

PR #45940 完整报告

vllm-project/vllm

[Doc] Update MiniMax-M3

合并时间: 2026-06-22 16:23

原文链接: <http://prhub.com.cn/vllm-project/vllm/pull/45940>

执行摘要

- 一句话: 更新 MiniMax-M3 文档与测试注册
- 推荐动作: 该 PR 为简单文档维护, 无需精读。可作为模型文档同步的参考示例。

功能与动机

PR 标题和 body 未明确说明动机, 但从变更内容推断, 目的是为 MiniMax-M3 模型添加文档条目, 并更新其在线可用性状态, 因为之前标记为 `is_available_online=False` 的约束已被移除。

实现拆解

1. 更新测试注册配置: 在 `tests/models/registry.py` 中, 移除了两处 `is_available_online=False` 行, 分别对应 `MiniMaxM3SparseForCausalLM` 和 `MiniMaxM3SparseForConditionalGeneration` 模型。
2. 更新文档表: 在 `docs/models/supported_models.md` 中添加了两行表格条目, 分别是 `MiniMaxM3SparseForConditionalGeneration` (MiniMax-M3) 和 `MiniMaxVL01ForConditionalGeneration` (MiniMax-VL), 并注明支持文本、图像和视频输入。

关键文件:

- `tests/models/registry.py` (模块 测试注册; 类别 test; 类型 test-coverage): 移除了两处 `is_available_online=False` 标记, 表明 MiniMax-M3 模型现已可以在线获取, 影响 CI 测试行为。
- `docs/models/supported_models.md` (模块 模型文档; 类别 docs; 类型 documentation): 新增了两行文档条目, 补充了 MiniMax-M3 和 MiniMax-VL 模型的支持信息, 是用户可见的变更。

关键符号: 未识别

评论区精华

无 review 评论, 仅有 mergify bot 的自动评论: 生成文档预览和提示存在合并冲突。无实质性技术讨论。

- 暂无高价值评论线程

风险与影响

- 风险：风险极低。变更仅涉及文档和测试配置，移除 `is_available_online=False` 意味着 CI 测试会尝试拉取该模型，若模型实际不可用则可能导致测试失败，但从 PR 被批准看，模型应该已可公开获取。
- 影响：对用户的影响：文档中新增了 MiniMax-M3 和 MiniMax-VL 模型条目，方便用户了解支持情况。对系统无影响。
- 风险标记：暂无

关联脉络

- PR #46352 Temporarily skip M3 on CI: 后续 PR 临时跳过了 M3 模型的 CI 测试，与本 PR 移除了 `is_available_online=False` 有关联，可能表明该模型在 CI 中存在问题。