

Designing and Implementing a Data Science Solution on Azure (beta)

1. DRAG DROP (Drag and Drop is not supported)

You are planning to host practical training to acquaint staff with Docker for Windows.

Staff devices must support the installation of Docker.

Which of the following are requirements for this installation? Answer by dragging the correct options from the list to the answer area.

Select and Place:

Options

2 GB of system RAM

4 GB of system RAM

BIOS-enabled virtualization

Microsoft Hardware-Assisted Virtualization Detection Tool

Windows 10 64-bit

Windows 10 32-bit

Answer

A. See Explanation section for answer.

Answer(s): A

2. HOTSPOT (Drag and Drop is not supported)

Complete the sentence by selecting the correct option in the answer area.

Hot Area:

Answer Area

	▼
SSD	
FPGA	
GPU	
Power BI	

is required for a Deep Learning Virtual Machine (DLVM) to support Compute Unified Device Architecture (CUDA) computations.

A. See Explanation section for answer.

Answer(s): A

3. You need to implement a Data Science Virtual Machine (DSVM) that supports the Caffe2 deep learning framework.

Which of the following DSVM should you create?

A. Windows Server 2012 DSVM

B. Windows Server 2016 DSVM

C. Ubuntu 16.04 DSVM

D. CentOS 7.4 DSVM

Answer(s): C

4. This question is included in a number of questions that depicts the identical set-up. However, every question has a distinctive result. Establish if the recommendation satisfies the requirements.

You have been tasked with employing a machine learning model, which makes use of a PostgreSQL database and needs GPU processing, to forecast prices.

You are preparing to create a virtual machine that has the necessary tools built into it.

You need to make use of the correct virtual machine type.

Solution: You make use of a Geo AI Data Science Virtual Machine (Geo-DSVM) Windows edition.
Will the requirements be satisfied?

A. Yes

B. No

Answer(s): B

5. This question is included in a number of questions that depicts the identical set-up. However, every question has a distinctive result. Establish if the recommendation satisfies the requirements.

You have been tasked with employing a machine learning model, which makes use of a PostgreSQL database and needs GPU processing, to forecast prices.

You are preparing to create a virtual machine that has the necessary tools built into it.

You need to make use of the correct virtual machine type.

Solution: You make use of a Deep Learning Virtual Machine (DLVM) Windows edition.

Will the requirements be satisfied?

A. Yes

B. No

Answer(s): B

6. This question is included in a number of questions that depicts the identical set-up. However, every question has a distinctive result. Establish if the recommendation satisfies the requirements.

You have been tasked with employing a machine learning model, which makes use of a PostgreSQL database and needs GPU processing, to forecast prices.

You are preparing to create a virtual machine that has the necessary tools built into it.

You need to make use of the correct virtual machine type.

Solution: You make use of a Data Science Virtual Machine (DSVM) Windows edition.

Will the requirements be satisfied?

A. Yes

B. No

Answer(s): A

7. DRAG DROP (Drag and Drop is not supported)

You have been tasked with moving data into Azure Blob Storage for the purpose of supporting Azure

Machine Learning.

Which of the following can be used to complete your task? Answer by dragging the correct options from the list to the answer area.

Select and Place:

Options

Answer

AzCopy

Bulk Copy Program
(BCP)

SSIS

Bulk Insert SQL Query

Azure Storage
Explorer

A. See Explanation section for answer.

Answer(s): A

8. You have been tasked with designing a deep learning model, which accommodates the most recent edition of Python, to recognize language.

You have to include a suitable deep learning framework in the Data Science Virtual Machine (DSVM).

Which of the following actions should you take?

A. You should consider including Rattle.

B. You should consider including TensorFlow.

C. You should consider including Theano.

D. You should consider including Chainer.

Answer(s): B

9. This question is included in a number of questions that depicts the identical set-up. However, every question has a distinctive result. Establish if the recommendation satisfies the requirements.

You have been tasked with evaluating your model on a partial data sample via k-fold cross-validation. You have already configured a k parameter as the number of splits. You now have to configure the k parameter for the cross-validation with the usual value choice.

Solution: You configure the use of the value $k=3$.

Will the requirements be satisfied?

A. Yes

B. No

Answer(s): B

10. This question is included in a number of questions that depicts the identical set-up. However, every question has a distinctive result. Establish if the recommendation satisfies the requirements.

You have been tasked with evaluating your model on a partial data sample via k-fold cross-validation. You have already configured a k parameter as the number of splits. You now have to configure the k parameter for the cross-validation with the usual value choice.

