

PR #5275 完整报告

verl-project/verl

[veomni] feat: Add GRPO training scripts for Qwen3-VL-30B-MOE (VeOmni Backends)

合并时间: 2026-05-19 15:32

原文链接: <http://prhub.com.cn/verl-project/verl/pull/5275>

执行摘要

- 一句话: 新增 Qwen3-VL-30B-MOE 的 VeOmni GRPO 训练脚本
- 推荐动作: 值得精读, 特别是 VeOmni 后端用户。注意根据实际硬件调整 max_model_len 和并行配置。脚本结构清晰, 可作为自定义训练任务的起点。

功能与动机

PR body: Add NPU/GPU GRPO training scripts for Qwen3-VL-30B-MOE (VeOmni backends). 目的是扩展 VeOmni 后端对 MoE 视觉语言模型的支持。

实现拆解

1. 新增训练脚本 `examples/grpo_trainer/run_qwen3_vl_30b_moe_veomni.sh`: 包含完整的 GRPO 训练参数配置, 特别添加了 `ACTOR_VEOMNI_CONFIG` 和 `REF_VEOMNI_CONFIG` 区块, 支持参数 / 优化器卸载、全分片、序列并行 (Ulysses)、专家并行以及注意力实现选择 (`veomni_flash_attention_2_with_sp`) 和 MoE 实现 (`fused`)。通过环境变量或直接修改变量调整超参数。
2. 更新示例文档 `examples/grpo_trainer/README.md`: 在模型支持表格中为 Qwen3-VL-30B-A3B 和 Qwen3-30B-A3B 增加 VeOmni 后端标记。
3. 更新算法支持表 `docs/ascend_tutorial/model_support/model_and_algorithm_support.md`: 在末尾新增三行记录, 分别是 Qwen3-1.7B、Qwen3-30B-A3B 和 Qwen3-VL-30B-A3B-Instruct, 均使用 VeOmni 后端, 并给出示例脚本链接和硬件信息。

关键文件:

- `examples/grpo_trainer/run_qwen3_vl_30b_moe_veomni.sh` (模块 示例脚本; 类别 other; 类型 entrypoint): 核心新增脚本, 包含 VeOmni 后端训练配置与启动逻辑。
- `examples/grpo_trainer/README.md` (模块 示例脚本; 类别 docs; 类型 documentation): 更新模型支持列表, 标明 VeOmni 后端。
- `docs/ascend_tutorial/model_support/model_and_algorithm_support.md` (模块 文档; 类别 docs; 类型 documentation): 在算法支持表中新增三条使用 VeOmni 后端的模型记录。

关键符号: 未识别

评论区精华

- 安全风险: gemini-code-assist[bot] 指出脚本传递 \$@ 允许任意参数覆盖, 可导致 data.trust_remote_code=True 等危险配置, 引发远程代码执行。作者移除了 \$@ 传递, 问题已解决。
- 配置错误: bot 发现 max_model_len=1024 小于 max_prompt_length+max_response_length=2048, 会导致序列截断; 同时存在重复参数 trainer.use_legacy_worker_impl=disable。作者调整了模型长度并去除了重复参数。
- NPU 验证询问: wucong25 询问脚本是否在 Ascend 200T A2 上验证 (num_gpus_per_node=8)。作者回复 NPU 默认 1 节点 16 卡, 脚本按 GPU 写法但配置参数通用。
- 文档更新要求: wucong25 要求刷新 grpo_trainer 的 README 和模型支持列表。作者已完成相应更新。
 - 安全风险 - 参数注入 (security): 作者移除了 \$@ 传递, 问题已解决。
 - max_model_len 配置错误 (correctness): 作者调整了模型长度至足够值并去除了重复参数。
 - NPU 验证询问 (question): 作者回复 NPU 默认 1 节点 16 卡, 脚本按 GPU 写法但配置参数通用, 用户可根据实际情况调整。
 - 文档更新要求 (documentation): 作者已更新了 README 和支持表。

风险与影响

- 风险:
 - 安全风险: 尽管已移除 \$@, 脚本仍允许通过环境变量大量覆盖配置, 用户需警惕来源不明的参数设置。
 - 配置错误: 若用户未根据硬件调整 max_model_len 和并行大小, 可能导致训练失败或性能损失。
 - 重复参数风险: 脚本中曾出现重复配置, 虽已修复, 但类似问题可能再次发生, 需持续 review。
 - 缺少测试覆盖: 脚本为示例性质, 未提供单元或集成测试, 用户需自行验证正确性。
- 影响:
 - 用户: 提供了开箱即用的 VeOmni 训练示例, 降低用户使用门槛, 尤其帮助用户在 Ascend NPU 上快速启动 GRPO 训练。
 - 系统: 新增脚本作为官方示例, 需与 VeOmni 后端版本保持同步, 增加维护成本。
 - 团队: 需要后续持续跟进 VeOmni 的更新并更新文档及脚本。
 - 风险标记: 安全风险 - 参数注入, 配置错误 -max_model_len, 重复参数, 缺少测试覆盖

关联脉络

- 暂无明显关联 PR