

PR #44469 完整报告

vllm-project/vllm

[XPU] Fix test_logprobs_e2e import error: pin lm-eval[api]>=0.4.12

合并时间: 2026-06-17 20:26

原文链接: <http://prhub.com.cn/vllm-project/vllm/pull/44469>

执行摘要

- 一句话: 修复 XPU CI 上 lm-eval 版本不兼容导致的 ImportError
- 推荐动作: 该 PR 是常规的 CI 维护变更, 不值得精读。关注点: requirements/test/xpu.in 中使用 >= 而非固定版本, 允许小版本升级, 在 CI 中通过显式安装确保最低版本。

功能与动机

XPU CI 上 test_logprobs_e2e 测试因 lm-eval 版本不兼容而抛出 ImportError, PR 旨在通过锁定最低版本并调整 CI 流水线来修复此问题。

实现拆解

1. 修改依赖约束: 在 requirements/test/xpu.in 中将 lm_eval[api] 改为 lm_eval[api]>=0.4.12, 确保安装的版本至少为 0.4.12。
2. 调整 CI 任务: 在 .buildkite/intel_jobs/misc_intel.yaml 的 V1 Sample + Logits 任务中, 在已有命令前添加 pip install lm_eval[api]>=0.4.12, 以显式安装满足约束的依赖, 并同时启用了之前被注释掉的 test_logprobs_e2e.py 测试。
3. 锁定依赖版本: 在 requirements/test/xpu.txt 中更新了 lm-eval 的固定版本从 0.4.11 到 0.4.12, 并移除了不再需要的旧版本依赖项 (如 jsonlines、zstandard), 反映依赖解析结果。

关键文件:

- requirements/test/xpu.in (模块 测试依赖; 类别 test; 类型 test-coverage): 核心变更文件, 修改了 lm-eval 的版本约束, 从无版本要求改为 >=0.4.12。
- .buildkite/intel_jobs/misc_intel.yaml (模块 CI 配置; 类别 config; 类型 configuration): CI 配置文件, 添加了显式的 pip install 命令并启用 test_logprobs_e2e 测试。
- requirements/test/xpu.txt (模块 测试依赖; 类别 docs; 类型 documentation): 依赖锁文件的自动更新结果, 反映 lm-eval 版本从 0.4.11 升级到 0.4.12, 并移除了冗余依赖。

关键符号: 未识别

关键源码片段

[.buildkite/intel_jobs/misc_intel.yaml](#)

CI 配置文件, 添加了显式的 pip install 命令并启用 test_logprobs_e2e 测试。

```
# .buildkite/intel_jobs/misc_intel.yaml (partial)
commands:
  - >-
    bash .buildkite/scripts/hardware_ci/run-intel-test.sh
    'pip install lm_eval[api]>=0.4.12 && # 显式安装满足约束的依赖
    export VLLM_WORKER_MULTIPROC_METHOD=spawn &&
    cd tests &&
    pytest -v -s v1/logits_processors --ignore=v1/logits_processors/test_custom_online.py --
    ignore=v1/logits_processors/test_custom_offline.py &&
    pytest -v -s v1/test_oracle.py &&
    pytest -v -s v1/test_request.py &&
    pytest -v -s v1/test_outputs.py &&
    pytest -v -s v1/sample/test_topk_topp_sampler.py &&
    pytest -v -s v1/sample/test_logprobs.py &&
    pytest -v -s v1/sample/test_logprobs_e2e.py' # 恢复运行此测试
```

评论区精华

无实质性 review 讨论。Mergify 机器人多次提示 pre-commit 检查和合并冲突，但最终由审核者 [jikunshang](#) 批准。无技术争议。

- 暂无高价值评论线程

风险与影响

- 风险：低风险。变更仅限于测试依赖和 CI 配置，不影响核心代码路径。潜在风险：
lm-eval>=0.4.12 可能与某些 XPU 环境完全兼容，但已有版本锁定确保一致性。如果新版本的 lm-eval 引入 API 不兼容，可能影响依赖它的其他测试，但当前 CI 中仅有 test_logprobs_e2e 使用。
- 影响：影响范围小。仅影响 XPU CI 上的 test_logprobs_e2e 测试。其他 CI 环境和用户不会受影响。
- 风险标记：暂无

关联脉络

- 暂无明显关联 PR