Solution: You configure the use of the value $k=10$.

Will the requirements be satisfied?

A. Yes

B. No

Answer(s): A

11. You construct a machine learning experiment via Azure Machine Learning Studio.

You would like to split data into two separate datasets.

Which of the following actions should you take?

A. You should make use of the Split Data module.

B. You should make use of the Group Categorical Values module.

C. You should make use of the Clip Values module.

D. You should make use of the Group Data into Bins module.

Answer(s): D

12. You have been tasked with creating a new Azure pipeline via the Machine Learning designer. You have to make sure that the pipeline trains a model using data in a comma-separated values (CSV) file that is published on a website. A dataset for the file for this file does not exist. Data from the CSV file must be ingested into the designer pipeline with the least amount of administrative effort as possible.

Which of the following actions should you take?

A. You should make use of the Convert to TXT module.

B. You should add the Copy Data object to the pipeline.

C. You should add the Import Data object to the pipeline.

D. You should add the Dataset object to the pipeline.

Answer(s): D

13. This question is included in a number of questions that depicts the identical set-up. However, every question has a distinctive result. Establish if the recommendation satisfies the requirements.

You are in the process of creating a machine learning model. Your dataset includes rows with null and missing values.

You plan to make use of the Clean Missing Data module in Azure Machine Learning Studio to detect and fix the null and missing values in the dataset.

Solution: You make use of the Replace with median option.

Will the requirements be satisfied?

A. Yes

B. No

Answer(s): B

14. This question is included in a number of questions that depicts the identical set-up. However, every question has a distinctive result. Establish if the recommendation satisfies the requirements.

You are in the process of creating a machine learning model. Your dataset includes rows with null and missing values.

You plan to make use of the Clean Missing Data module in Azure Machine Learning Studio to detect and fix the null and missing values in the dataset.

Solution: You make use of the Custom substitution value option.

Will the requirements be satisfied?

A. Yes

B. No

Answer(s): B

15. This question is included in a number of questions that depicts the identical set-up. However, every question has a distinctive result. Establish if the recommendation satisfies the requirements.

You are in the process of creating a machine learning model. Your dataset includes rows with null and missing values.

You plan to make use of the Clean Missing Data module in Azure Machine Learning Studio to detect and fix the null and missing values in the dataset.

Solution: You make use of the Remove entire row option.

Will the requirements be satisfied?

A. Yes

B. No

Answer(s): A

16. You need to consider the underlined segment to establish whether it is accurate.

To transform a categorical feature into a binary indicator, you should make use of the Clean Missing Data module.

Select `No adjustment required` if the underlined segment is accurate. If the underlined segment is inaccurate, select the accurate option.

A. No adjustment required.

B. Convert to Indicator Values

C. Apply SQL Transformation

D. Group Categorical Values

Answer(s): B

17. You need to consider the underlined segment to establish whether it is accurate.

To improve the amount of low incidence cases in a dataset, you should make use of the SMOTE module. Select `No adjustment required` if the underlined segment is accurate. If the underlined segment is inaccurate, select the accurate option.

A. No adjustment required.

B. Remove Duplicate Rows

C. Join Data

D. Edit Metadata

Answer(s): A

18. HOTSPOT (Drag and Drop is not supported)

You need to consider the underlined segment to establish whether it is accurate.

Hot Area:

Answer Area

The

	▼
Venn diagram	
Box plot	
Gradient descent	
Violin plot	

visualization can be used to reveal outliers in your data.

A. See Explanation section for answer.

Answer(s): A

19. You are planning to host practical training to acquaint learners with data visualization creation using Python. Learner devices are able to connect to the internet.

Learner devices are currently NOT configured for Python development. Also, learners are unable to install software on their devices as they lack administrator permissions. Furthermore, they are unable to access Azure subscriptions.

It is imperative that learners are able to execute Python-based data visualization code.

Which of the following actions should you take?

A. You should consider configuring the use of Azure Container Instance.

B. You should consider configuring the use of Azure BatchAI.

C. You should consider configuring the use of Azure Notebooks.

D. You should consider configuring the use of Azure Kubernetes Service.

Answer(s): C

20. You have recently concluded the construction of a binary classification machine learning model. You are currently assessing the model. You want to make use of a visualization that allows for precision to be used as the measurement for the assessment.

Which of the following actions should you take?

A. You should consider using Venn diagram visualization.

B. You should consider using Receiver Operating Characteristic (ROC) curve visualization.

C. You should consider using Box plot visualization.

D. You should consider using the Binary classification confusion matrix visualization.

Answer(s): D
