

PR #7337 完整报告

verl-project/verl

[ci] chore: add three baselines for npu's nightly ci

合并时间: 2026-08-10 20:14

原文链接: <http://prhub.com.cn/verl-project/verl/pull/7337>

执行摘要

- 一句话: 为 NPU nightly CI 新增三个基线对比检查
- 推荐动作: 值得 CI/ 基础设施维护者快速浏览, 了解 NPU nightly 的基线对比模式。设计上简单直接, 但存在任务名称与实际脚本不一致的小问题, 建议后续清理。无需深入精读。

功能与动机

PR body 明确说明目标为“add three baselines for npu's nightly ci”, 即给 NPU 的 nightly CI 增加基线对比机制。结合此前多个 NPU CI 相关 PR (如 #7338、#7311、#7305 等), 团队在持续完善 NPU 平台的持续集成与回归检测能力, 基线对比用于自动发现训练结果偏离历史水平的情况。

实现拆解

1. 修改 .github/workflows/nightly_ascend.yml: 在现有三个 nightly 任务 (dapo-moonlight-16b-megatron-vllm、gspo-qwen3-30b-megatron-vllm、grpo-qwen3-8b-fsdp2-vllm) 的脚本执行之后, 各新增一个名为“Running nightlyCI_* checking script”的 step。每个 step 先 cd 到 /root/.cache/nightly_log/, 再执行 `python check_npu.py --log <日志路径> --base <基线文件路径>`, 日志和基线均存放在对应任务目录下。该脚本会对比本次运行日志与基线文件, 输出差异并决定 CI 是否失败。
2. 修改 tests/trainer/ppo/test_reinforce_pp_multiturn_on_cpu.py: 将 test_observation_span_does_not_block_returns 中的函数调用格式由多行改为单行, 并重新排版 test_batch_dimension 中的 mask/rewards 张量列表。这属于 pre-commit 自动格式化, 无逻辑变化。
3. 未引入新的测试、配置文件或部署配套, 基线文件 (baseline_*.txt) 由外部生成存放于 nightly log 目录, check_npu.py 的具体实现不在本次变更范围内。

关键文件:

- .github/workflows/nightly_ascend.yml (模块 CI 配置; 类别 infra; 类型 infrastructure): 核心变更文件, 为三个 nightly 任务添加基线检查 step, 是 PR 的主体。
- tests/trainer/ppo/test_reinforce_pp_multiturn_on_cpu.py (模块 单元测试; 类别 test; 类型 test-coverage): 被 pre-commit 格式化, 无逻辑变化, 但作为 PR 文件之一体现变更被清理后的最终状态。

关键符号: 未识别

关键源码片段

[.github/workflows/nightly_ascend.yml](#)

核心变更文件，为三个 nightly 任务添加基线检查 step，是 PR 的主体。

```
# .github/workflows/nightly_ascend.yml 中新增的基线检查 step 示例
# 每个任务在训练脚本执行完后，进入 nightly_log 目录运行 check_npu.py
# --log 指定本次运行日志，--base 指定预存的基线文件
- name: Running nightlyCI_dapo-moonlight-16b-megatron-vllm_ascend checking script
  run: |
    cd /root/.cache/nightly_log/
    python check_npu.py --log run_dapo_moonlight-16b_megatron_npu/run_dapo_moonlight-16b_megatron_npu.log --base run_dapo_moonlight-16b_megatron_npu/baseline_dapo_moonlight-16b_megatron_npu.txt
```

[tests/trainer/ppo/test_reinforce_pp_multiturn_on_cpu.py](#)

被 pre-commit 格式化，无逻辑变化，但作为 PR 文件之一体现变更被清理后的最终状态。

```
# tests/trainer/ppo/test_reinforce_pp_multiturn_on_cpu.py
# 格式化后的调用示例：函数调用合并为单行，语义完全不变
def test_observation_span_does_not_block_returns(self):
    """Core regression test: observation tokens must not block reward propagation."""
    config = _config(gamma=1.0)
    compact_mask = torch.ones(1, 4)
    compact_rewards = torch.tensor([[0.0, 0.0, 0.0, 1.0]])
    expanded_mask = torch.tensor([[1.0, 1.0, 0.0, 0.0, 1.0, 1.0]])
    expanded_rewards = torch.tensor([[0.0, 0.0, 0.0, 0.0, 0.0, 1.0]])

    # 对比紧凑序列与插入 observation 后的序列，合法 token 的 return 应一致
    _, compact_returns = compute_reinforce_plus_plus_outcome_advantage(compact_rewards,
        compact_mask, config=config)
    _, expanded_returns = compute_reinforce_plus_plus_outcome_advantage(expanded_rewards,
        expanded_mask, config=config)

    compact_valid = compact_returns[compact_mask.bool()]
    expanded_valid = expanded_returns[expanded_mask.bool()]
    torch.testing.assert_close(compact_valid, expanded_valid)
```

评论区精华

该 PR 无 review 评论，仅得到 wucong25 的 APPROVED 批准。整个变更过程通过 40 个 commit 反复调试（如“测试基线 -1”“新增对标 GPU 的 UT/ST”等），但最终合并时只保留了格式化后的测试文件和 CI 配置的净改动，没有公开的设计讨论或分歧记录。

- 暂无高价值评论线程

风险与影响

- 风险：

1. 基线文件缺失或过期: `check_npu.py` 依赖对应的 `baseline_*.txt` 文件, 若基线文件未同步生成或内容陈旧, 可能导致 nightly CI 误报失败。
1. 任务名与实际脚本不一致: `patch` 中 `gs-po-qwen3-30b` 任务新增 `step` 的日志路径指向 `run_gs-po_qwen3_30b_megatron_sglang_npu`, 但 `step` 名写为 `vllm`, 虽然不影响脚本执行, 但会造成后续排查混淆。
2. `check_npu.py` 的判定逻辑未知: 该脚本不在本 PR 内, 其对比阈值和失败标准未明确, 存在对比过严或过松的风险。
3. 无测试覆盖: 本变更纯粹是 CI 配置, 没有对应的本地验证方式, 错误只能在 nightly 运行中发现。- 影响: 影响范围局限于 NPU nightly CI 流水线, 为三个训练任务增加回归检测能力, 提升 CI 对训练异常 (如损失爆炸、性能下降) 的感知。对用户无直接功能影响, 对维护团队则需要在 nightly 运行前确保基线文件就绪。- 风险标记: 基线文件可用性依赖, 任务名与脚本不一致, `check_npu.py` 判定逻辑未评审

关联脉络

- PR #7338 [ci] chore: Update npu docker image cann version: 同为 NPU CI 基础设施变更, 升级 CANN 版本, 与本 PR 属于同一 NPU CI 演进线。
- PR #7311 [ci] chore: Fix docker image upload: 修复 NPU Docker 镜像上传, 与本 PR 共同完善 NPU CI 流程。
- PR #7305 [ci] chore: Update npu docker build file: 调整 NPU Docker 构建参数, 与 nightly CI 配置强相关。
- PR #7258 [ci] chore: remove RAY_DEDUP_LOGS=0 from GSPO Qwen3-8B FSDP2
NPU nightly test: 直接修改同一 `nightly_ascend.yml` 中的 `gs-po-qwen3-8b` 任务, 与本 PR 改动位置高度重合。
- PR #7238 [doc] feat: Update vllm/vllm-ascend and torch version: 涉及 NPU 环境依赖升级, 影响 nightly CI 的稳定性, 与基线对比目标一致。