

PR #28072 完整报告

sgl-project/sglang

Minor refactorings to the LFM2.5 cookbook for accuracy

合并时间: 2026-06-16 05:39

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/28072>

执行摘要

- 一句话: 修正 LFM2.5 文档描述以提高准确性
- 推荐动作: 该 PR 是纯文档改进, 可以快速合并。对于技术管理者, 无需精读; 对于工程师, 如果涉及 LFM2.5 模型的集成或调试, 建议参考更新后的 cookbook 以获取准确架构信息。

功能与动机

PR 标题和 body 指出这是对 LFM2.5 cookbook 的 "Minor refactorings for accuracy", 旨在修正文档中不准确的技术描述, 例如将 "double-gated LIV convolution blocks" 更正为 "gated short convolution blocks", 并补充更精确的模型架构细节。

实现拆解

该 PR 仅修改了 3 个文档文件, 没有代码变更:

1. docs_new/cookbook/autoregressive/LiquidAI/LFM2.5.mdx (核心修改): 重写了描述开头, 修正架构术语为 "gated short convolution blocks", 增加了预训练数据量 (10T→28T 等)、块详细结构 (depthwise conv kernel size 3、GQA KV groups 8、head size 64 等)、稀疏 MoE 路由机制 (top-4 SwiGLU experts、sigmoid router 等) 以及新特性 (滑动窗口、训练后量化等)。
2. docs_new/docs/supported-models/generative_models.mdx: 更新了 LFM2 和 LFM2-MoE 在支持模型表格中的描述, 与 cookbook 新术语保持一致。
3. docs_new/docs/supported-models/multimodal_language_models.mdx: 更新了 LFM2-VL 的描述, 将 "SigLip2" 更正为 "SigLIP2 NaFlex", 并统一架构术语。

关键文件:

- docs_new/cookbook/autoregressive/LiquidAI/LFM2.5.mdx (模块 文档; 类别 other; 类型 core-logic): 核心文档, LFM2.5 cookbook 的主要修改, 包括架构描述重构和详细模型细节补充。
- docs_new/docs/supported-models/generative_models.mdx (模块 文档; 类别 other; 类型 data-contract): 更新了 LFM2 和 LFM2-MoE 在支持模型表格中的描述, 与 cookbook 新术语保持一致。
- docs_new/docs/supported-models/multimodal_language_models.mdx (模块 文档; 类别 other; 类型 data-contract): 更新 LFM2-VL 描述, 将 SigLip2 更正为 SigLIP2 NaFlex,

并统一架构术语。

关键符号：未识别

评论区精华

该 PR 没有任何 review 评论，只有作者的一条 `/tag-and-rerun-ci` 命令，以及 `gemini-code-assist` 的配额警告。合并者 `zijiexia` 直接批准。

- 暂无高价值评论线程

风险与影响

- 风险：该 PR 仅涉及文档修改，不包含任何代码变更，因此技术风险极低。唯一潜在风险是文档描述可能存在技术不准确，但该 PR 旨在提升准确性，且已由领域专家 `zijiexia` 批准。
- 影响：
 - 用户影响：阅读 LFM2.5 相关文档的用户将获得更准确的技术细节，有助于正确理解模型架构。
 - 系统影响：无。
 - 团队影响：维护者需要确保后续文档更新与当前风格一致。
 - 风险标记：暂无

关联脉络

- 暂无明显关联 PR