

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 123 DISCUSSI...

Actual exam question from Google's Professional Machine Learning Engineer

Question #: 123

Topic #: 1

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

You are developing an ML model to predict house prices. While preparing the data, you discover that an important predictor variable, distance from the closest school, is often missing and does not have high variance. Every instance (row) in your data is important. How should you handle the missing data?

- A. Delete the rows that have missing values.
- B. Apply feature crossing with another column that does not have missing values.
- C. Predict the missing values using linear regression.
- D. Replace the missing values with zeros.

Show Suggested Answer

by [pshemol](#) at Dec. 21, 2022, 10:48 a.m.

Comments

Type your comment...

Submit

[fitri001](#) 6 months, 2 weeks ago

Selected Answer: C

Preserves Information: Deleting rows (Option A) throws away valuable data, especially since every instance is important.
Not Applicable Technique: Feature crossing (Option B) creates new features by multiplying existing features. It wouldn't address missing values directly.


Zero Imputation Might Bias: Replacing missing values with zeros (Option D) can introduce bias if zeros have a specific meaning in the data (e.g., distance cannot be zero).

   upvoted 3 times

  **Voyager2** 1 year, 5 months ago

Went with A: Predict the missing values using linear regression as the data does not have high variance.



   upvoted 3 times

  **M25** 1 year, 6 months ago

Selected Answer: C

Went with C

   upvoted 1 times

  **TNT87** 1 year, 8 months ago

Selected Answer: C

Answer is C

Predicting the missing values using linear regression can be a good approach, especially if the variable is important for the prediction. The values can be imputed using regression, where the missing variable can be the dependent variable, and other relevant variables can be used as predictors

   upvoted 1 times

  **John_Pongthorn** 1 year, 9 months ago

Selected Answer: C

Regression <https://cran.r-project.org/web/packages/miceRanger/vignettes/miceAlgorithm.html>

- Find linear or non-linear relationships between the missing feature and other features

- Most advanced technique: MICE (Multiple Imputation by Chained Equations)

   upvoted 1 times

  **ares81** 1 year, 10 months ago

Selected Answer: C

It's C.

   upvoted 1 times

  **daran** 1 year, 10 months ago

My answer was based on the below article



<https://towardsdatascience.com/7-ways-to-handle-missing-values-in-machine-learning-1a6326adf79e>

   upvoted 1 times

  **daran** 1 year, 10 months ago

One of the ways to handle missing data is deleting the rows. but question here says that every row is important. so I think another possible option could be to predict the missing value. Option C could be correct !

   upvoted 2 times

  **hiromi** 1 year, 10 months ago

Selected Answer: C

C (not sure)

   upvoted 2 times

  **pshemol** 1 year, 10 months ago

Selected Answer: C

A no - Every row is important

B no - product of other feature values with no values makes no sense to me

D no - zero value would bias the model as zero distance from school has the highest value to model

C yes - there is an approach using linear regression to predict missing values

   upvoted 2 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.