structured data not yet registered as a dataset, create a TabularDataset and use it directly in your training script for your local or remote experiment.
To load the TabularDataset to pandas DataFrame
df = dataset.to_pandas_dataframe()
Note: TabularDataset represents data in a tabular format created by parsing the provided file or list of files.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-train-with-datasets

NEW QUESTION: 42
You need to implement a new cost factor scenario for the ad response models as illustrated in the performance curve exhibit.
Which technique should you use?
A. Set the threshold to 0.5 and retrain if weighted Kappa deviates +/- 5% from 0.45.
B. Set the threshold to 0.05 and retrain if weighted Kappa deviates +/- 5% from 0.5.
C. Set the threshold to 0.2 and retrain if weighted Kappa deviates +/- 5% from 0.6.
D. Set the threshold to 0.75 and retrain if weighted Kappa deviates +/- 5% from 0.15.
Answer: (SHOW ANSWER)
Scenario:
Performance curves of current and proposed cost factor scenarios are shown in the following diagram:
The ad propensity model uses a cut threshold is 0.45 and retrains occur if weighted Kappa deviated from 0.1
+/- 5%.

NEW QUESTION: 43
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.
Answer:
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: 44
You create an Azure Machine Learning workspace. You use the Azure Machine Learning SDK for Python.
You must create a dataset from remote paths. The dataset must be reusable within the workspace.
You need to create the dataset.
How should you complete the following code segment? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.
Answer:

NEW QUESTION: 45
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
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:
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: 46
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
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:

NEW QUESTION: 47
You use Azure Machine Learning Studio to build a machine learning experiment.
You need to divide data into two distinct datasets.
Which module should you use?
A. Partition and Sample
B. Assign Data to Clusters
C. Group Data into Bins
D. Test Hypothesis Using t-Test
Answer: (SHOW ANSWER)
Partition and Sample with the Stratified split option outputs multiple datasets, partitioned using the rules you specified.
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/partition-and-sample

NEW QUESTION: 48
You need to implement source control for scripts in an Azure Machine Learning workspace. You use a terminal window in the Azure Machine Learning Notebook tab You must authenticate your Git account with SSH.
You need to generate a new SSH key.
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 m the correct order.
Answer:
Explanation

NEW QUESTION: 49
You are creating a binary classification by using a two-class logistic regression model.
You need to evaluate the model results for imbalance.
Which evaluation metric should you use?
A. Relative Absolute Error
B. AUC Curve
C. Mean Absolute Error
D. Relative Squared Error
Answer: (SHOW ANSWER)
One can inspect the true positive rate vs. the false positive rate in the Receiver Operating Characteristic (ROC) curve and the corresponding Area Under the Curve (AUC) value. The closer this curve is to the upper left corner, the better the classifier's performance is (that is maximizing the true positive rate while minimizing the false positive rate). Curves that are close to the diagonal of the plot, result from classifiers that tend to make predictions that are close to random guessing.
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio/evaluate-model-performance#evaluating-a-binary-classification-model

NEW QUESTION: 50
You are training machine learning models in Azure Machine Learning. You use Hyperdrive to tune the hyperparameter.
In previous model training and tuning runs, many models showed similar performance.
You need to select an early termination policy that meets the following requirements:
* accounts for the performance of all previous runs when evaluating the current run
* avoids comparing the current run with only the best performing run to date Which two early termination policies should you use? Each correct answer presents part of the solution.
NOTE: Each correct selection is worth one point.
A. Median stopping
B. Bandit
C. Default
D. Truncation selection
Answer: A,C (LEAVE A REPLY)
The Median Stopping policy computes running averages across all runs and cancels runs whose best performance is worse than the median of the running averages.
If no policy is specified, the hyperparameter tuning service will let all training runs execute to completion.
Incorrect Answers:
B: BanditPolicy defines an early termination policy based on slack criteria, and a frequency and delay interval for evaluation.
The Bandit policy takes the following configuration parameters:
slack_factor: The amount of slack allowed with respect to the best performing training run. This factor specifies the slack as a ratio.
D: The Truncation selection policy periodically cancels the given percentage of runs that rank the lowest for their performance on the primary metric. The policy strives for fairness in ranking the runs by accounting for improving model performance with training time. When ranking a relatively young run, the policy uses the corresponding (and earlier) performance of older runs for comparison. Therefore, runs aren't terminated for having a lower performance because they have run for less time than other runs.
Reference:

https://docs.microsoft.com/en-us/python/api/azureml-train-core/azureml.train.hyperdrive.medianstoppingpolicy
https://docs.microsoft.com/en-us/python/api/azureml-train-core/
azureml.train.hyperdrive.truncationselectionpolicy
https://docs.microsoft.com/en-us/python/api/azureml-train-core/azureml.train.hyperdrive.banditpolicy

