EXAMTOPICS

- 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 58 DISCUSSIO..**

Actual exam question from Google's Professional Machine Learning Engineer

Question #: 58

Topic #: 1

[All Professional Machine Learning Engineer Questions]

You are working on a Neural Network-based project. The dataset provided to you has columns with different ranges. While preparing the data for model training, you discover that gradient optimization is having difficulty moving weights to a good solution. What should you do?

    A. Use feature construction to combine the strongest features.

    B. Use the representation transformation (normalization) technique.

    C. Improve the data cleaning step by removing features with missing values.

    D. Change the partitioning step to reduce the dimension of the test set and have a larger training set.

Show Suggested Answer

by 👤 ralf_cc at *July 10, 2021, 6:07 a.m.*

Comments

Type your comment…

Submit

⊟ 👤 **kurasaki** [Highly Voted 👍] 3 years, 3 months ago

Vote for B. We could impute instead of remove the column to avoid loss of information

👍 ↩ 🚩 upvoted 25 times

⊟ 👤 **pddddd** `Highly Voted 👍` 3 years ago

I also think it is B:
"The presence of feature value X in the formula will affect the step size of the gradient descent. The difference in ranges of features will cause different step sizes for each feature. To ensure that the gradient descent moves smoothly towards the minima and that the steps for gradient descent are updated at the same rate for all the features, we scale the data before feeding it to the model."

👍 ↩ 🚩 upvoted 10 times

⊟ 👤 **jsalvasoler** `Most Recent ⊘` 2 months, 2 weeks ago

`Selected Answer: B`

clearly B

👍 ↩ 🚩 upvoted 1 times

⊟ 👤 **PhilipKoku** 4 months, 2 weeks ago

`Selected Answer: B`

B) Option B (Use the representation transformation technique) is the most relevant choice. Normalizing the features will help gradient descent converge efficiently, leading to better weight updates and improved model performance.
Remember that feature scaling is crucial for gradient optimization, especially when dealing with features that have different ranges. By ensuring consistent scales, you'll enhance the effectiveness of your Neural Network training process.

👍 ↩ 🚩 upvoted 2 times

⊟ 👤 **MultiCloudIronMan** 6 months, 3 weeks ago

`Selected Answer: B`

Because the range needs to normalize

👍 ↩ 🚩 upvoted 2 times

⊟ 👤 **fragkris** 10 months, 3 weeks ago

`Selected Answer: B`

B - The key phrase is "different ranges", therefore we need to normalize the values.

👍 ↩ 🚩 upvoted 2 times

⊟ 👤 **M25** 1 year, 5 months ago

`Selected Answer: B`

Went with B

👍 ↩ 🚩 upvoted 1 times

⊟ 👤 **SergioRubiano** 1 year, 5 months ago

`Selected Answer: B`

Normalization

👍 ↩ 🚩 upvoted 1 times

⊟ 👤 **ares81** 1 year, 9 months ago

`Selected Answer: B`

Normalization is the word.

👍 ↩ 🚩 upvoted 2 times

⊟ 👤 **ares81** 1 year, 9 months ago

`Selected Answer: C`

Normalization is the word.

👍 ↩ 🚩 upvoted 1 times

⊟ 👤 **hiromi** 1 year, 10 months ago

`Selected Answer: B`

B
"Normalization" is the keyword

👍 ↩ 🚩 upvoted 1 times

⊟ 👤 **ggorzki** 2 years, 9 months ago

`Selected Answer: B`

normalization
https://developers.google.com/machine-learning/data-prep/transform/transform-numeric

👍 ↩ 🚩 upvoted 4 times

⊟ 👤 **MK_Ahsan** 2 years, 9 months ago

B. The problem does not mention anything about missing values. It needs to normalize the features with different ranges.

👍 ↩ 🚩 **upvoted 4 times**

⊟ 👤 **NamitSehgal** 2 years, 9 months ago

Looking at explanation I would choose C as well

👍 ↩ 🚩 **upvoted 1 times**

⊟ 👤 **kaike_reis** 2 years, 11 months ago

(B)
- NN models needs features with close ranges
- SGD converges well using features in [0, 1] scale
- The question specifically mention "different ranges"
Documentation - https://developers.google.com/machine-learning/data-prep/transform/transform-numeric

👍 ↩ 🚩 **upvoted 3 times**

⊟ 👤 **Y2Data** 3 years, 1 month ago

When gradient descent fails, it's out of the lacking of a powerful feature. Using normalization would make it worse.
Instead, using either A or C would increase the strength of certain feature.
But, C should come first since A is only feasible after at least 1 meaningful training.
So C.

👍 ↩ 🚩 **upvoted 2 times**

⊟ 👤 **ralf_cc** 3 years, 3 months ago

B - remove the outliers?

👍 ↩ 🚩 **upvoted 3 times**

⊟ 👤 **omar_bh** 3 years, 3 months ago

Normalization is more complicated than that.

Normalization changes the values of dataset's numeric fields to be in a common scale, without impacting differences in the ranges of values. Normalization is required only when features have different ranges.

👍 ↩ 🚩 **upvoted 4 times**

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.