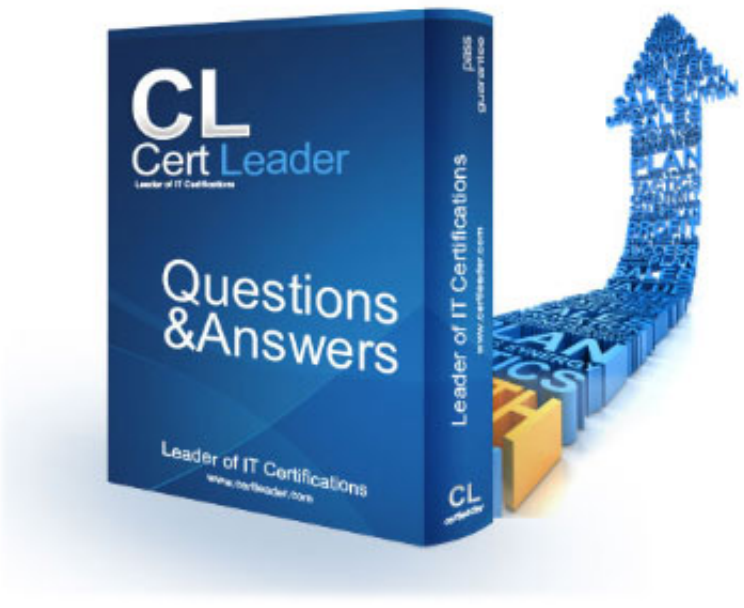
70-774 Dumps

Perform Cloud Data Science with Azure Machine Learning (beta)

https://www.certleader.com/70-774-dumps.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 sections, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.
You are designing an Azure Machine Learning workflow.
You have a dataset that contains two million large digital photographs.
You plan to detect the presence of trees in the photographs. You need to ensure that your model supports the following:
Solution: You create an Azure notebook that supports the Microsoft Cognitive Toolkit. Does this meet the goal?

A. Yes
B. No

Answer: B

NEW QUESTION 2
Note: This question is part of a series of questions that use the same scenario. For your convenience, the scenario is repeated in each question. Each question presents a different goal and answer choices, but the text of the scenario is exactly the same in each question in this series.
You plan to create a predictive analytics solution for credit risk assessment and fraud prediction in Azure Machine Learning. The Machine Learning workspace for the solution will be shared with other users in your organization. You will add assets to projects and conduct experiments in the workspace.
The experiments will be used for training models that will be published to provide scoring from web services. The experiment for fraud prediction will use Machine Learning modules and APIs to train the models and will predict probabilities in an Apache Hadoop ecosystem.
End of repeated scenario.
You need to alter the list of columns that will be used for predicting fraud for an input web service endpoint. The columns from the original data source must be retained while running the Machine Learning experiment.
Which module should you add after the web service input module and before the prediction module?

A. Edit Metadata
B. Import Data
C. SMOTE
D. Select Columns in Dataset

Answer: D

NEW QUESTION 3
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.
After you answer a question in this sections, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.
You are designing an Azure Machine Learning workflow.
You have a dataset that contains two million large digital photographs. You plan to detect the presence of trees in the photographs.
You need to ensure that your model supports the following:
Solution: You create a Machine Learning experiment that implements the Multiclass Decision Jungle module. Does this meet the goal?

A. Yes
B. No

Answer: B

NEW QUESTION 4
You have a dataset that is missing values in a column named Column3. Column3 is correlated to two columns named Column4 and Column5.
You need to improve the accuracy of the dataset, while minimizing data loss. What should you do?

A. Replace the missing values in Column3 by using probabilistic Principal Component Analysis (PCA).
B. Remove all of the rows that have the missing values in Column4 and Column5.
C. Replace the missing values in Column3 with a mean value.
D. Remove the rows that have the missing values in Column3.

Answer: A

NEW QUESTION 5
Note: This question is part of a series of questions that use the same scenario. For your convenience, the scenario is repeated in each question. Each question presents a different goal and answer choices, but the text of the scenario is exactly the same in each question in this series.
You plan to create a predictive analytics solution for credit risk assessment and fraud prediction in Azure Machine Learning. The Machine Learning workspace for the solution will be shared with other users in your organization. You will add assets to projects and conduct experiments in the workspace.
The experiments will be used for training models that will be published to provide scoring from web services. The experiment for fraud prediction will use Machine Learning modules and APIs to train the models and will predict probabilities in an Apache Hadoop ecosystem.
You plan to configure the resources for part of a workflow that will be used to preprocess data from files stored in Azure Blob storage. You plan to use Python to preprocess and store the data in Hadoop.
You need to get the data into Hadoop as quickly as possible.
Which three actions should you perform? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

A. Create an Azure virtual machine (VM), and then configure MapReduce on the VM.
B. Create an Azure HDInsight Hadoop cluster.
C. Create an Azure virtual machine (VM), and then install an IPython Notebook server.
D. Process the files by using Python to store the data to a Hadoop instance.
E. Create the Machine learning experiment, and then add an Execute Python Script module.