NEW QUESTION: 51
You plan to use automated machine learning to train a regression model. You have data that has features which have missing values, and categorical features with few distinct values.
You need to configure automated machine learning to automatically impute missing values and encode categorical features as part of the training task.
Which parameter and value pair should you use in the AutoMLConfig class?
A. featurization = 'auto'
B. enable_voting_ensemble = True
C. task = 'classification'
D. exclude_nan_labels = True
E. enable_tf = True
Answer: A (LEAVE A REPLY)
Featurization str or FeaturizationConfig
Values: 'auto' / 'off' / FeaturizationConfig
Indicator for whether featurization step should be done automatically or not, or whether customized featurization should be used.
Column type is automatically detected. Based on the detected column type preprocessing/featurization is done as follows:
Categorical: Target encoding, one hot encoding, drop high cardinality categories, impute missing values.
Numeric: Impute missing values, cluster distance, weight of evidence.
DateTime: Several features such as day, seconds, minutes, hours etc.
Text: Bag of words, pre-trained Word embedding, text target encoding.
Reference:
https://docs.microsoft.com/en-us/python/api/azureml-train-automl-client/azureml.train.automl.automlconfig.auto

NEW QUESTION: 52
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:
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:
Does the solution meet the goal?
A. Yes
B. No
Answer: A (LEAVE A REPLY)
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: 53
You plan to build a team data science environment. Data for training models in machine learning pipelines will be over 20 GB in size.
You have the following requirements:
Models must be built using Caffe2 or Chainer frameworks.
Data scientists must be able to use a data science environment to build the machine learning pipelines
and train models on their personal devices in both connected and disconnected network environments.
Personal devices must support updating machine learning pipelines when connected to a network.
You need to select a data science environment.
Which environment should you use?
A. Azure Machine Learning Service
B. Azure Machine Learning Studio
C. Azure Databricks
D. Azure Kubernetes Service (AKS)
Answer: A (LEAVE A REPLY)
Explanation/Reference:
Explanation:
The Data Science Virtual Machine (DSVM) is a customized VM image on Microsoft's Azure cloud built specifically for doing data science. Caffe2 and Chainer are supported by DSVM.
DSVM integrates with Azure Machine Learning.
Incorrect Answers:
B: Use Machine Learning Studio when you want to experiment with machine learning models quickly and easily, and the built-in machine learning algorithms are sufficient for your solutions.
References:
https://docs.microsoft.com/en-us/azure/machine-learning/data-science-virtual-machine/overview

NEW QUESTION: 54
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: B,C,D (LEAVE A REPLY)

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: 55
You need to correct the model fit issue.
Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.
Answer:
Explanation
Step 1: Augment the data
Scenario: Columns in each dataset contain missing and null values. The datasets also contain many outliers.
Step 2: Add the Bayesian Linear Regression module.
Scenario: You produce a regression model to predict property prices by using the Linear Regression and Bayesian Linear Regression modules.
Step 3: Configure the regularization weight.
Regularization typically is used to avoid overfitting. For example, in L2 regularization weight, type the value to use as the weight for L2 regularization. We recommend that you use a non-zero value to avoid overfitting.
Scenario:
Model fit: The model shows signs of overfitting. You need to produce a more refined regression model that reduces the overfitting.

NEW QUESTION: 56
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.
After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.
You are a data scientist using Azure Machine Learning Studio.
You need to normalize values to produce an output column into bins to predict a target column.
Solution: Apply a Quantiles binning mode with a PQuantile normalization.
Does the solution meet the goal?
A. Yes
B. No
Answer: B (LEAVE A REPLY)
Use the Entropy MDL binning mode which has a target column.
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/group-data-into-bins

NEW QUESTION: 57
You plan to preprocess text from CSV files. You load the Azure Machine Learning Studio default stop words list.

