

Huawei H13-731

Huawei Certified ICT Expert-Big Data-Data Mining

For More Information – Visit link below:

<https://www.examsempire.com/>

Product Version

- 1. Up to Date products, reliable and verified.**
- 2. Questions and Answers in PDF Format.**



<https://examsempire.com/>

Visit us at: <https://www.examsempire.com/h13-731>

Latest Version: 6.0

Question: 1

You need to perform L2 regularization (Ridge) on a LogisticRegression model in PySpark. How should you configure the elasticNetParam?

- A. Set elasticNetParam = 1.0
- B. Set elasticNetParam = 0.5
- C. Set elasticNetParam = 0.0
- D. Set elasticNetParam = -1.0

Answer: C

Question: 2

Which of the following are distinct advantages of using Huawei MRS ClickHouse compared to standard distributed databases?

(Select all that apply)

- A. High-performance columnar storage and data compression.
- B. Support for massive parallel processing (MPP).
- C. Native support for complex recursive transactions (ACID).
- D. Extremely fast real-time aggregation and analytical query response.

Answer: A,B,D

Question: 3

Before performing PCA, which preprocessing steps are highly recommended?

(Select all that apply)

- A. Feature Scaling (Standardization).
- B. Mean Centering (subtracting the mean).
- C. One-Hot Encoding of high-cardinality strings.
- D. Removing all outliers.

Answer: A,B

Question: 4

A data analyst needs to find where a specific "customer_id" column originated and which downstream reports it influences. Which DataArts Studio feature provides this visualization?

- A. Data Standards
- B. Data Service API
- C. Data Lineage (within Data Map)
- D. Data Masking

Answer: C

Question: 5

For which of the following datasets would DBSCAN be a better choice than K-Means?

- A. A dataset with three perfectly spherical, well-separated clusters.
- B. A dataset where clusters are shaped like concentric circles (rings).
- C. A dataset where all features are binary (0 and 1).
- D. A dataset where the number of clusters is known to be exactly four.

Answer: B

Question: 6

In a classification problem, the evaluation metric defined as the ratio of the number of correct predictions (both Positive and Negative) to the total number of input samples is: _____

- A. Precision
- B. Recall
- C. F1-Score
- D. Accuracy

Answer: D

Question: 7

Given a transaction set of 100 baskets: 40 baskets contain Bread, 30 contain Butter, and 20 contain both. What is the Confidence of the rule {Bread} -> {Butter}?

- A. 0.2
- B. 0.5
- C. 0.66
- D. 0.3

Answer: B

Question: 8

During the Data Integration phase of a large-scale project in Huawei DataArts Studio, you identify two attributes, purchase_price and tax_amount, that have a Pearson correlation coefficient of 0.98. What is the most appropriate expert action to take?

- A. Remove one of the features or combine them to mitigate redundancy and multicollinearity.
- B. Keep both features to provide more data to the model.
- C. Apply One-Hot Encoding to both features.
- D. Increase the number of samples in the dataset to compensate for the correlation.

Answer: A

Question: 9

In a multi-class classification problem (e.g., classifying 10 different types of fruit), you use the "One-vs-Rest" (OvR) strategy. How many binary classifiers will be trained?

- A. 1
- B. 9
- C. 10
- D. 45

Answer: C

Question: 10

When performing Binning (Discretization), why might an expert choose "Equal Frequency" (Quantiles) over "Equal Width"?

- A. Because it is computationally simpler.
- B. To ensure the bins are easier for humans to interpret (e.g., 0-10, 10-20).
- C. To ensure the bins have the same numerical range.
- D. To handle skewed data where most values are clustered in a narrow range.

Answer: D

Thank You for Trying Our Product
Special 16 USD Discount Coupon: NSZUBG3X

Email: support@examsempire.com

**Check our Customer Testimonials and ratings
available on every product page.**

Visit our website.

<https://examsempire.com/>