DS-200^Q&As

Data Science Essentials

Pass Cloudera DS-200 Exam with 100% Guarantee

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

https://www.pass2lead.com/ds-200.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Cloudera 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

Under what two conditions does stochastic gradient descent outperform 2nd-order optimization techniques such as iteratively reweighted least squares?

A. When the volume of input data is so large and diverse that a 2nd-order optimization technique can be fit to a sample of the data

B. When the model\\'s estimates must be updated in real-time in order to account for new observations.

C. When the input data can easily fit into memory on a single machine, but we want to calculate confidence intervals for all of the parameters in the model.

D. When we are required to find the parameters that return the optimal value of the objective function.

Correct Answer: AB

QUESTION 2

Given the following sample of numbers from a distribution: 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89

What are two benefits of using the five-number summary of sample percentiles to summarize a data set?

A. You can calculate unbiased estimators for the parameters of the distribution

B. It\\'s robust to outliers

C. It\\'s well-defined for any probability distribution

D. You can calculate it quickly using a relational database like MySQL, even when we have a large sample

Correct Answer: D

QUESTION 3

In what format are web server log files usually generated and how must you transform them in order to make them usable for analysis in Hadoop?

A. XML files that you need to convert to JSON

B. Text files that require parsing into useful fields

C. CSV files that require parsing into useful fields

D. HTML files that you need to convert to plain text or CSV

E. Binary files that may require decompression and conversion using AVRO

Correct Answer: AB

QUESTION 4

When optimizing a function using stochastic gradient descent, how frequently should you update your estimate of the gradient?

A. Once after every pass through the data set

B. Once per observation

C. For each observation with a probability that you choose ahead of time

D. After a random number of observations

E. Once every N observations, where you decide N ahead of time

Correct Answer: AC

QUESTION 5

What are three benefits of running feature selection analysis before filtering a classification model?

A. Eliminates the need to include a regularization term

B. Reduces the number of subjective decisions required to construct the model

C. Guarantee the optimally of the final model

D. Speeds up the model fitting process

E. Develops an understanding of the importance of different features

F. Improves the predictive performance of the model

Correct Answer: DEF

[DS-200 VCE Dumps](#)   [DS-200 Practice Test](#)   [DS-200 Braindumps](#)