

# PR #5733 完整报告

verl-project/verl

[model] fix: An end-to-end script for the 235b model is provided for the 256k long sequence

合并时间: 2026-03-24 23:04

原文链接: <http://prhub.com.cn/verl-project/verl/pull/5733>

## 执行摘要

此 PR 新增了用于 Qwen3-235B 模型的 256k 长序列端到端训练脚本，并修复了相关配置错误，扩展了模型在长上下文任务中的训练能力，同时通过 review 讨论提升了脚本健壮性。

## 功能与动机

主要动机是提供标准化的训练脚本来支持 Qwen3-235B 模型在 256k 长序列上的训练，以应对长序列推理需求。PR body 中描述为“An end-to-end script for the 235b model is provided for the 256k long sequence”，旨在简化大型模型在长上下文场景下的部署和训练流程。

## 实现拆解

- 新增脚本: `examples/grpo_trainer/run_qwen3_235b_256k_megatron_npu.sh`，配置了 256k 序列长度（如 `max_prompt_length=$((1024 * 1))` 和 `max_response_length=$((1024 * 255))`）、Megatron 并行参数（`train_tp=4`，`train_ep=32` 等）和 vLLM 配置（`gen_tp=4`，`gen_dp=32`）。
- 修改现有脚本: `examples/grpo_trainer/run_qwen3_235b_megatron_npu.sh`，删除 `actor_rollout_ref.rollout.enable_expert_parallel=True` 并设置 `actor_rollout_ref.rollout.enforce_eager=False`，优化了训练引擎行为。
- 调整数据配置: `examples/sft/gsm8k/run_qwen3_8b_sft_peft_sp2_npu.sh`，添加 `data.messages_key=messages` 和 `data.ignore_input_ids_mismatch=True`，可能用于处理特定数据格式，但需注意验证风险。

## 评论区精华

gemini-code-assist[bot] 和 wucong25 的评论焦点:

“The ++ prefix is not valid Hydra syntax... It should be +” – 纠正了关键配置错误，避免脚本执行失败。“Setting `data.ignore_input_ids_mismatch=True` disables a crucial sanity check...” – 警告了数据 tokenization 问题可能被掩盖，建议优先调查根因。建议启用 `set -xeuo pipefail`、创建日志目录和重定向错误流，以提升脚本健壮性和可调试性。这些讨论强调了配置正确性和数据验证的重要性，部分问题已修复，但数据风险需持续关注。

## 风险与影响

- 风险: 脚本中的 Hydra 语法 typo（已修复）曾可能导致训练失败；禁用数据验证（`ignore_input_ids_mismatch=True`）在 `run_qwen3_8b_sft_peft_sp2_npu.sh` 中可能引入

数据质量问题，影响模型性能；日志处理不完善（如目录缺失和错误流未重定向）可能增加调试难度。

- 影响：用户现在可以使用新脚本进行长序列训练，扩展了模型应用范围，但需谨慎处理数据配置；现有脚本的修改可能优化性能，但也可能引入不稳定因素，建议测试验证。

## 关联脉络

与历史 PR #5732 相关，后者也修改了相同脚本文件（如 `run_qwen3_235b_megatron_npu.sh`）以纠正错误，表明团队持续优化训练脚本的可靠性和功能扩展。这反映了对大型模型训练流程标准化的重视，以及跨 PR 协作维护脚本质量的趋势。