You need to configure the Preprocess Text module to meet the following requirements:
* Ensure that multiple related words from a single canonical form.
* Remove pipe characters from text.
* Remove words to optimize information retrieval.
Which three options should you select? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.
Answer:
Explanation
Box 1: Remove stop words
Remove words to optimize information retrieval.
Remove stop words: Select this option if you want to apply a predefined stopword list to the text column. Stop word removal is performed before any other processes.
Box 2: Lemmatization
Ensure that multiple related words from a single canonical form.
Lemmatization converts multiple related words to a single canonical form Box 3: Remove special characters Remove special characters: Use this option to replace any non-alphanumeric special characters with the pipe | character.
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/preprocess-text

NEW QUESTION: 58
You create an experiment in Azure Machine Learning Studio- You add a training dataset that contains 10.000 rows. The first 9.000 rows represent class 0 (90 percent). The first 1.000 rows represent class 1 (10 percent).
The training set is unbalanced between two Classes. You must increase the number of training examples for class 1 to 4,000 by using data rows. You add the Synthetic Minority Oversampling Technique (SMOTE) module to the experiment.
You need to configure the module.
Which values should you use? To answer, select the appropriate options in the dialog box in the answer area.
NOTE: Each correct selection is worth one point.
Answer:

NEW QUESTION: 59
You have a multi-class image classification deep learning model that uses a set of labeled photographs. You create the following code to select hyperparameter values when training the model.
For each of the following statements, select Yes if the statement is true. Otherwise, select No.
NOTE: Each correct selection is worth one point.
Answer:
Explanation:
Box 1: Yes
Hyperparameters are adjustable parameters you choose to train a model that govern the training process itself. Azure Machine Learning allows you to automate hyperparameter exploration in an efficient manner, saving you significant time and resources. You specify the range of hyperparameter values and a maximum number of training runs. The system then automatically launches multiple simultaneous runs with different parameter configurations and

finds the configuration that results in the best performance, measured by the metric you choose. Poorly performing training runs are automatically early terminated, reducing wastage of compute resources. These resources are instead used to explore other hyperparameter configurations.
Box 2: Yes
uniform(low, high) - Returns a value uniformly distributed between low and high Box 3: No Bayesian sampling does not currently support any early termination policy.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-tune-hyperparameters

NEW QUESTION: 60
You register the following versions of a model.
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:
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:
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-deploy-and-where

NEW QUESTION: 61
You create a multi-class image classification deep learning experiment by using the PyTorch framework. You plan to run the experiment on an Azure Compute cluster that has nodes with GPU's.
You need to define an Azure Machine Learning service pipeline to perform the monthly retraining of the image classification model. The pipeline must run with minimal cost and minimize the time required to train the model.
Which three pipeline steps should you run in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.
Answer:
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: 62
You are tuning a hyperparameter for an algorithm. The following table shows a data set with different hyperparameter, training error, and validation errors.
Use the drop-down menus to select the answer choice that answers each question based on the information presented in the graphic.
Answer:
Reference:
https://medium.com/comet-ml/organizing-machine-learning-projects-project-management-guidelines-2d2b85651bbd

NEW QUESTION: 63
You need to define an evaluation strategy for the crowd sentiment models.
Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.
Answer:
Explanation:
Step 1: Define a cross-entropy function activation
When using a neural network to perform classification and prediction, it is usually better to use cross-entropy error than classification error, and somewhat better to use cross-entropy error than mean squared error to evaluate the quality of the neural network.
Step 2: Add cost functions for each target state.
Step 3: Evaluated the distance error metric.
References:
https://www.analyticsvidhya.com/blog/2018/04/fundamentals-deep-learning-regularization-techniques/

NEW QUESTION: 64
You deploy a model as an Azure Machine Learning real-time web service using the following code.
The deployment fails.
You need to troubleshoot the deployment failure by determining the actions that were performed during deployment and identifying the specific action that failed.
Which code segment should you run?
A. service.get_logs()
B. service.state
C. service.serialize()
D. service.update_deployment_state()
Answer: A (LEAVE A REPLY)
Explanation
You can print out detailed Docker engine log messages from the service object. You can view the log for ACI, AKS, and Local deployments. The following example demonstrates how to print the logs.
# if you already have the service object handy
print(service.get_logs())
# if you only know the name of the service (note there might be multiple services with the same name but different version number)
print(ws.webservices['mysvc'].get_logs()) Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-troubleshoot-deployment

NEW QUESTION: 65
You have a dataset that includes home sales data for a city. The dataset includes the following columns.
Each row in the dataset corresponds to an individual home sales transaction.
You need to use automated machine learning to generate the best model for predicting the sales price based on the features of the house.
Which values should you use? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.
Answer:
Explanation:
Box 1: Regression
Regression is a supervised machine learning technique used to predict numeric values.
Box 2: Price
Reference:
https://docs.microsoft.com/en-us/learn/modules/create-regression-model-azure-machine-learning-designer

