PR #48513 完整报告

vllm-project/vllm

[ROCm][CI] Unblock `AMD: Language Models Test (Extended Pooling)`

合并时间：2026-07-14 02:38

原文链接：http://prhub.com.cn/vllm-project/vllm/pull/48513

执行摘要

- 一句话：跳过 ROCm 失败测试
- 推荐动作：该 PR 仅用于临时解除 CI 阻塞，不值得精读。但关联的 Issue #48510 和修复 PR #48179 值得关注，因为它们涉及 ROCm 上 Workspace 增长分配的深层问题。

功能与动机

ROCm CI 中测试 test_openai_privacy_filter 因 Workspace growth allocation after lock 而失败，该问题由 transformers 5.10 升级触发（详见 Issue #48510）。PR #48513 添加跳过标记以解除 CI 阻塞，同时关联 PR #48179 正在尝试根本修复。

实现拆解

在测试函数 test_openai_privacy_filter 前添加 @pytest.mark.skipif 装饰器，条件为 current_platform.is_rocm()，skip 原因指向 Issue #48510。仅修改 tests/models/language/pooling/test_token_classification.py 一个文件，无其他源码或配置变更。

关键文件：

- tests/models/language/pooling/test_token_classification.py（模块 测试；类别 test；类型 test-coverage）：唯一变更文件，添加了 ROCm 平台的 pytest skip 标记。

关键符号：未识别

评论区精华

无 reviewer 讨论。仅 Claude bot 自动评论说明从 fork 发起无法自动审查，最终由 AndreasKaratzas 批准合并。

- 暂无高价值评论线程

风险与影响

- 风险：风险极低：仅跳过 ROCm 平台上的单个测试用例，不影响其他测试或生产代码。唯一风险是该跳过可能被遗忘，导致 ROCm 上的回归问题漏测——此风险依赖 Issue #48510 被跟踪修复。

- 影响：影响范围极小，仅影响 ROCm CI 流水线：跳过 AMD: Language Models Test (Extended Pooling) 中的 test_openai_privacy_filter 测试。其他平台无影响。
- 风险标记：临时跳过可能遗忘

关联脉络

- PR #48179 [WIP] Potential fix for workspace growth allocation issue on ROCm: 该 PR 尝试从根本上修复 Issue #48510 中的 Workspace 增长分配问题，PR #48513 的跳过是临时措施等待该修复。
- PR #48510 [CI Failure]: AMD: Language Models Test (Extended Pooling) Workspace growth allocation after lock: 关联的 Issue，记录了该测试失败的根因和复现步骤，PR #48513 直接引用此 issue。