Answer: BDE

NEW QUESTION 6
You plan to use Azure Machine Learning to develop a predictive model. You plan to include an Execute Python Script module.
What capability does the module provide?

A. importing Python modules from a ZIP file for execution in a Machine Learning experiment
B. performing interactive debugging of a Python script
C. saving the results of a Python script run in a Machine Learning environment to a local file
D. returning multiple data frames

Answer: A

NEW QUESTION 7
Note: This question is part of a series of questions that use the same scenario. For your convenience, the scenario is repeated in each question. Each question presents a different goal and answer choices, but the text of the scenario is exactly the same in each question in this series.
A travel agency named Margie's Travel sells airline tickets to customers in the United States.
Margie's Travel wants you to provide insights and predictions on flight delays. The agency is considering implementing a system that will communicate to its customers as the flight departure nears about possible delays due to weather conditions. The flight data contains the following attributes:
The weather data contains the following attributes: AirportID, ReadingDate (YYYY/MM/DD HH), SkyConditionVisibility, WeatherType, WindSpeed, StationPressure, PressureChange, and HourlyPrecip.
You need to use historical data about on-time flight performance and the weather data to predict whether the departure of a scheduled flight will be delayed by more than 30 minutes.
Which method should you use?

A. clustering
B. linear regression
C. classification
D. anomaly detection

Answer: C

Explanation: References:
https://gallery.cortanaintelligence.com/Experiment/Binary-Classification-Flight-delay-prediction-3

NEW QUESTION 8
Note: This question is part of a series of questions that use the same or similar answer choices. An answer choice may be correct for more than one question in the series. Each question is independent of the other questions in this series. Information and details provided in a question apply only to that question.
You need to transform the columns in a dataset. The resulting data must be mean centered and have a variance of L The solution must use a native module.
Which module should you use?

A. Execute Python Script
B. Import Data
C. Edit Metadata
D. Select Columns in Dataset
E. Clean Missing Data
F. Tune Model Hyperparameters
G. Clip Values
H. Normalize Data

Answer: H

NEW QUESTION 9
You are building an Azure Machine Learning experiment.
You need to transform a string column into a label column for a Multiclass Decision Jungle module. Which module should you use?

A. Select Columns Transform
B. Group Categorical Values
C. Convert to Indicator Values
D. Edit Metadata

Answer: D

NEW QUESTION 10
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.
After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.
You have an Azure ML experiment that contains an intermediate dataset. You need to explore data from the intermediate dataset by using Jupyter.
Solution: You add a web service input to retrieve the data for the data source, and then add the Execute R Script module.
Does this meet the goal?

A. Yes
B. No

Answer: B

NEW QUESTION 10
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 sections, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.
You are designing an Azure Machine Learning workflow.
You have a dataset that contains two million large digital photographs. You plan to detect the presence of trees in the photographs.
You need to ensure that your model supports the following: Solution: You create an endpoint to the Computer vision API. Does this meet the goal?

A. Yes
B. No

Answer: B

NEW QUESTION 15
You plan to use the Import Data module to import data from a web URL by using HTTP. Which file format can you use as the source of the import operation?

A. Optimized Row Columnar (ORQ
B. Extensible Markup Language (XML)
C. comma-separated value (CSV)
D. JavaScript Object Notation (JSON)

Answer: D

NEW QUESTION 18
You are building a classification experiment in Azure Machine Learning.
You need to ensure that you can use the Evaluate Model module the experiment.
Which two actions should you perform? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

A. Connect the input of the Score Model modules to the output of the Evaluate Model module.
B. Connect the input of the Score Model modules to the output of the Train Model modules and the output Split Data modules.
C. Connect the output of the Score Model modules to the input of the Evaluate Model module.
D. Connect the output of the Score Model modules to the input of the Train Model modules and the input of the Split Data modules.

