



Free Questions for C1000-059 by dumpshq

Shared by Hughes on 12-12-2023

For More Free Questions and Preparation Resources

Check the Links on Last Page

Question 1

Question Type: MultipleChoice

Which is an example of a nominal scale data?

Options:

- A- a variable industry with categorical values such as financial, engineering, and retail
- B- a variable mood with a scale of values unhappy, ok, and happy
- C- a variable bank account balance whose possible values are \$5, \$10, and \$15
- D- a variable temperature with a scale of values low, medium, and high

Answer:

C

Question 2

Question Type: MultipleChoice

A data scientist is exploring transaction data from a chain of stores with several locations. The data includes store number, date of sale, and purchase amount.

If the data scientist wants to compare total monthly sales between stores, which two options would be good ways to aggregate the data? (Choose two.)

Options:

- A- Find the sum of the transaction prices
- B- Select the largest transaction amount by month and store
- C- Write a GROUP BY query
- D- Plot a time series plot of transaction amounts
- E- Generate a pivot table

Answer:

B, D

Question 3

Question Type: MultipleChoice

A data analyst creates a term-document matrix for the following sentence: I saw a cat, a dog and another cat.

Assuming they used a binary vectorizer, what is the resulting weight for the word cat?

Options:

A- 0

B- 1

C- 3

D- 2

Answer:

B

Question 4

Question Type: MultipleChoice

In a hyperparameter search, whether a single model is trained or a lot of models are trained in parallel is largely determined by?

Options:

- A- The number of hyperparameters you have to tune.
- B- The presence of local minima in your neural network.
- C- The amount of computational power you can access.
- D- Whether you use batch or mini-batch optimization.

Answer:

C

Question 5

Question Type: MultipleChoice

What is the main difference between traditional programming and machine learning?

Options:

- A- Machine learning models take less time to train.

- B-** Machine learning takes full advantage of SDKs and APIs.
- C-** Machine learning is optimized to run on parallel computing and cloud computing.
- D-** Machine learning does not require explicit coding of decision logic.

Answer:

D

Question 6

Question Type: MultipleChoice

What is the name of the design thinking work product that contains a summary description of a particular person or role?

Options:

- A-** persona
- B-** snapshot
- C-** My Sticky Note
- D-** user summary report

Answer:

A

Question 7

Question Type: MultipleChoice

What are two methods used to detect outliers in structured data? (Choose two.)

Options:

A- multi-label classification

B- isolation forest

C- gradient descent

D- one class Support Vector Machine (SVM)

E- Word2Vec

Answer:

B, D

Question 8

Question Type: MultipleChoice

A classification task has examples that are labeled as belonging to one of two classes:

- * 90% of the examples belong to class-1
- * 10% belong to class-2

Which two techniques are appropriate to deal with the class imbalance? (Choose two.)

Options:

- A-** apply dimensionality reduction to the features before training
- B-** impose an additional cost on the model for making classification mistakes on the minority class during training
- C-** lower the detection threshold of the minority class after training
- D-** oversample the minority class and/or undersample the majority class
- E-** after training, divide the model accuracy of each class by the proportion that they represent in the dataset

Answer:

B, E

Question 9

Question Type: MultipleChoice

What is meant by the curse of dimensionality?

Options:

- A- The number of available algorithms for a given task is high.
- B- The number of available data sources for a given task is high.
- C- The data sparsity becomes more severe as the number of features is increased.
- D- The data sparsity becomes more severe as the number of samples is increased.

Answer:

B

Question 10

Question Type: MultipleChoice

The least squares optimization technique (The Method of Least Squares) is used in which algorithm?

Options:

- A- Support Vector Machines
- B- Naive Bayes classification
- C- Logistic regression
- D- Linear regression

Answer:

D

Question 11

Question Type: MultipleChoice

What are three elements that are typically part of a machine learning pipeline in scikit-learn or pyspark? (Choose three.)

Options:

A- model building

B- data preprocessing

C- model prediction

D- business understanding

E- use case selection

F- data exploration

Answer:

B, C, F

To Get Premium Files for C1000-059 Visit

<https://www.p2pexams.com/products/c1000-059>

For More Free Questions Visit

<https://www.p2pexams.com/ibm/pdf/c1000-059>