NEW QUESTION: 66
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:
Does the solution meet the goal?
A. Yes
B. No
Answer: A (LEAVE A REPLY)
The scikit-learn estimator provides a simple way of launching a scikit-learn training job on a compute target. It is implemented through the SKLearn class, which can be used to support single-node CPU training.
Example:
from azureml.train.sklearn import SKLearn
}
estimator = SKLearn(source_directory=project_folder,
compute_target=compute_target,
entry_script='train_iris.py'
)
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-train-scikit-learn

NEW QUESTION: 67
You are creating an experiment by using Azure Machine Learning Studio.
You must divide the data into four subsets for evaluation. There is a high degree of missing values in the data.
You must prepare the data for analysis.
You need to select appropriate methods for producing the experiment.
Which three modules 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.
NOTE: More than one order of answer choices is correct. You will receive credit for any of the correct orders you select.
Answer:
Explanation
The Clean Missing Data module in Azure Machine Learning Studio, to remove, replace, or infer missing values.

NEW QUESTION: 68
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.
After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.
You use Azure Machine Learning designer to load the following datasets into an experiment:
You need to create a dataset that has the same columns and header row as the input datasets and contains all rows from both input datasets.
Solution: Use the Execute Python Script module.
Does the solution meet the goal?
A. No
B. Yes
Answer: A (LEAVE A REPLY)

NEW QUESTION: 69
You are authoring a notebook in Azure Machine Learning studio.
You must install packages from the notebook into the currently running kernel. The installation must be limited to the currently running kernel only.
You need to install the packages.
Which magic function should you use?
A. !conda
B. !pjp
C. %load
D. %pip
Answer: D (LEAVE A REPLY)

NEW QUESTION: 70
You create an Azure Machine Learning workspace.
You need to detect data drift between a baseline dataset and a subsequent target dataset by using the DataDriftDetector class.
How should you complete the code segment? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.
Answer:

Reference:
https://docs.microsoft.com/en-us/python/api/azureml-datadrift/azureml.datadrift.datadriftdetector(class)

NEW QUESTION: 71
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 contain 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: Use the last Observation Carried Forward (IOCF) method to impute the missing data points.
Does the solution meet the goal?
A. Yes
B. No
Answer: B (LEAVE A REPLY)

NEW QUESTION: 72
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:
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. No
B. Yes
Answer: B (LEAVE A REPLY)

NEW QUESTION: 73
You have a dataset that contains 2,000 rows. You are building a machine learning classification model by using Azure Learning Studio. You add a Partition and Sample module to the experiment.
You need to configure the module. You must meet the following requirements:
Divide the data into subsets
Assign the rows into folds using a round-robin method
Allow rows in the dataset to be reused
How should you configure the module? To answer, select the appropriate options in the dialog box in the answer area.
NOTE: Each correct selection is worth one point.

Answer:
Explanation
Use the Split data into partitions option when you want to divide the dataset into subsets of the data. This option is also useful when you want to create a custom number of folds for cross-validation, or to split rows into several groups.
Add the Partition and Sample module to your experiment in Studio (classic), and connect the dataset.
For Partition or sample mode, select Assign to Folds.
Use replacement in the partitioning: Select this option if you want the sampled row to be put back into the pool of rows for potential reuse. As a result, the same row might be assigned to several folds.
If you do not use replacement (the default option), the sampled row is not put back into the pool of rows for potential reuse. As a result, each row can be assigned to only one fold.
Randomized split: Select this option if you want rows to be randomly assigned to folds.
If you do not select this option, rows are assigned to folds using the round-robin method.
References:
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/partition-and-sample

NEW QUESTION: 74
You are evaluating a completed binary classification machine learning model.
You need to use the precision as the evaluation metric.
Which visualization should you use?
A. Violin plot
B. Gradient descent
C. Box plot
D. Binary classification confusion matrix
Answer: D (LEAVE A REPLY)
Incorrect Answers:
A: A violin plot is a visual that traditionally combines a box plot and a kernel density plot.
B: Gradient descent is a first-order iterative optimization algorithm for finding the minimum of a function. To find a local minimum of a function using gradient descent, one takes steps proportional to the negative of the gradient (or approximate gradient) of the function at the current point.
C: A box plot lets you see basic distribution information about your data, such as median, mean, range and quartiles but doesn't show you how your data looks throughout its range.
References:
https://machinelearningknowledge.ai/confusion-matrix-and-performance-metrics-machine-learning/