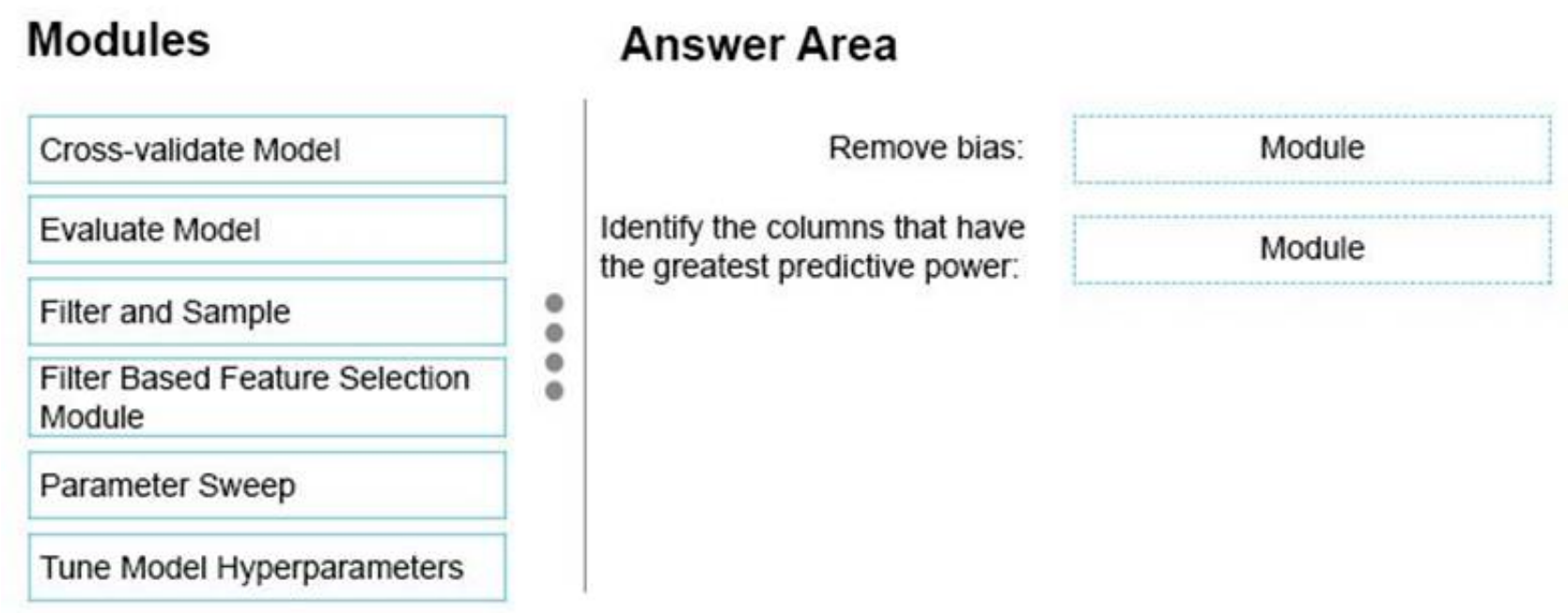
Answer: AB

NEW QUESTION 21
Note: This question is part of a series of questions that use the same scenario. For your convenience, the scenario is repeated in each question. Each question presents a different goal and answer choices, but the text of the scenario is exactly the same in each question in this scries.
Start of repeated scenario
You plan to create a predictive analytics solution for credit risk assessment and fraud prediction in Azure Machine Learning. The Machine Learning workspace for the solution will be shared with other users in your organization. You will add assets to projects and conduct experiments in the workspace.
The experiments will be used for training models that will be published to provide scoring from web services. The experiment for fraud prediction will use Machine Learning modules and APIs to train the models and will predict probabilities in an Apache Hadoop ecosystem.
End of repeated scenario.
You plan to share the Machine Learning workspace with the other users.
You are evaluating whether to assign the User role or the Owner role to several of the users.
Which three actions can be performed by the users who are assigned the User role? Each correct answer presents a complete solution.
NOTE: Each correct selection is worth one point.

A. Create, open, modify, and delete datasets.
B. Create, open, modify, and delete experiments.
C. Invite users to the workspace.
D. Delete users from the workspace.
E. Create, open, modify, and delete web services.
F. Access notebooks.

