



DATABRICKS-CERTIFIED- PR OFESIONAL-DATA-SCIENTIST^{Q&As}

Databricks Certified Professional Data Scientist Exam

**Pass Databricks DATABRICKS-CERTIFIED-
PROFESSIONAL-DATA-SCIENTIST Exam with 100%
Guarantee**

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

<https://www.passapply.com/databricks-certified-professional-data-scientist.html>

100% Passing Guarantee
100% Money Back Assurance





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



VCE & PDF

PassApply.com

<https://www.passapply.com/databricks-certified-professional-data-scientist.2024> Latest passapply DATABRICKS-CERTIFIED-PROFESSIONAL-DATA-SCIENTIST PDF and VCE dumps Download

-  **Instant Download** After Purchase
-  **100% Money Back** Guarantee
-  **365 Days** Free Update
-  **800,000+** Satisfied Customers





QUESTION 1

- A. Naive Bayes classifier
- B. Collaborative filtering
- C. Logistic Regression
- D. Content-based filtering

Correct Answer: B

Explanation: One scenario of collaborative filtering application is to recommend interesting or popular information as judged by the community. As a typical example, stories appear in the front page of Digg as they are "voted up" (rated positively) by the community. As the community becomes larger and more diverse, the promoted stories can better reflect the average interest of the community members.

QUESTION 2

You are working as a data science consultant for a gaming company. You have three member team and all other stake holders are from the company itself like project managers and project sponsored, data team etc. During the discussion project managed asked you that when can you tell me that the model you are using is robust enough, after which step you can consider answer for this question?

- A. Data Preparation
- B. Discovery
- C. Operationalize
- D. Model planning
- E. Model building

Correct Answer: E

To answer whether the model you are building is robust enough or not you need to have answer below questions at least

- Model is performing as expected with the test data or not?
 - Whatever hypothesis defined in the initial phase is being tested or not?
 - Do we need more data?
 - Domain experts are convinced or not with the model? And all these can be answered when you have built the model and tested with the test data sets. Hence, correct option will be Model Building.
-

QUESTION 3

A denote the event 'student is female' and let B denote the event 'student is French'. In a class of 100 students



suppose 60 are French, and suppose that 10 of the French students are females. Find the probability that if I pick a French student, it will be a girl, that is, find $P(A|B)$.

- A. $1/3$
- B. $2/3$
- C. $1/6$
- D. $2/6$

Correct Answer: C

Explanation: Since 10 out of 100 students are both French and female, then $P(A \text{ and } B) = 10/100$ Also. 60 out of the 100 students are French, so $P(B) = 60/100$ So the required probability is: $P(A|B) = P(A \text{ and } B) / P(B) = 10/100 \cdot 100/60 = 1/6$

QUESTION 4

Which is an example of supervised learning?

- A. PCA
- B. k-means clustering
- C. SVD
- D. EM
- E. SVM

Correct Answer: E

Explanation: SVMs can be used to solve various real world problems:

SVMs are helpful in text and hypertext categorization as their application can significantly reduce the need for labeled training instances in both the standard inductive and transductive settings.

Classification of images can also be performed using SVMs. Experimental results show that SVMs achieve significantly higher search accuracy than traditional query refinement schemes after just three to four rounds of relevance feedback.

SVMs are also useful in medical science to classify proteins with up to 90% of the compounds classified correctly.

Hand-written characters can be recognized using SVM

QUESTION 5

Which method is used to solve for coefficients b_0, b_1, \dots, b_n in your linear regression model:

$$Y = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n$$



- A. Apriori Algorithm
- B. Ridge and Lasso
- C. Ordinary Least squares
- D. Integer programming

Correct Answer: C

Explanation: : $RY = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n$ In the linear model, the b_i 's represent the unknown p parameters. The estimates for these unknown parameters are chosen so that, on average, the model provides a reasonable estimate of a person's income based on age and education. In other words, the fitted model should minimize the overall error between the linear model and the actual observations. Ordinary Least Squares (OLS) is a common technique to estimate the parameters

[DATABRICKS-CERTIFIED-PROFESSIONAL-DATA-SCIENTIST PDF Dumps](#)

[DATABRICKS-CERTIFIED-PROFESSIONAL-DATA-SCIENTIST Exam Questions](#)

[DATABRICKS-CERTIFIED-PROFESSIONAL-DATA-SCIENTIST Braindumps](#)