

100% Money Back
Guarantee

Vendor:Microsoft

Exam Code:70-774

Exam Name:Perform Cloud Data Science with Azure
Machine Learning

Version:Demo

QUESTION 1

You are building an Azure Machine Learning solution for an online retailer.

When a customer selects a product, you need to recommend products that the customer might like to purchase at the same time. The recommendation should be based on what other customers purchased when they purchased the same product.

Which model should you use?

- A. Collaborative filtering
- B. Boosted Decision Tree Regression model
- C. Two-Class boosted decision tree
- D. K-Means Clustering

Correct Answer: A

QUESTION 2

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 (ORC)
- B. Extensible Markup Language (XML)
- C. comma-separated value (CSV)
- D. JavaScript Object Notation (JSON)

Correct Answer: D

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 an Azure notebook that supports the Microsoft Cognitive Toolkit.

Does this meet the goal?

A. Yes

B. No

Correct Answer: B

QUESTION 4

You have the following HiveQL query in an Import Data module.

```
from Student_Table
) a
where state_rank <= state_cnt*'{hiveconf:sampleRate}'

(
select
  field1, field2, field3, ..., fieldN, state,
  count(*) over (partition by state) as state_cnt,
  rank() over (partition by state order by rand()) as state_rank

from Student_Table
) a
where state_rank <= state_cnt*'{hiveconf:sampleRate}'
```

Which type of operation is being performed?

A. sampling a bucketized table

B. random sampling by groups

C. uniform random sampling

D. stratified sampling

Correct Answer: D

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.

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

Correct Answer: A

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

QUESTION 6

DRAG DROP

You need to create a training set for use with a linear regression model, and then to test the training set.

How should you create the training set? To answer, drag the appropriate modules to the correct locations. 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

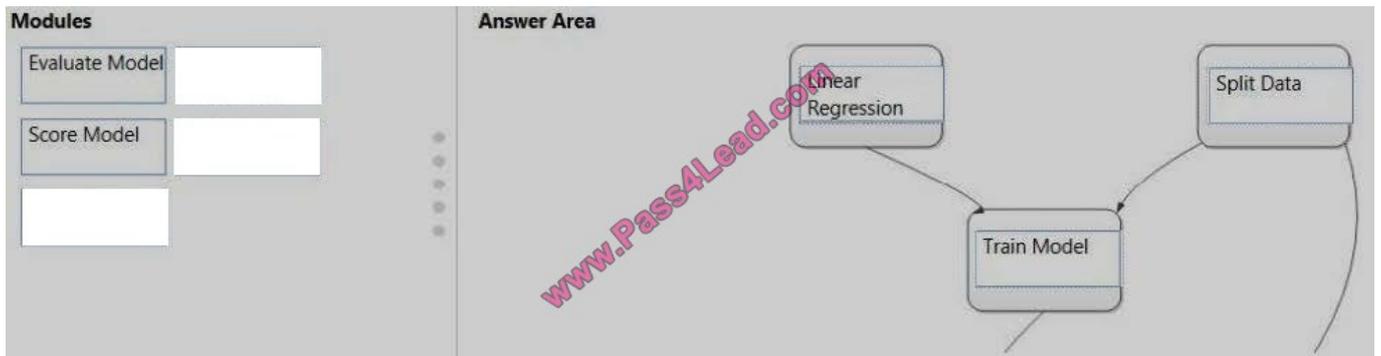
content.

NOTE: Each correct selection is worth one point.

Select and Place:



Correct Answer:



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.

Start of repeated scenario You plan to use Azure platform tools to detect and analyze food items in smart refrigerators. To provide families with an integrated experience for grocery shopping and cooking, the refrigerators will connect to other smart appliances, such as

stoves and microwave ovens, on a LAN.

You plan to build an object recognition model by using the Microsoft Cognitive Toolkit. The object recognition model will receive input from the connected devices and send results to applications.

The training data will be derived from more than 10 TB of images. You will convert the raw images to the sparse format.

End of repeated scenario.

You need to ensure that a web service endpoint can receive image data and use an object recognition model to return the expected object and the confidence level of the model. The solution must minimize the effort required to generate the

client code to access the web service.

Which resource should you use?

- A. the edX Data Science Learning Dashboard
- B. Azure Machine Learning Studio
- C. Cortana Intelligence Gallery
- D. the Data Science Virtual Machine

Correct Answer: B

QUESTION 8

You are analyzing taxi trips in New York City. You leverage the Azure Data Factory to create data pipelines and to orchestrate data movement.

You plan to develop a predictive model for 170 million rows (37 GB) of raw data in Apache Hive by using Microsoft R Server to identify which factors contribute to the passenger tipping behavior.

All of the platforms that are used for the analysis are the same. Each worker node has eight processor cores and 26 GB of memory.

Which type of Azure HDInsight cluster should you use to produce results as quickly as possible?

- A. Hadoop
- B. HBase
- C. Interactive Hive
- D. Spark

Correct Answer: D

References:

<https://azure.microsoft.com/en-gb/blog/general-availability-of-hdinsight-interactive-query- blazing-fast-data-warehouse-style-queries-on-hyper-scale-data-2/>

QUESTION 9

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. Outputting a file to a network location.
- 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. Visualizing univariate and multivariate summaries by using Python code.

Correct Answer: D

QUESTION 10

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 plan to predict flight delays that are 30 minutes or more.

You need to build a training model that accurately fits the data. The solution must minimize over fitting and minimize data leakage.

Which attribute should you remove?

- A. OriginAirportID
- B. DepDel
- C. DepDel30
- D. Carrier
- E. DestAirportID

Correct Answer: C

QUESTION 11

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

Correct Answer: B

QUESTION 12

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

Correct Answer: D

To Read the [Whole Q&As](#), please purchase the [Complete Version](#) from [Our website](#).

Try our product !

100% Guaranteed Success

100% Money Back Guarantee

365 Days Free Update

Instant Download After Purchase

24x7 Customer Support

Average **99.9%** Success Rate

More than **800,000** Satisfied Customers Worldwide

Multi-Platform capabilities - **Windows, Mac, Android, iPhone, iPod, iPad, Kindle**

Need Help

Please provide as much detail as possible so we can best assist you.

To update a previously submitted ticket:



 <p>One Year Free Update Free update is available within One Year after your purchase. After One Year, you will get 50% discounts for updating. And we are proud to boast a 24/7 efficient Customer Support system via Email.</p>	 <p>Money Back Guarantee To ensure that you are spending on quality products, we provide 100% money back guarantee for 30 days from the date of purchase.</p>	 <p>Security & Privacy We respect customer privacy. We use McAfee's security service to provide you with utmost security for your personal information & peace of mind.</p>
---	---	--

Any charges made through this site will appear as Global Simulators Limited.

All trademarks are the property of their respective owners.