

PR #6616 完整报告

verl-project/verl

[rollout] chore: change default to calculate_log_probs=True

合并时间: 2026-06-05 12:25

原文链接: <http://prhub.com.cn/verl-project/verl/pull/6616>

执行摘要

- 一句话: rollout 默认开启 calculate_log_probs
- 推荐动作: 该 PR 简单明了, 适合合并。建议确认下游用户是否知晓此默认变更, 可在发版说明中提及。

功能与动机

PR body 仅说明 "As title", 但根据配置文件注释, 此变更支持 rollout 概率日志记录调试, 并且当启用截断重要性采样时, 需要设置 `calculate_log_probs` 为 `true`。

实现拆解

1. 在 `verl/trainer/config/rollout/rollout.yaml` 中将 `calculate_log_probs` 默认值从 `False` 改为 `True`。
2. 同步更新 4 个自动生成的 PPO 训练器配置文件: `_generated_ppo_megatron_trainer.yaml`、`_generated_ppo_torch titan_trainer.yaml`、`_generated_ppo_trainer.yaml`、`_generated_ppo_veomni_trainer.yaml`, 将各自 `calculate_log_probs` 值从 `false` 改为 `true`。
3. 每个文件仅修改一行, 确保默认行为一致。

关键文件:

- `verl/trainer/config/rollout/rollout.yaml` (模块 Rollout 配置; 类别 config; 类型 configuration) : Rollout 配置的原始 YAML 文件, 变更了 `calculate_log_probs` 默认值, 是配置的源头。
- `verl/trainer/config/_generated_ppo_trainer.yaml` (模块 训练器配置; 类别 config; 类型 configuration) : PPO 训练器自动生成的配置, 同步了默认值变更。
- `verl/trainer/config/_generated_ppo_megatron_trainer.yaml` (模块 训练器配置; 类别 config; 类型 configuration) : Megatron 训练器自动生成的配置, 同步了默认值变更。
- `verl/trainer/config/_generated_ppo_torch titan_trainer.yaml` (模块 训练器配置; 类别 config; 类型 configuration) : TorchTitan 训练器自动生成的配置, 同步了默认值变更。
- `verl/trainer/config/_generated_ppo_veomni_trainer.yaml` (模块 训练器配置; 类别 config; 类型 configuration) : VeOmni 训练器自动生成的配置, 同步了默认值变更。

关键符号: 未识别

评论区精华

无人工 review 讨论。仅有一个机器人自动 review 无反馈。

- 暂无高价值评论线程

风险与影响

- 风险：低风险。`calculate_log_probs` 开启后会增加计算开销（计算对数概率），但不会影响功能正确性。对于不需要日志概率的用户，他们仍可通过显式配置 `calculate_log_probs: False` 来关闭。
- 影响：影响用户所有使用默认配置的 PPO 训练。默认开启后，每次 rollout 会额外计算对数概率，轻微增加计算量。对于调试和截断重要性采样场景是预期行为。
- 风险标记：暂无

关联脉络

- 暂无明显关联 PR