

PR #1968 完整报告

THUDM/slime

[ci] add check for train_rollout_logprob_abs_diff

合并时间: 2026-05-28 13:56

原文链接: <http://prhub.com.cn/THUDM/slime/pull/1968>

执行摘要

在 CI 测试中增加对 `train_rollout_logprob_abs_diff` 指标的断言检查, 确保其值不超过 0.1。该 PR 仅修改一个文件, 增加 3 行代码, 无其他副作用。

功能与动机

为了防止训练过程中 rollout logprob 数值异常偏离导致训练不稳定或结果不可用, 通过 CI 自动捕获该指标的漂移。

实现拆解

1. 变更入口: `slime/backends/megatron_utils/model.py` 的 `train` 函数。
2. 插入断言: 在日志记录之后、原有的 KL 检查之前, 添加一段条件断言:
 - 条件: `args.ci_test` 为真 且 `"train/train_rollout_logprob_abs_diff"` in `log_dict`
 - 动作: `assert log_dict["train/train_rollout_logprob_abs_diff"] <= 0.1`, 失败时打印完整 `log_dict` 便于排查。
3. 影响范围: 仅在 CI 测试 (`args.ci_test=True`) 且日志包含该指标时生效, 完全独立于正常训练流程。

无必要源码片段, 变更仅为 3 行断言。

评论区精华

无 review 讨论。

风险与影响

- 风险: 极低, 仅影响 CI 测试, 不改变训练逻辑。阈值 0.1 可能需要根据实际训练数据微调。
- 影响: 正面——增强 CI 对训练稳定性的监控能力; 负面——无。

关联脉络

该 PR 是 CI 基础设施增强的一部分, 与近期 #1963 (修复轨迹合并逻辑) 等 CI 相关 PR 一脉相承, 持续强化自动化质量门禁。