NEW QUESTION: 75
You are analyzing a dataset by using Azure Machine Learning Studio.
YOU need to generate a statistical summary that contains the p value and the unique value count for each feature column.
Which two modules can you users? Each correct answer presents a complete solution.
NOTE: Each correct selection is worth one point.
A. Compute linear Correlation
B. Execute Python Script
C. Convert to Indicator Values

D. Summarize Data
E. Export Count Table
Answer: A,E (LEAVE A REPLY)

NEW QUESTION: 76
You are producing a multiple linear regression model in Azure Machine Learning Studio.
Several independent variables are highly correlated.
You need to select appropriate methods for conducting effective feature engineering on all the data.
Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.
Answer:
Explanation:
Step 1: Use the Filter Based Feature Selection module
Filter Based Feature Selection identifies the features in a dataset with the greatest predictive power.
The module outputs a dataset that contains the best feature columns, as ranked by predictive power. It also outputs the names of the features and their scores from the selected metric.
Step 2: Build a counting transform
A counting transform creates a transformation that turns count tables into features, so that you can apply the transformation to multiple datasets.
Step 3: Test the hypothesis using t-Test
References:
https://docs.microsoft.com/bs-latn-ba/azure/machine-learning/studio-module-reference/filter-based-feature-selection
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/build-counting-transform

NEW QUESTION: 77
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:
Does the solution meet the goal?
A. Yes
B. No

Answer: A (LEAVE A REPLY)
Explanation
The two steps are present: process_step and train_step
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.
PipelineData objects are also used when constructing Pipelines to describe step dependencies. To specify that a step requires the output of another step as input, use a PipelineData object in the constructor of both steps.
For 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: 78
You create an Azure Machine Learning workspace and install the MLflow library.
You need to tog different types of data by using the MLflow library.
Which method should you use? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.
Answer:

NEW QUESTION: 79
You create an Azure Machine Learning workspace named workspaces. You create a Python SDK v2 notebook to perform custom model training in workspace1. You need to run the notebook from Azure Machine Learning Studio in workspace1. What should you provision first?
A. default storage account
B. Azure Machine Learning compute cluster
C. Azure Machine Learning compute instance
D. real-time endpoint
Answer: C (LEAVE A REPLY)

NEW QUESTION: 80
You need to define an evaluation strategy for the crowd sentiment models.
Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.
Answer:
1 - Add new features for retraining supervised models.
2 - Evaluate the changes in correlation between...

3 - Filter labeled cases for retraining using...
Reference:
https://en.wikipedia.org/wiki/Nearest_centroid_classifier
https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/sweep-clustering

NEW QUESTION: 81
You are using hyperparameter tuning in Azure Machine Learning Python SDK v2 to train a model. You configure the hyperparameter tuning experiment by running the following code:
For each of the following statements select Yes if the statement is true. Otherwise, select No. NOTE: Fach correct selection is worth one paint.
Answer:

NEW QUESTION: 82
You create a new Azure Machine Learning workspace with a compute cluster.
You need to create the compute cluster asynchronously by using the Azure Machine Learning Python SDK v2.
How should you complete the code segment? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point
Answer:
Explanation:

NEW QUESTION: 83
A coworker registers a datastore in a Machine Learning services workspace by using the following code:
You need to write code to access the datastore from a notebook.
Answer:
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: 84
You deploy a model in Azure Container Instance.
You must use the Azure Machine Learning SDK to call the model API.
You need to invoke the deployed model using native SDK classes and methods.
How should you complete the command? To answer, select the appropriate options in the answer areas.
NOTE: Each correct selection is worth one point.
Answer:

Explanation
Box 1: from azureml.core.webservice import Webservice
The following code shows how to use the SDK to update the model, environment, and entry script for a web service to Azure Container Instances:
from azureml.core import Environment
from azureml.core.webservice import Webservice
from azureml.core.model import Model, InferenceConfig
Box 2: predictions = service.run(input_json)
Example: The following code demonstrates sending data to the service:
import json
test_sample = json.dumps({'data': [
[1, 2, 3, 4, 5, 6, 7, 8, 9, 10],
[10, 9, 8, 7, 6, 5, 4, 3, 2, 1]
]})
test_sample = bytes(test_sample, encoding='utf8')
prediction = service.run(input_data=test_sample)
print(prediction)
Reference:
https://docs.microsoft.com/bs-latn-ba/azure/machine-learning/how-to-deploy-azure-container-instance
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-troubleshoot-deployment

