DATABRICKS-CERTIFIED-PROFESSIONAL-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.pass4itsure.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

QUESTION 1

Let\\'s say you have two cases as below for the movie ratings

1.

You recommend to a user a movie with four stars and he really doesn\\'t like it and he\\'d rate it two stars

2.

You recommend a movie with three stars but the user loves it (he\\'d rate it five stars). So which statement correctly applies?

A. In both cases, the contribution to the RMSE is the same

B. In both cases, the contribution to the RMSE is the different

C. In both cases, the contribution to the RMSE, could varies

D. None of the above

Correct Answer: A

QUESTION 2

Spam filtering of the emails is an example of

A. Supervised learning

B. Unsupervised learning

C. Clustering

D. 1 and 3 are correct

E. 2 and 3 are correct

Correct Answer: A

Explanation: Clustering is an example of unsupervised learning. The clustering algorithm finds groups within the data without being told what to look for upfront. This contrasts with classification, an example of supervised machine learning, which is the process of determining to which class an observation belongs. A common application of classification is spam filtering. With spam filtering we use labeled data to train the classifier: e-mails marked as spam or ham.

QUESTION 3

Which of the following skills a data scientists required?

A. Web designing to represent best visuals of its results from algorithm.

B. He should be creative

C. Should possess good programming skills

D. Should be very good at mathematics and statistic

E. He should possess database administrative skills.

Correct Answer: BCD

Explanation: Yes a data scientists should have combination of skills like to solve the complex problem he should be creative as well as able to find new solutions and use of existing data. And solve the problem skills required are programming as currently we see SAS, R: Python, Spark, Java and SPSS even day by day new technologies are coming. To apply various existing and new algorithm using Machine Learning, or Al it require good mathematics and statistics skills (Where the programmer feels, weaknesses). Another skill required is using visualization techniques like Qlik, Tableau etc

QUESTION 4

Suppose you have been given a relatively high-dimension set of independent variables and you are asked to come up with a model that predicts one of Two possible outcomes like "YES" or "NO", then which of the following technique best fit?

A. Support vector machines

B. Naive Bayes

C. Logistic regression

D. Random decision forests

E. All of the above

Correct Answer: E

Explanation: In this problem you have been given high-dimensional independent variables like yeS; nO; no English words , test results etc. and you have to predict either valid or not valid (One of two). So all of the below technique can be applied to this problem. Support vector machines Naive Bayes Logistic regression Random decision forests

QUESTION 5

You are using one approach for the classification where to teach the agent not by giving explicit categorizations, but by using some sort of reward system to indicate success, where agents might be rewarded for doing certain actions and

punished for doing others.

Which kind of this learning?

A. Supervised

B. Unsupervised

C. Regression

D. None of the above

Correct Answer: B

Explanation: Unsupervised learning seems much harder: the goal is to have the computer learn how to do something that we don\\'t tell it how to do! The approach is to teach the agent not by giving explicit categorizations, but by using some sort of reward system to indicate success. Note that this type of training will generally fit into the decision problem framework because the goal is not to produce a classification but to make decisions that maximize rewards. This approach nicely generalizes to the real world, where agents might be rewarded for doing certain actions and punished fordoing others.