

PR #2031 完整报告

THUDM/slime

Allow only set rollout_id for prompt that return multiple responses

合并时间: 2026-06-08 11:38

原文链接: <http://prhub.com.cn/THUDM/slime/pull/2031>

执行摘要

- 一句话: 修复多响应 prompt 的 rollout_id 分配逻辑
- 推荐动作: 值得合并, 修复逻辑清晰, 且不影响已有行为。建议后续增加针对多响应 prompt 的测试用例以覆盖此修复路径。

功能与动机

PR title 和 body 表明需要支持只为返回多个响应的 prompt 设置 rollout_id, 原实现如果 samples[0].rollout_id 不为 None 但部分 samples 未设置 rollout_id, 会导致这些 sample 的 rollout_id 仍然为 None, 破坏 rollout 聚合逻辑。

实现拆解

在 slime/ray/rollout.py 的 _convert_samples_to_train_data 方法中, 修改 rollout_ids 的生成逻辑:

1. 先收集所有 sample 的 rollout_id 到一个列表。
2. 用集合记录所有非 None 的 rollout_id。
3. 遍历列表, 对值为 None 的元素, 分配一个不与已有 ID 冲突的临时自增 ID (从 0 开始, 跳过已存在的 ID), 并加入集合。
4. 最终 rollout_ids 列表每个元素均有有效值, 后续写入 train_data 字典。

关键文件:

- slime/ray/rollout.py (模块 Rollout 引擎; 类别 source; 类型 core-logic; 符号 _convert_samples_to_train_data): 核心变更文件, 修改了 rollout_id 生成逻辑, 影响样本聚合和 loss 计算正确性。

关键符号: _convert_samples_to_train_data

关键源码片段

`slime/ray/rollout.py`

核心变更文件, 修改了 rollout_id 生成逻辑, 影响样本聚合和 loss 计算正确性。

```
# slime/ray/rollout.py (head 版本) _convert_samples_to_train_data 方法中的 rollout_id 生成部分
rollout_ids = [sample.rollout_id for sample in samples]
```

```
# 收集所有已存在的 rollout_id (非 None)
existed_rollout_id_values = set(rid for rid in rollout_ids if rid is not None)
tmp_id = 0
for i in range(len(rollout_ids)):
    if rollout_ids[i] is None:
        # 寻找一个不冲突的临时 ID
        while tmp_id in existed_rollout_id_values:
            tmp_id += 1
        rollout_ids[i] = tmp_id
        existed_rollout_id_values.add(tmp_id)
```

评论区精华

该 PR 无 review 评论，变更由作者自行合并，故无讨论记录。

- 暂无高价值评论线程

风险与影响

- 风险：风险较低：变更集中在单个方法内部，逻辑从全量判断改为逐元素处理，与外部模块无接口变化。唯一潜在风险是临时 ID 生成逻辑与其他已知 ID 冲突，但通过 `existed_rollout_id_values` 集合避免了此问题。
- 影响：影响范围有限：仅影响多响应 prompt 场景下 `rollout_id` 的生成，修复了可能的数据聚合错误。对单响应 prompt 场景行为不变。对后续 `loss reducer` 和 `reward` 归一化有间接正确性影响。
- 风险标记：缺少测试覆盖

关联脉络

- 暂无明显关联 PR