

PR #5901 完整报告

verl-project/verl

[model, fsdp] fix: Fix modeling_qwen2_5_vl missing attribute 'Qwen2RMSNorm'

合并时间: 2026-05-13 17:26

原文链接: <http://prhub.com.cn/verl-project/verl/pull/5901>

执行摘要

- 一句话: 修复 Qwen2.5-VL RMSNorm 属性缺失
- 推荐动作: 值得精读, 是一个典型且优秀的兼容性修复范例。重点关注如何通过 hasattr 而非版本号比较来解决模型属性变更问题。

功能与动机

Issue #5788 报告了在 transformers 5.3.0 下 FSDP 训练时, NPU 补丁因 `modeling_qwen2_5_vl` 没有 `Qwen2RMSNorm` 属性而全部失效。PR 作者在 body 中说明需要添加版本检查来解决此问题。

实现拆解

1. 修改文件: `verl/models/transformers/npu_patch.py`, 位于 "Patches for Qwen2.5-VL Model" 区域。
2. 原代码: 直接赋值 `modeling_qwen2_5_vl.Qwen2RMSNorm.forward = rms_norm_forward_npu`, 在 transformers $\geq 5.2.0$ 时该属性不存在。
3. 修改方式: 使用 `hasattr` 检查 `modeling_qwen2_5_vl` 模块中是否存在 `Qwen2RMSNorm` 属性。如果存在 (老版本), 则 patching 它; 否则 (新版本), patching 重命名后的 `Qwen2_5_VLRMSNorm`。
4. 优势: 避免了不可靠的版本字符串比较, 直接通过属性存在性判断, 更健壮。
5. 影响范围: 仅该文件中一段代码 4 行替换, 不影响其他模型补丁。

关键文件:

- `verl/models/transformers/npu_patch.py` (模块 NPU 补丁; 类别 source; 类型 data-contract): 核心修改文件, 修复了 Qwen2.5-VL 模型 NPU 补丁的属性缺失问题。

关键符号: 未识别

关键源码片段

`verl/models/transformers/npu_patch.py`

核心修改文件, 修复了 Qwen2.5-VL 模型 NPU 补丁的属性缺失问题。

```
# 位于 verl/models/transformers/npu_patch.py
# 修改后的补丁区域: Qwen2.5-VL Model
```

```
# Patches for Qwen2.5-VL Model
# 使用 hasattr 检查属性存在性，避免版本字符串比较的陷阱
if hasattr(modeling_qwen2_5_vl, "Qwen2RMSNorm"):
    # 旧版本 transformers (如 4.57.0) 使用 Qwen2RMSNorm
    modeling_qwen2_5_vl.Qwen2RMSNorm.forward = rms_norm_forward_npu
else:
    # 新版本 transformers (如 5.3.0) 使用重命名后的 Qwen2_5_VLRMSNorm
    modeling_qwen2_5_vl.Qwen2_5_VLRMSNorm.forward = rms_norm_forward_npu

# MLP 补丁保持不变
modeling_qwen2_5_vl.Qwen2_5_VLMPLP.forward = silu_forward_npu
```

评论区精华

AI 审查者 [gemini-code-assist\[bot\]](#) 指出初始提交中使用的字符串版本比较（`transformers_version <= "5.1.0"`）不可靠，因为 Python 的字符串比较会按字典序，例如 `"10.0.0" < "5.1.0"` 为真。建议改用 `hasattr` 进行更健壮的检查。该建议被采纳，最终提交使用了 `hasattr` 方案。

- 版本号字符串比较的可靠性 (correctness): 采纳建议，最终提交使用 `hasattr(modeling_qwen2_5_vl, "Qwen2RMSNorm")` 代替版本字符串比较。

风险与影响

- 风险：风险极低。变更仅涉及一个条件判断，不存在回归风险。但需确认 `Qwen2_5_VLRMSNorm` 在所有新版本 transformers 中都存在，否则仍需额外 fallback。
- 影响：影响范围仅限于使用 transformers 新版本 ($\geq 5.2.0$) 并启用 NPU 补丁的用户。修复后这些用户可以正常进行 FSDP 训练，不会影响旧版本用户或其他模型补丁。
- 风险标记：低风险，核心路径变更

关联脉络

- PR #5788 Failed to apply NPU patches: module 'transformers.models.qwen2_5_vl.modeling_qwen2_5_vl' has no attribute 'Qwen2RMSNorm': 该 issue 是此 PR 的直接触发原因，提供了 Bug 复现和预期行为。