Answer: CDF

NEW QUESTION 25
Note: This question is part of a series of questions that use the same scenario. For your convenience, the scenario is repeated in each question. Each question presents a different goal and answer choices, but the text of the scenario is exactly the same in each question in this series.
A travel agency named Margie's Travel sells airline tickets to customers in the United States.
Margie's Travel wants you to provide insights and predictions on flight delays. The agency is considering implementing a system that will communicate to its customers as the flight departure nears about possible delays due to weather conditions. The flight data contains the following attributes:
The weather data contains the following attributes: AirportID, ReadingDate (YYYY/MM/DD HH), SkyConditionVisibility, WeatherType, WindSpeed, StationPressure, PressureChange, and HourlyPrecip.
You need to remove the bias and to identify the columns in the input dataset that have the greatest predictive power.
Which module should you use for each requirement? To answer, drag the appropriate modules to the correct requirements. Each module 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: References:
https://gallery.cortanaintelligence.com/Experiment/Binary-Classification-Flight-delay-prediction-3
https://msdn.microsoft.com/library/azure/038d91b6-c2f2-42a1-9215-1f2c20ed1b40

NEW QUESTION 26
You are working on an Azure Machine Learning experiment that uses four different logistic regression algorithms. You are evaluating the algorithms based on the data in the following table.

Metric	Model 1	Model 2	Model 3	Model 4
Mean absolute error (MAE)	.2	.63	.41	.46
Relative absolute error (RAE)	.40	.30	.60	.65
Root mean squared error (RMSE)	.25	.37	.15	.46

Which model produces predictions that are the closest to the actual outcomes?

A. Model 1
B. Model 2
C. Model 3
D. Model 4

Answer: A

NEW QUESTION 30
Note: This question is part of a series of questions that use the same scenario. For your convenience, the scenario is repeated in each question. Each question presents a different goal and answer choices, but the text of the scenario is exactly the same in each question in this series.
A travel agency named Margie's Travel sells airline tickets to customers in the United States.
Margie's Travel wants you to provide insights and predictions on flight delays. The agency is considering implementing a system that will communicate to its customers as the flight departure nears about possible delays due to weather conditions. The flight data contains the following attributes:
The weather data contains the following attributes: AirportID, ReadingDate (YYYY/MM/DD HH), SkyConditionVisibility, WeatherType, WindSpeed, StationPressure, PressureChange, and HourlyPrecip.
You have an untrained Azure Machine Learning model that you plan to train to predict flight delays.
You need to assess the variability of the dataset and the reliability of the predictions from the model. Which module should you use?

A. Cross-Validate Model
B. Evaluate Model
C. Tune Model Hyperparameters
D. Train Model
E. Score Model

Answer: A

Explanation: References:
https://msdn.microsoft.com/en-us/library/azure/dn905852.aspx

NEW QUESTION 35
You are building an Azure Machine Learning experiment.
You need to transform a string column that has 47 distinct values into a binary indicator column. The solution must use the One-vs-All Multiclass model.
Which module should you use?

A. Select Column Transform
B. Convert to Indicator Values

C. Group Categorical Values
D. Edit Metadata

Answer: B

NEW QUESTION 37
You are performing exploratory analysis of files that are encoded in a complex proprietary format. The format requires disk intensive access to several dependent files in HDFS.
You need to build an Azure Machine Learning model by using a canopy clustering algorithm. You must ensure that changes to proprietary file formats can be maintained by using the least amount of effort.
Which Machine Learning library should you use?

A. MicrosoftML
B. scikit-learn
C. SparkR
D. Mahout

Answer: C

NEW QUESTION 39
Note: This question is part of a series of questions that use the same or similar answer choices. An answer choice may be correct for more than one question in the series. Each question is independent of the other questions in this series. Information and details provided in a question apply only to that question.
You process some data by using Azure Machine Learning Studio. You have an intermediate dataset. The dataset has a column that contains date values stored in a format of MM/DD/YYYY.
ft You need to split the column into three separate columns by year, month, and day.
Which module should you use?

A. Edit Metadata
B. Normalize Data
C. Clean Missing Data
D. Import Data
E. Execute Python Script
F. Clip Values
G. Clip Values
H. Execute Python Script

Answer: A

NEW QUESTION 40
Note: This question is part of a series of questions that use the same or similar answer choices. An
answer choice may be correct for more than one question in the series. Each question is independent of the other questions in this series. Information and details provided in a question apply only to that question.
You need to change a column name to a friendly name. The solution must use a native module. Which module should you use?