NEW QUESTION: 85
You have an Azure Machine Learning workspace.
You plan to tune a model hyperparameter when you train the model.
You need to define a search space that returns a normally distributed value.
Which parameter should you use?
A. LogUniform
B. Uniform
C. QUniform
D. LogNormal
Answer: (SHOW ANSWER)

NEW QUESTION: 86
You are retrieving data from a large datastore by using Azure Machine Learning Studio.
You must create a subset of the data for testing purposes using a random sampling seed based on the system clock.
You add the Partition and Sample module to your experiment.
You need to select the properties for the module.
Which values should you select? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.
Answer:
Explanation
Box 1: Sampling

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

NEW QUESTION: 87
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:
The IT department creates an Azure Kubernetes Service (AKS)-based inference compute target named aks- cluster in the Azure Machine Learning workspace.
You have a Microsoft Surface Book computer with a GPU. Python 3.6 and Visual Studio Code are installed.
You need to run a script that trains a deep neural network (DNN) model and logs the loss and accuracy metrics.
Solution: Install the Azure ML SDK on the Surface Book. Run Python code to connect to the workspace and then run the training script as an experiment on local compute.
Does the solution meet the goal?
A. Yes
B. No
Answer: B (LEAVE A REPLY)
Need to attach the mlvm virtual machine as a compute target in the Azure Machine Learning workspace.
Reference:
https://docs.microsoft.com/en-us/azure/machine-learning/concept-compute-target

NEW QUESTION: 88
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.
Answer:
Reference:

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

NEW QUESTION: 89
You need to build a feature extraction strategy for the local models.
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:

NEW QUESTION: 90
You need to use the Python language to build a sampling strategy for the global penalty detection models.
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:
Reference:
https://github.com/pytorch/pytorch/blob/master/torch/utils/data/distributed.py
Topic 2, Overview
Datasets
There are two datasets in CSV format that contain property details for two cities, London and Paris, with the following columns:
The two datasets have been added to Azure Machine Learning Studio as separate datasets and included as the starting point of the experiment.
Dataset issues
The AccessibilityToHighway column in both datasets contains missing values. The missing data must be replaced with new data so that it is modeled conditionally using the other variables in the data before filling in the missing values.
Columns in each dataset contain missing and null values. The dataset also contains many outliers. The Age column has a high proportion of outliers. You need to remove the rows that have outliers in the Age column. The MedianValue and AvgRoomsinHouse columns both hold data in numeric format. You need to select a feature selection algorithm to analyze the relationship between the two columns in more detail.
Model fit
The model shows signs of overfitting. You need to produce a more refined regression model that reduces the overfitting.
Experiment Requirements
You must set up the experiment to cross-validate the Linear Regression and Bayesian Linear Regression modules to evaluate performance.
In each case, the predictor of the dataset is the column named MedianValue. An initial investigation showed that the datasets are identical in structure apart from the MedianValue column. The smaller Paris dataset contains the MedianValue in text format, whereas the larger London dataset contains the MedianValue in numerical format. You must ensure that the datatype of the MedianValue column of the Paris dataset matches the structure of the London dataset.
You must prioritize the columns of data for predicting the outcome. You must use non-parameters statistics to measure the relationships.
You must use a feature selection algorithm to analyze the relationship between the MedianValue and AvgRoomsinHouse columns.
Model training
Given a trained model and a test dataset, you need to compute the permutation feature importance scores of feature variables. You need to set up the Permutation Feature Importance module to select the correct metric to investigate the model's accuracy and replicate the findings.
You want to configure hyperparameters in the model learning process to speed the learning phase by using hyperparameters. In addition, this configuration should cancel the lowest performing runs at each evaluation interval, thereby directing effort and resources towards models that are more likely to be successful.

You are concerned that the model might not efficiently use compute resources in hyperparameter tuning. You also are concerned that the model might prevent an increase in the overall tuning time. Therefore, you need to implement an early stopping criterion on models that provides savings without terminating promising jobs.
Testing
You must produce multiple partitions of a dataset based on sampling using the Partition and Sample module in Azure Machine Learning Studio. You must create three equal partitions for cross-validation. You must also configure the cross-validation process so that the rows in the test and training datasets are divided evenly by properties that are near each city's main river. The data that identifies that a property is near a river is held in the column named NextToRiver. You want to complete this task before the data goes through the sampling process.
When you train a Linear Regression module using a property dataset that shows data for property prices for a large city, you need to determine the best features to use in a model. You can choose standard metrics provided to measure performance before and after the feature importance process completes. You must ensure that the distribution of the features across multiple training models is consistent.
Data visualization
You need to provide the test results to the Fabrikam Residences team. You create data visualizations to aid in presenting the results.
You must produce a Receiver Operating Characteristic (ROC) curve to conduct a diagnostic test evaluation of the model. You need to select appropriate methods for producing the ROC curve in Azure Machine Learning Studio to compare the Two-Class Decision Forest and the Two-Class Decision Jungle modules with one another.

