

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 273 DISCUSSI...

Actual exam question from Google's Professional Machine Learning Engineer

Question #: 273

Topic #: 1

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

You work for a rapidly growing social media company. Your team builds TensorFlow recommender models in an on-premises CPU cluster. The data contains billions of historical user events and 100,000 categorical features. You notice that as the data increases, the model training time increases. You plan to move the models to Google Cloud. You want to use the most scalable approach that also minimizes training time. What should you do?

- A. Deploy the training jobs by using TPU VMs with TPUv3 Pod slices, and use the TPUEmbedding API
- B. Deploy the training jobs in an autoscaling Google Kubernetes Engine cluster with CPUs
- C. Deploy a matrix factorization model training job by using BigQuery ML
- D. Deploy the training jobs by using Compute Engine instances with A100 GPUs, and use the tf.nn.embedding_lookup API

Show Suggested Answer

by [daidai75](#) at Jan. 8, 2024, 12:12 p.m.

Comments

Type your comment...

Submit

🗄️ 👤 **daidai75** Highly Voted 10 months ago

Selected Answer: A

TPU (Tensor Processing Units) VMs are specialized hardware accelerators designed by Google specifically for machine learning tasks.
TPUv3 Pod slices offer high scalability and are excellent for distributed training tasks.
The TPUEmbedding API is optimized for handling large volumes of categorical features, which fits your scenario with 100,000 categorical features.
This option is likely to offer the fastest training times due to specialized hardware and optimized APIs for large-scale machine learning tasks.

👍 ↩️ 🚩 upvoted 6 times

🗄️ 👤 **omer mahgoub** Most Recent 6 months, 3 weeks ago

Selected Answer: A

Addressing Bottleneck: As data size increases, CPU-based training becomes increasingly slow. TPUs are specifically designed to address this challenge, significantly accelerating training.
Large Categorical Features: TPUEmbedding API efficiently handles embedding lookups for a vast number of categorical features, a common characteristic of recommender system data.

👍 ↩️ 🚩 upvoted 3 times

🗄️ 👤 **JG123** 8 months ago

Option C

👍 ↩️ 🚩 upvoted 1 times

🗄️ 👤 **guilhermebutzke** 8 months, 2 weeks ago

Selected Answer: A

My Answer:

A: most scalable approach that also minimizes training time: TPU using TPUEmbedding API

https://www.tensorflow.org/api_docs/python/tf/tpu/experimental/embedding/TPUEmbedding

👍 ↩️ 🚩 upvoted 1 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.