DP-100^{Q&As}

Designing and Implementing a Data Science Solution on Azure

Pass Microsoft DP-100 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

https://www.pass2lead.com/dp-100.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Microsoft Official Exam Center

⚙ **Instant Download** After Purchase

⚙ **100% Money Back** Guarantee

⚙ **365 Days** Free Update

⚙ **800,000+** Satisfied Customers

![Pass2Lead](https://Pass2Lead.com)
QUESTION 1

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.

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

Will the requirements be satisfied?

A. Yes

B. No

Correct Answer: B

Reference: https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/clean-missing-data

QUESTION 2

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.

Recommendation: You configure the use of the value k=10.

Will the requirements be satisfied?

A. Yes

B. No

Correct Answer: A

Leave One Out (LOO) cross-validation

Setting K = n (the number of observations) yields n-fold and is called leave-one out cross-validation (LOO), a special case of the K-fold approach.

LOO CV is sometimes useful but typically doesn\\'t shake up the data enough. The estimates from each fold are highly correlated and hence their average can have high variance.

This is why the usual choice is K=5 or 10. It provides a good compromise for the bias-variance tradeoff.

QUESTION 3

HOTSPOT

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.

Correct Answer:

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 Deep Learning Virtual Machine is a pre-configured environment for deep learning using GPU instances.

QUESTION 4

You have been tasked with ascertaining if two sets of data differ considerably. You will make use of Azure Machine Learning Studio to complete your task.

You plan to perform a paired t-test.

Which of the following are conditions that must apply to use a paired t-test? (Choose all that apply.)

A. All scores are independent from each other.

B. You have a matched pairs of scores.

C. The sampling distribution of d is normal.

D. The sampling distribution of x1- x2 is normal.

Correct Answer: BC

Reference: https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/test-hypothesis-using-t-test

---

QUESTION 5

You are implementing hyperparameter tuning by using Bayesian sampling for a model training from a notebook. The notebook is in an Azure Machine Learning workspace that uses a compute cluster with 20 nodes.

The code implements Bandit termination policy with slack factor set to 0.2 and the HyperDriveConfig class instance with max_concurrent_runs set to 10.

You must increase effectiveness of the tuning process by improving sampling convergence.

You need to select which sampling convergence to use.

What should you select?

A. Set the value of slack factor of early_termination_policy to 0.9.

B. Set the value of max_concurrent_runs of HyperDriveConfig to 4.

C. Set the value of slack factor of early_termination_policy to 0.1.

D. Set the value of max_concurrent_runs of HyperDriveConfig to 20.

Correct Answer: B

Bayesian sampling

Bayesian sampling is based on the Bayesian optimization algorithm. It picks samples based on how previous samples did, so that new samples improve the primary metric.

The number of concurrent jobs has an impact on the effectiveness of the tuning process. A smaller number of concurrent jobs may lead to better sampling convergence, since the smaller degree of parallelism increases the number of jobs

that benefit from previously completed jobs.

Reference: https://learn.microsoft.com/en-us/azure/machine-learning/how-to-tune-hyperparameters

DP-100 VCE Dumps   **DP-100 Exam Questions**   **DP-100 Braindumps**