

PR #5653 完整报告

verl-project/verl

[fully_async] chore: Add fully async dapo qwen3-30b npu script

合并时间: 2026-03-24 21:41

原文链接: <http://prhub.com.cn/verl-project/verl/pull/5653>

执行摘要

- 一句话: 新增 fully async DAPO 训练脚本, 支持 Qwen3-30B 模型在 NPU 硬件上运行。
- 推荐动作: 建议关注 review 中指出的 `overlong_buffer_len` 问题, 在使用脚本前验证配置正确性。PR 本身变更简单, 无需深入代码阅读, 但可作为示例学习配置参数设置, 并注意配置陷阱。

功能与动机

PR body 中说明 'Add fully async dapo qwen3-30b npu script', 旨在提供针对特定模型和硬件的训练脚本, 支持在 NPU 上运行 fully async 策略的实验, 可能是为了验证性能或扩展硬件兼容性。

实现拆解

实现集中于添加单个 shell 脚本文件: `verl/experimental/fully_async_policy/shell/dapo_30b_a3b_math_fsdp_npu.sh`。脚本设置环境变量 (如 `ENABLE_FULLY_ASYNC=1`)、数据文件路径 (`TRAIN_FILE`、`TEST_FILE`)、模型 ID (`MODEL_ID`)、算法参数 (如 `adv_estimator=grpo`、`max_prompt_length`、`max_response_length`) 和硬件相关配置 (如 `n_gpus_rollout=16`、`sp_size=8`)。最后调用 `python3 -m verl.experimental.fully_async_policy.fully_async_main` 启动训练命令。

关键文件:

- `verl/experimental/fully_async_policy/shell/dapo_30b_a3b_math_fsdp_npu.sh` (模块 `fully_async_policy`): 唯一变更文件, 包含所有训练配置、算法参数和硬件设置, 是 PR 功能的核心载体。

关键符号: 未识别

评论区精华

review 中主要讨论围绕 `overlong_buffer_len` 的配置错误: `gemini-code-assist[bot]` 指出该参数被设置为 `max_response_length`, 根据 `DAPORewardManager` 逻辑会导致 `expected_len=0`, 使所有有效响应都受惩罚, 影响训练奖励函数。`wucong25` 补充说明此错误可能严重干扰训练过程。此外, `psyloy` 建议删除注释并进行全量检查。这些讨论强调了配置正确性的重要性, 可能已在后续提交中修复。

- `overlong_buffer_len` 配置错误 (correctness): 问题被识别并可能在提交中修复; review 评论强调了配置正确性对训练的重要性。

风险与影响

- 风险: 主要风险是 `overlong_buffer_len` 配置错误, 可能导致奖励函数计算异常, 影响训练收敛性和实验结果准确性。脚本参数硬编码, 缺乏灵活性, 且无自动化测试覆盖, 依赖手动验证; 如果用户直接使用错误配置, 可能浪费计算资源。脚本位于 `experimental` 目录, 可能存在不稳定依赖或未充分测试。
- 影响: 影响范围限于使用此脚本进行实验的用户, 对核心系统无直接影响。影响程度中等, 因为配置错误可能导致实验失败或结果偏差, 但脚本作为示例, 可能被其他开发者参考, 错误传播风险需关注。
- 风险标记: 配置错误风险, 奖励函数异常, 缺少测试覆盖

关联脉络

- PR #5740 [misc] fix: supplement the dependencies that are missing in the `requirements-npu.txt`: 涉及 NPU 依赖补充, 与本 PR 的 NPU 脚本硬件配置相关。
- PR #5732 [model] chore: Corrected the description of errors related to the 235b script and fixed the error in running the sft script.: 类似脚本修复, 关注训练脚本中的配置错误和调试, 展示脚本维护的共通模式。
- PR #5733 [model] fix: An end-to-end script for the 235b model is provided for the 256k long sequence: 新增模型脚本, 与本 PR 同属示例脚本添加, 显示仓库中扩展硬件和模型支持的演进趋势。