NEW QUESTION: 91
You are building an experiment using the Azure Machine Learning designer.
You split a dataset into training and testing sets. You select the Two-Class Boosted Decision Tree as the algorithm.
You need to determine the Area Under the Curve (AUC) of the model.
Which three modules should you use in sequence? To answer, move the appropriate modules from the list of modules to the answer area and arrange them in the correct order.
Answer:
Explanation:
Step 1: Train Model
Two-Class Boosted Decision Tree
First, set up the boosted decision tree model.
1. Find the Two-Class Boosted Decision Tree module in the module palette and drag it onto the canvas.
2. Find the Train Model module, drag it onto the canvas, and then connect the output of the Two-Class Boosted Decision Tree module to the left input port of the Train Model module.
The Two-Class Boosted Decision Tree module initializes the generic model, and Train Model uses training data to train the model.
3. Connect the left output of the left Execute R Script module to the right input port of the Train Model module (in this tutorial you used the data coming from the left side of the Split Data module for training).
This portion of the experiment now looks something like this:
Step 2: Score Model
Score and evaluate the models
You use the testing data that was separated out by the Split Data module to score our trained models. You can then compare the results of the two models to see which generated better results.
Add the Score Model modules
1. Find the Score Model module and drag it onto the canvas.

2. Connect the Train Model module that's connected to the Two-Class Boosted Decision Tree module to the left input port of the Score Model module.
3. Connect the right Execute R Script module (our testing data) to the right input port of the Score Model module.
Step 3: Evaluate Model
To evaluate the two scoring results and compare them, you use an Evaluate Model module.
1. Find the Evaluate Model module and drag it onto the canvas.
2. Connect the output port of the Score Model module associated with the boosted decision tree model to the left input port of the Evaluate Model module.
3. Connect the other Score Model module to the right input port.

NEW QUESTION: 92
You have an Azure Machine learning workspace. The workspace contains a dataset with data in a tabular form.
You plan to use the Azure Machine Learning SDK for Python vl to create a control script that will load the dataset into a pandas dataframe in preparation for model training The script will accept a parameter designating the dataset You need to complete the script.
How should you complete the script? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.
Answer:
Explanation

NEW QUESTION: 93
You have an existing GitHub repository containing Azure Machine Learning project files.
You need to clone the repository to your Azure Machine Learning shared workspace file system.
Which four actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.
NOTE: More than one order of answer choices is correct. You will receive credit for any of the correct orders you select.
Answer:
1 - From the terminal window in the Azure Machine Learning interface, run the git...
2 - From the terminal window in the Azure Machine Learning....
3 - From the terminal window in the Azure Machine Learning...
4 - Add a public key to the GitHub account.

NEW QUESTION: 94
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.

Answer:

1 - Add a Two-Class Support Vector Machine module to initialize the SVM classifier.
2 - Add a dataset to the experiment
3 - Add a Split Data module to create training and test dataset.
4 - Add a Permutation Feature Importance module and connect to the trained model and test dataset.
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

NEW QUESTION: 95

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.

Answer:

Explanation
Box 1: Attached compute
Box 2: Inference cluster
Box 3: Training cluster
Box 4: Attached compute

NEW QUESTION: 96

You need to set up the Permutation Feature Importance module according to the model training requirements.
Which properties should you select? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.

Answer:

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

NEW QUESTION: 97

You need to visually identify whether outliers exist in the Age column and quantify the outliers before the outliers are removed.
Which three Azure Machine Learning Studio modules should you use in sequence? To answer, move the appropriate modules from the list of modules to the answer area and arrange them in the correct order.

Answer:

Explanation

Create Scatterplot

Summarize Data

Clip Values

You can use the Clip Values module in Azure Machine Learning Studio, to identify and optionally replace data values that are above or below a specified threshold. This is useful when you want to remove outliers or replace them with a mean, a constant, or other substitute value.

References:

https://blogs.msdn.microsoft.com/azuredev/2017/05/27/data-cleansing-tools-in-azure-machine-learning/

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

NEW QUESTION: 98

You manage an Azure Machine Learning workspace. You submit a training job with the Azure Machine Learning Python SDK v2. You must use MLflow to log metrics, model parameters, and mode! artifacts automatically when training a model.

