EXAM PROFESSIONAL MACHINE LEARNING ENGINEER TOPIC 1 QUESTION 183 DISCUSSI...

Actual exam question from Google's Professional Machine Learning Engineer
Question #: 183
Topic #: 1
[All Professional Machine Learning Engineer Questions]

You are training an ML model on a large dataset. You are using a TPU to accelerate the training process. You notice that the training process is taking longer than expected. You discover that the TPU is not reaching its full capacity. What should you do?

A. Increase the learning rate

B. Increase the number of epochs

C. Decrease the learning rate

D. Increase the batch size

Show Suggested Answer

by pikachu007 at *Jan. 11, 2024, 11:58 a.m.*

Comments

Type your comment...

Submit

fitri001 6 months, 2 weeks ago

Selected Answer: D

A common reason for underutilized TPUs is a small batch size. TPUs are designed for high throughput, and feeding them small batches doesn't leverage their full potential.
Try increasing the batch size while monitoring model performance. A larger batch size can lead to faster training but might also affect accuracy. Experiment to find the optimal balance.

upvoted 3 times

36bdc1e 9 months, 3 weeks ago

D
taking big batch size allows to use more memory and decrease the train time

upvoted 1 times

BlehMaks 9 months, 3 weeks ago

Selected Answer: D

Batch size is too small because of sharding
https://cloud.google.com/tpu/docs/performance-guide

upvoted 1 times

pikachu007 9 months, 4 weeks ago

Selected Answer: D

D, the bigger the batch size, the more resource is taken up

upvoted 2 times