

PR #50447 完整报告

vllm-project/vllm

[XPU][CI] skip kimi-k3 test

合并时间: 2026-07-30 23:27

原文链接: <http://prhub.com.cn/vllm-project/vllm/pull/50447>

执行摘要

- 一句话: XPU CI 跳过 Kimi-K3 测试
- 推荐动作: 该 PR 是简单的 CI 修复, 值得快速合并以恢复 XPU CI 稳定性。但建议关联一个 issue 来跟踪未来对 Kimi-K3 的 XPU 支持, 以便及时移除该排除项。

功能与动机

PR body 指出目的是 'fix CI by skip unsupported kimi-k3 model', 即在 Intel XPU 平台上跳过对 Kimi-K3 模型的测试, 避免 CI 失败。

实现拆解

在 `tests/models/multimodal/processing/test_common.py` 文件中的 `_XPU_EXCLUDED_MODEL_IDS` 集合里添加了 "moonshotai/Kimi-K3" 条目。该集合用于在 XPU 平台上排除特定模型 ID, 避免执行不支持的测试。

关键文件:

- `tests/models/multimodal/processing/test_common.py` (模块 测试配置; 类别 test; 类型 test-coverage) : 唯一修改文件: 在 XPU 排除列表中添加 Kimi-K3 模型 ID。

关键符号: 未识别

关键源码片段

`tests/models/multimodal/processing/test_common.py`

唯一修改文件: 在 XPU 排除列表中添加 Kimi-K3 模型 ID。

```
# 文件 : tests/models/multimodal/processing/test_common.py
# 在 XPU 平台上排除不支持的模型 ID, 避免 CI 测试失败
_XPU_EXCLUDED_MODEL_IDS = {
    "baidu/Unlimited-OCR",
    "mistralai/Mistral-Large-3-675B-Instruct-2512-NVFP4",
    "moonshotai/Kimi-K3", # 新增 : 跳过 Kimi-K3 模型测试
    "Qwen/Qwen2.5-Omni-7B-AWQ",
    "thinkingmachines/Inkling-NVFP4",
}

def _get_model_ids_to_test(model_arch_list: AbstractSet[str]):
```

```
model_ids = list(_iter_model_ids_to_test(model_arch_list))
if current_platform.is_xpu():
    for excluded_model_id in _XPU_EXCLUDED_MODEL_IDS:
        while excluded_model_id in model_ids:
            model_ids.remove(excluded_model_id)
return model_ids
```

评论区精华

该 PR 没有 review 讨论。只有 Claude bot 自动评论（无实质内容）和 maintainer [DarkLight1337](#) 的批准。

- 暂无高价值评论线程

风险与影响

- 风险：风险极低。变更仅涉及测试配置，新增一个字符串条目到排除列表，不会影响任何核心逻辑或运行时行为。唯一潜在风险是未来如果 Kimi-K3 在 XPU 上得到支持，需要手动从排除列表中移除。
- 影响：影响范围极小，仅 XPU CI 流水线中的多模态处理测试：Kimi-K3 模型将不再在 XPU 上被测试，避免 CI 报错。对其他平台（如 NVIDIA、AMD）无影响。
- 风险标记：CI 稳定性

关联脉络

- PR #50000 [New model] Kimi K3: Kimi-K3 模型引入 PR，本 PR 在 XPU 上跳过该模型测试。