

# 论文 Efficient Fault Tolerance for Recommendation Model Training via Erasure Coding 中纠删码的存储介质

集群设置。我们在AWS上对5台Type5N.8X Large服务器进行评估，每台服务器配备32个虚拟CPU、**256 GB内存**和25 Gbps网络带宽(由于内存需求，R5N.12xLarge和R5N.24xLarge分别用于大于440 GB和880 GB的DLRM)。我们使用15个P3.2xLarge类型的工作器，每个工作器配一个V100 GPU和8个虚拟CPU。以及10 Gbps的网络带宽。虽然由于成本原因，我们无法进一步扩展集群规模，但我们选择了基于现实部署的工作节点到服务器节点(例如XDL[21])。因此，我们捕获了2048个使用批次大小的工作人员的相关交易。**为了检查点，我们使用了另外15个类型为I3EN.XLarge的HDFSNode，每个节点都具有NVMe固态硬盘和25 Gbps的网络带宽。**所有实例都使用AWS ENA网络

## 嵌入表在推荐系统中的作用以及是否会频繁更新的问题

- **特征工程阶段：**
  - 对收集到的原始数据进行处理和转换，提取出有用的特征。这些特征将作为推荐系统的输入，用于训练模型。在这个阶段，可以利用嵌入表将类别信息等离散特征转换为连续的向量表示。例如，对于用户的性别特征，可以将“男”和“女”分别映射为不同的向量。这样做的好处是可以让模型更好地处理离散特征，提高模型的性能
  - **嵌入表在这个阶段的作用是将高维的离散特征映射为低维的连续向量，从而减少特征的维度，降低模型的复杂度，同时也可以提高模型的泛化能力。**
- **模型选择与构建阶段：**
  - 根据具体的业务需求和数据特点，选择合适的推荐模型。常见的推荐模型包括协同过滤、内容基于推荐、深度学习推荐等。在构建模型的过程中，**嵌入表可以作为模型的一部分，用于存储和表示用户和物品的特征向量。**
  - 例如，**在深度学习推荐模型中，嵌入表可以作为输入层的一部分，将用户和物品的 ID 等离散特征映射为连续的向量表示，然后输入到神经网络中进行训练。**嵌入表的作用是为模型提供有效的特征表示，使得模型能够更好地学习用户和物品之间的关系。
- **模型训练阶段：**
  - 使用准备好的训练数据对选择的模型进行训练。在训练过程中，模型会不断调整参数，以最小化损失函数。**嵌入表在这个阶段的作用是为模型提供初始的特征表示，并且随着模型的训练不断更新和优化。**

- 例如，在协同过滤模型中，嵌入表可以存储用户和物品的潜在因子向量，通过不断地更新这些向量，使得具有相似兴趣的用户在嵌入空间中的距离更近，从而提高推荐的准确性。

- **模型部署与应用阶段：**

- 将训练好的模型部署到实际的生产环境中，为用户提供推荐服务。在这个阶段，嵌入表可以作为模型的一部分，存储在内存或数据库中，以便快速地为用户生成推荐。

- 当用户请求推荐时，模型可以根据用户的特征和物品的特征，从嵌入表中查找相应的向量表示，然后进行计算，生成推荐结果。嵌入表的作用是为模型提供快速的特征查询和计算能力，提高推荐系统的响应速度和性能。

## **嵌入表在模型训练过程中并非频繁更新，而是呈现出稀疏更新的特点。**

- **从数据访问角度：**每个训练样本通常仅访问嵌入表中的少数条目。以在Criteo数据集上的训练为例，一批2048个训练样本平均仅访问约8900个嵌入表条目，而该数据集对应的DLRM中约有2亿个条目，访问的条目数仅占极小比例。
- **从更新频率对比角度：**与神经网络参数不同，神经网络参数在每次训练批次中都会被更新，而嵌入表条目不会每次都更新。这种差异体现了嵌入表更新的稀疏性。

# **论文 Efficient Fault Tolerance for Recommendation Model Training via Erasure Coding 被哪些论文引用**

**[SOSP23] Oobleck: Resilient Distributed Training of Large Models Using Pipeline Templates**

**[HPCA25] Revisiting Reliability in Large-Scale Machine Learning Research Clusters**

**[NAS24] TranLogs: Lossless Failure Recovery Empowered by Training Logs**

**[HotStorage24] Rethinking Erasure-Coding Libraries in the Age of Optimized Machine Learning**

**[TPDS24] Enabling Efficient Erasure Coding in Disaggregated Memory Systems**

**[preprint] Fault-Tolerant Hybrid-Parallel Training at Scale with Reliable and Efficient In-memory Checkpointing**