

PR #1762 完整报告

THUDM/slime

[Fix] Initialize grad_norm before found_inf skip path

合并时间: 2026-04-07 14:59

原文链接: <http://prhub.com.cn/THUDM/slime/pull/1762>

执行摘要

- 一句话: 修复 Megatron 训练中 found_inf=True 时 grad_norm 未初始化导致的 UnboundLocalError。
- 推荐动作: 该 PR 变更简单直接, 适合快速浏览以了解 Megatron 训练中异常处理机制。值得关注的是, 它揭示了梯度检查与训练步骤之间的耦合关系, 以及如何通过变量初始化避免边界条件错误。对于维护类似训练循环的工程师有参考价值。

功能与动机

根据 PR body 描述, 当启用 --no-check-for-nan-in-loss-and-grad 参数时, 如果 optimizer.prepare_grads() 返回 found_inf=True, 代码会跳过 optimizer.step(), 但 grad_norm 变量在该分支中未初始化, 导致 train_one_step() 函数返回时引发 UnboundLocalError。修复目的是避免训练崩溃, 确保异常处理路径的健壮性。

实现拆解

仅修改了 slime/backends/megatron_utils/model.py 文件中的 forward_step 函数。在 found_inf 分支检查之前, 添加一行代码将 grad_norm 初始化为 float("nan")。这样, 无论是否进入 found_inf 分支, grad_norm 都有定义, 避免 UnboundLocalError。

关键文件:

- slime/backends/megatron_utils/model.py (模块 Megatron 训练后端): 这是 Megatron 训练的核心文件, 包含 forward_step 函数, 负责单步训练逻辑。修复在此函数中确保 grad_norm 变量在所有路径下都有定义。

关键符号: forward_step

评论区精华

本次 PR 没有 review 评论, 因此没有讨论记录。

- 暂无高价值评论线程

风险与影响

- 风险: 风险极低。变更仅涉及单个变量的初始化, 不影响核心训练逻辑。grad_norm 被初始化为 NaN, 在 found_inf=True 时表示无效梯度范数, 符合语义。需注意: 如果其他代码依

赖 `grad_norm` 为有效数值，在 `found_inf` 路径下可能产生意外行为，但根据 PR 描述，该路径本应跳过参数更新，因此风险可控。

- 影响：对用户：修复了特定配置下的训练崩溃问题，提升系统稳定性。对系统：确保异常处理路径不会导致进程中断，支持更健壮的大规模训练。对团队：维护了 Megatron 训练模块的代码质量，避免潜在 bug。影响范围限于使用 `--no-check-for-nan-in-loss-and-grad` 参数且遇到梯度异常的用户。
- 风险标记：边界条件处理

关联脉络

- PR #1809 fix missing position_ids in log-prob forward step: 同样修改了 `slime/backends/megatron_utils/model.py` 文件，涉及 Megatron 模型前向传播的 bug 修复。
- PR #1807 sync from internal: 同样修改了 `slime/backends/megatron_utils/model.py` 文件，涉及 Megatron 模型 forward 参数构建逻辑的重构。
- PR #1775 [Fix] Fix duplicate Megatron LR scheduler resume when optimizer state is not loaded: 同样修改了 `slime/backends/megatron_utils/model.py` 文件，涉及 Megatron 训练中的 bug 修复。