A. Normalize Data
B. Select Columns in Dataset
C. Import Data
D. Edit Metadata
E. Tune Model Hyperparameters
F. Clean Missing Data
G. Clip Values
H. Execute Python Script

Answer: D

NEW QUESTION 41
You have an Apache Spark cluster in Azure HDinsight. The cluster includes 200 TB in five Apache Hive tables that have multiple foreign key relationships.
You have an Azure Machine Learning model that was built by using SPARK Accelerated Failure Time (AFT) Survival Regression Model (spark-survreg).
You need to prepare the Hive data into a single table as input for the Machine Learning model. The Hive data must be prepared in the least amount of time possible.
What should you use to prepare the data?

A. a Hive user-defined function (UDF)
B. Spark SQL
C. the GPU
D. Java Mapreduce jobs

Answer: A

NEW QUESTION 43
Note: This question is part of a series of questions that use the same or similar answer choices. An answer choice may be correct for more than one question in the series. Each question is independent of the other questions in this series. Information and details provided in a question apply only to that question.
You need to use only one percent of an Apache Hive data table by conducting random sampling by groups. Which module should you use?

A. Execute Python Script
B. Tune Model Hyperparameters
C. Normalize Data
D. Select Columns in Dataset

E. Import Data
F. Edit Metadata
G. Clip Values
H. Clean Missing Data

Answer: A

Explanation: References:
https://docs.microsoft.com/en-us/azure/machine-learning/team-data-science-process/sample-data-hive

NEW QUESTION 44
Note: This question is part of a series of questions that use the same or similar answer choices. An answer choice may be correct for more than one question in the series. Each question is independent of the other questions in this series. Information and details provided in a question apply only to that question.
You have a dataset that contains a column named Column1. Some of the values in Column1 are empty. You need to replace the empty values by using probabilistic Principal Component Analysis (PCA). The
solution must use a native module.
Which module should you use?

A. Execute Python Script
B. Clean Missing Data
C. Select Columns in Dataset
D. Import Data
E. Normalize Data
F. Edit Metadata
G. Tune Model Hyperparameters

Answer: B

NEW QUESTION 48
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.
After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.
You have an Azure ML experiment that contains an intermediate dataset. You need to explore data from the intermediate dataset by using Jupyter.
Solution: You add a Convert to ARFF module, and then add the Execute R Script module. Does this meet the goal?

A. Yes
B. No

Answer: B

NEW QUESTION 53
You have an Azure Machine Learning environment. You are evaluating whether to use R code or Python.
Which three actions can you perform by using both R code and Python in the Machine Learning environment? Each correct answer presents a complete solution.
NOTE: Each correct selection is worth one point.

A. Preprocess, cleanse, and group data.
B. Score a training model.
C. Create visualizations.
D. Create an untrained model that can be used with the Train Model module.
E. Implement feature ranking.

Answer: ABC

NEW QUESTION 56
Note: This question is part of a series of questions that use the same or similar answer choices. An answer choice may be correct for more than one question in the series. Each question is independent of the other questions in this series. Information and details provided in a question apply only to that question.
You need to remove rows that have an empty value in a specific column. The solution must use a native module.
Which module should you use?

A. Execute Python Script
B. Tune Model Hyperparameters
C. Normalize Data
D. Select Columns in Dataset
E. Import Data
F. Edit Metadata
G. Clip Values
H. Clean Missing Data

Answer: H

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

NEW QUESTION 60
You are building an Azure Machine Learning experiment.
You are preparing the output of a Boosted Decision Tree Regression module. You add a Normalize Data module to the experiment.

You need to ensure that the range of the transformation method produces an output on a scale of -1 to 1. Which transformation method should you use?

A. MinMax
B. TanH
C. Logistic
D. Zscore
E. LogNormal

Answer: D

NEW QUESTION 63
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 sections, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.
You are working on an Azure Machine Learning experiment. You have the dataset configured as shown in the following table.

Model	Mean absolute error (MAE)
Boosted decision tree	.2
Relative absolute error (RAE)	.43

You need to ensure that you can compare the performance of the models and add annotations to the results. Solution: You consolidate the output of the Score Model modules by using the Add Rows module, and then
use the Execute R Script module.
Does this meet the goal?

A. Yes
B. No

Answer: A

Explanation: References:
https://msdn.microsoft.com/en-us/library/azure/dn905915.aspx

NEW QUESTION 68
.....

Thank You for Trying Our Product

* 100% Pass or Money Back

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

* One year free update

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

* Trusted by Millions

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

* Shop Securely

All transactions are protected by VeriSign!

100% Pass Your 70-774 Exam with Our Prep Materials Via below:

https://www.certleader.com/70-774-dumps.html