



- Expert Verified, Online, Free.

MENU



Google Discussions



Exam Professional Machine Learning Engineer All Questions

View all questions & answers for the Professional Machine Learning Engineer exam

Go to Exam

EXAM PROFESSIONAL MACHINE LEARNING ENGINEER TOPIC 1 QUESTION 106 DISCUSSI...

Actual exam question from Google's Professional Machine Learning Engineer

Question #: 106

Topic #: 1

[\[All Professional Machine Learning Engineer Questions\]](#)

You are building a linear regression model on BigQuery ML to predict a customer's likelihood of purchasing your company's products. Your model uses a city name variable as a key predictive component. In order to train and serve the model, your data must be organized in columns. You want to prepare your data using the least amount of coding while maintaining the predictable variables. What should you do?

- A. Use TensorFlow to create a categorical variable with a vocabulary list. Create the vocabulary file, and upload it as part of your model to BigQuery ML.
- B. Create a new view with BigQuery that does not include a column with city information
- C. Use Cloud Data Fusion to assign each city to a region labeled as 1, 2, 3, 4, or 5, and then use that number to represent the city in the model.
- D. Use Dataprep to transform the state column using a one-hot encoding method, and make each city a column with binary values.



Show Suggested Answer

by [hiromi](#) at Dec. 20, 2022, 10:03 p.m.

Comments

Type your comment...

Submit

  **fitri001** 6 months, 2 weeks ago

Selected Answer: D

A. Using TensorFlow: This is an overkill for this scenario. BigQuery ML can handle one-hot encoding natively within Dataprep.
B. Excluding City Information: This removes a potentially important predictive variable, reducing model accuracy.
C. Assigning Region Labels: This approach loses granularity and might not capture the specific variations between cities.

   upvoted 3 times

  **andresvelasco** 1 year, 1 month ago

Selected Answer: D

D by elimination but ...
Does not bigquery automatically do one-hot encoding of categorical features for you?
Also the wording of the question does not seem right: a linear regression model to predict the likelihood that the customer ...
isn't that a classification model?

   upvoted 1 times

  **M25** 1 year, 6 months ago

Selected Answer: D

Went with D

   upvoted 1 times

  **Yajnas_arpohc** 1 year, 7 months ago

Is it correct to say that A is technically a better way to do things if the ask was for separate columns?

   upvoted 1 times

  **tavva_prudhvi** 1 year, 7 months ago

"least amount of coding"

   upvoted 4 times

  **guilhermebutzke** 1 year, 7 months ago

Selected Answer: D

One-hot is a good way to use categorical variables in regressions problems
<https://academic.oup.com/rheumatology/article/54/7/1141/1849688>
<https://cloud.google.com/bigquery-ml/docs/reference/standard-sql/bigqueryml-auto-preprocessing>

   upvoted 3 times

  **TNT87** 1 year, 8 months ago

Selected Answer: D

Answer D

   upvoted 1 times

  **abneural** 1 year, 8 months ago

Selected Answer: C

for a fuller answer, D--> transforms "state" column not city column
C--> at least works with city column

   upvoted 1 times

  **tavva_prudhvi** 1 year, 7 months ago

Read smarques comment

   upvoted 1 times

  **John_Pongthorn** 1 year, 9 months ago

Selected Answer: D

<https://docs.trifacta.com/display/SS/Prepare+Data+for+Machine+Processing>

   upvoted 2 times

  **smarques** 1 year, 9 months ago

Selected Answer: D

This will allow you to maintain the city name variable as a predictor while ensuring that the data is in a format that can be used to train a linear regression model on BigQuery ML.



   upvoted 1 times

  **Abhijat** 1 year, 10 months ago

Selected Answer: D

Answer D

   upvoted 1 times

  **Abhijat** 1 year, 10 months ago

Answer is D



   upvoted 1 times

  **mymy9418** 1 year, 10 months ago

Selected Answer: D

one-hot encoding makes sense to me



   upvoted 2 times

  **hiromi** 1 year, 10 months ago

Selected Answer: C

I vote for C

   upvoted 2 times

  **hiromi** 1 year, 10 months ago

Changing my vote to D

   upvoted 3 times

Start Learning for free



Social Media

[Facebook](#) , [Twitter](#)

[YouTube](#) , [Reddit](#)

[Pinterest](#)



We are the biggest and most updated IT certification exam material website.

Using our own resources, we strive to strengthen the IT professionals community for free.



© 2024 ExamTopics

ExamTopics doesn't offer Real Microsoft Exam Questions. ExamTopics doesn't offer Real Amazon Exam Questions. ExamTopics Materials do not contain actual questions and answers from Cisco's Certification Exams.

CFA Institute does not endorse, promote or warrant the accuracy or quality of ExamTopics. CFA® and Chartered Financial Analyst® are registered trademarks owned by CFA Institute.