

PR #6351 完整报告

verl-project/verl

[ci] fix: change model_path to local cache dir

合并时间: 2026-05-14 20:51

原文链接: <http://prhub.com.cn/verl-project/verl/pull/6351>

执行摘要

- 一句话: 修复 CI 测试中模型路径使用本地缓存
- 推荐动作: 简单修复, 无需精读。但体现了在 CI 测试中使用本地缓存路径的最佳实践。

功能与动机

PR 标题和描述均为 "As title", 结合 review 评论可知, 原始路径可能无法在 CI 环境中解析, 而文件中已存在 `custom_model_prefix` 变量用于统一路径前缀, 因此修改以保持一致性并适配 CI。

实现拆解

1. 修改 `tests/utils/dataset/test_multiturn_sft_dataset_on_cpu.py` 第 242 行的 `pytest.mark.parametrize` 中 `model_path` 参数, 从 `"openai/gpt-oss-20b"` 替换为 `f"{custom_model_prefix}/openai/gpt-oss-20b"`。
2. 该变量在文件顶部定义 (例如 `custom_model_prefix = "/path/to/local/cache"`), 确保路径解析正确且符合 CI 环境。

关键文件:

- `tests/utils/dataset/test_multiturn_sft_dataset_on_cpu.py` (模块测试专用; 类别 `test`; 类型 `test-coverage`): 唯一变更文件, 修改了测试参数中的模型路径以适应 CI 本地缓存。

关键符号: `test_multiturn_sft_dataset_with_chat_template_kwargs`

关键源码片段

`tests/utils/dataset/test_multiturn_sft_dataset_on_cpu.py`

唯一变更文件, 修改了测试参数中的模型路径以适应 CI 本地缓存。

```
@pytest.mark.parametrize(
    "model_path, apply_chat_template_kwargs",
    [
        # 使用 custom_model_prefix 变量替代硬编码路径, 确保 CI 环境能正确加载本地缓存模型
        (f"{custom_model_prefix}/openai/gpt-oss-20b", {"model_identity": "You are a helpful assistant."}),
    ],
)
```

```
def test_multiturn_sft_dataset_with_chat_template_kwargs(model_path: str, apply_chat_template_kwargs: dict):  
    """Test that custom apply_chat_template_kwargs are forwarded to system prompt  
    measurement so the loss mask is not shifted when kwargs change tokenization."""  
    # ... 测试体保持不变 ...
```

评论区精华

仅有一条来自 [gemini-code-assist\[bot\]](#) 的 review 评论，建议使用 `custom_model_prefix` 而非 `os.path.expanduser`，以保持文件内一致性并避免 symlink 问题。作者采纳并提交了修改。

- 模型路径使用 `custom_model_prefix` 而非 `os.path.expanduser (style)`: 作者采纳建议，修改为 `f"{custom_model_prefix}/openai/gpt-oss-20b"`。

风险与影响

- 风险：风险极低。仅修改测试参数中的模型路径字符串，不涉及任何业务逻辑。
- 影响：影响范围仅限该测试函数，使其在 CI 中更可靠地运行。对其他用户无影响。
- 风险标记：低风险

关联脉络

- PR #6305 [data] fix: forward apply_chat_template_kwargs to system prompt measurement: 该 PR 修改了同一个测试文件 `test_multiturn_sft_dataset_on_cpu.py`，本次 PR 是在其基础上进一步修复 CI 路径问题。