

PR #22137 完整报告

sgl-project/sglang

Remove flaky TestToolChoiceLfm2Moe from test_tool_choice

合并时间: 2026-04-05 13:53

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/22137>

执行摘要

- 一句话: 删除不稳定的 LFM2-MoE 模型工具调用测试类, 解决 CI flaky 问题。
- 推荐动作: 该 PR 变更简单, 适合快速浏览以了解 CI 优化策略。值得关注的是团队如何处理 flaky 测试: 通过移除不可靠的模型特定测试, 同时确保核心逻辑 (lfm2 解析器) 仍有其他测试覆盖。

功能与动机

PR body 明确指出 LiquidAI/LFM2-8B-A1B 模型在工具调用方面不可靠, 导致 `test_tool_choice_specific_function_non_streaming` 测试频繁失败 (如提供的 CI 运行链接所示)。移除该测试类是为了解决 CI flaky 问题, 确保测试稳定性。

实现拆解

仅修改一个文件 `test/registered/openai_server/function_call/test_tool_choice.py`, 完全删除 `TestToolChoiceLfm2Moe` 类 (28 行代码)。该类原本继承自 `TestToolChoiceLlama32`, 使用 lfm2 解析器, 并配置了 `flaky_tests` 集合。移除后, 测试套件中不再包含针对 LFM2-MoE 模型的工具调用测试。

关键文件:

- `test/registered/openai_server/function_call/test_tool_choice.py` (模块 `openai_server/function_call` 测试): 唯一修改的文件, 完全删除 `TestToolChoiceLfm2Moe` 测试类, 直接解决 CI flaky 问题。

关键符号: 未识别

评论区精华

review 中仅 `gemini-code-assist[bot]` 发表了一条评论, 确认 PR 移除了针对 LFM2-MoE 模型的测试类, 没有提供实质性反馈。没有其他 review 评论或争议点, 表明变更简单直接, 团队对移除 flaky 测试达成共识。

- 移除 flaky 测试的合理性 (testing): 团队默认接受移除 flaky 测试以提升 CI 稳定性。

风险与影响

- 风险：风险较低：1. 移除测试可能掩盖 LFM2-MoE 模型在工具调用方面的潜在问题，但 PR body 指出 lfm2 解析器逻辑已通过 test_function_call_parser.py 独立覆盖，降低了回归风险。2. 如果未来 LFM2-MoE 模型工具调用功能有变化，缺少专门测试可能无法及时发现兼容性问题。
- 影响：对用户无直接影响，仅影响内部 CI 稳定性。移除 flaky 测试可减少 CI 失败频率，提升开发效率。对系统无功能影响，因为测试本身不改变产品代码。团队可更专注于稳定测试，但需注意潜在模型特定问题的监控。
- 风险标记：潜在模型特定问题掩盖

关联脉络

- PR #22138 [CI]Temporary ban auto benchmark tool test: 类似地处理 CI 不稳定性问题，通过临时禁用测试解决 flaky 问题。
- PR #22100 Relax spec decoding accuracy threshold to fix flaky test: 同为修复 CI flaky 测试的策略，但通过放宽阈值而非移除测试。
- PR #21230 Add LFM2-VL (Liquid Foundation Model 2 Vision-Language) support: 添加 LFM2-VL 模型支持，与本 PR 涉及的 LFM2-MoE 模型同属 LFM2 系列，可能共享类似测试问题。