You start by writing the following code segment:

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

Answer:

Explanation

NEW QUESTION: 99

You are using Azure Machine Learning to train machine learning models. You need a compute target on which to remotely run the training script. You run the following Python code:

Answer:

NEW QUESTION: 100

You are developing deep learning models to analyze semi-structured, unstructured, and structured data types.

You have the following data available for model building:

Video recordings of sporting events

Transcripts of radio commentary about events

Logs from related social media feeds captured during sporting events

You need to select an environment for creating the model.

Which environment should you use?

A. Azure Cognitive Services

B. Azure Data Lake Analytics

C. Azure HDInsight with Spark MLib

D. Azure Machine Learning Studio

Answer: A (LEAVE A REPLY)

Explanation/Reference:
Explanation:
Azure Cognitive Services expand on Microsoft's evolving portfolio of machine learning APIs and enable developers to easily add cognitive features - such as emotion and video detection; facial, speech, and vision recognition; and speech and language understanding - into their applications. The goal of Azure Cognitive Services is to help developers create applications that can see, hear, speak, understand, and even begin to reason. The catalog of services within Azure Cognitive Services can be categorized into five main pillars - Vision, Speech, Language, Search, and Knowledge.
References:
https://docs.microsoft.com/en-us/azure/cognitive-services/welcome

NEW QUESTION: 101
You need to implement source control for scripts in an Azure Machine Learning workspace. You use a terminal window in the Azure Machine Learning Notebook tab You must authenticate your Git account with SSH.
You need to generate a new SSH key.
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 m the correct order.
Answer:
Explanation

NEW QUESTION: 102
You train classification and regression models by using automated machine learning.
You must evaluate automated machine learning experiment results. The results include how a classification model is making systematic errors in its predictions and the relationship between the target feature and the regression model's predictions. You must use charts generated by automated machine learning.
You need to choose a chart type for each model type.
Which chart types should you use? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.
Answer:
See below image

NEW QUESTION: 103
You need to implement early stopping criteria as suited in the model training requirements.
Which three code segments should you use to develop the solution? To answer, move the appropriate code segments from the list of code segments to the answer area and arrange them in the correct order.
NOTE: More than one order of answer choices is correct. You will receive credit for any of the correct orders you select.
Answer:
Explanation
You need to implement an early stopping criterion on models that provides savings without terminating promising jobs.
Truncation selection cancels a given percentage of lowest performing runs at each evaluation interval. Runs are compared based on their performance on the primary metric and the lowest X% are terminated.
Example:

from azureml.train.hyperdrive import TruncationSelectionPolicy
early_termination_policy = TruncationSelectionPolicy(evaluation_interval=1, truncation_percentage=20, delay_evaluation=5)

NEW QUESTION: 104
You train a model by using Azure Machine Learning. You use Azure Blob Storage to store production data.
The model must be re-trained when new data is uploaded to Azure Blob Storage. You need to minimize development and coding.
You need to configure Azure services to develop a re-training solution.
Which Azure services should you use? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.
Answer:
See below image

NEW QUESTION: 105
You register a file dataset named csvjolder that references a folder. The folder includes multiple com ma-separated values (CSV) files in an Azure storage blob container. You plan to use the following code to run a script that loads data from the file dataset. You create and instantiate the following variables:
You have the following code:
You need to pass the dataset to ensure that the script can read the files it references. Which code segment should you insert to replace the code comment?
A)
B)
C)
D)
A. Option A
B. Option B
C. Option C
D. Option D
Answer: (SHOW ANSWER)
Example:
from azureml.train.estimator import Estimator
script_params = {
# to mount files referenced by mnist dataset
'--data-folder': mnist_file_dataset.as_named_input('mnist_opendataset').as_mount(),
'--regularization': 0.5
}
est = Estimator(source_directory=script_folder,
script_params=script_params,
compute_target=compute_target,
environment_definition=env,
entry_script='train.py')
Reference:

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

NEW QUESTION: 106
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.
Answer:
Explanation

Valid DP-100 Dumps shared by BraindumpsPass.com for Helping Passing DP-100 Exam! BraindumpsPass.com now offer the **newest DP-100 exam dumps**, the BraindumpsPass.com DP-100 exam **questions have been updated** and **answers have been corrected** get the **newest** BraindumpsPass.com DP-100 dumps with Test Engine here: https://www.braindumpspass.com/Microsoft/DP-100-practice-exam-dumps.html (**519** Q&As Dumps, **40%OFF** **Special Discount:** **Exam-Tests**)