

PR #2492 完整报告

radixark/miles

docs: list Kimi-K2.6 as a P2P-supported model

合并时间: 2026-08-13 08:27

原文链接: <http://prhub.com.cn/radixark/miles/pull/2492>

执行摘要

PR 2492 是一行纯文档变更: 在 P2P 权重传输文档的支持模型表格中, 将 Kimi-K2.6 与 Kimi-K2 并列列为 **DeepseekV3ForCausalLM** 路径下的支持模型。变更不涉及任何源码、测试或配置, 1 个 commit 即获 approve 并合并, 对运行时无影响。

功能与动机

PR body 说明: Kimi-K2.6 已经用 P2P weight transfer 训练过, 但支持模型架构表只列出 Kimi-K2; 两者都走同一个 **DeepseekV3ForCausalLM** 路径。因此原文档会误导用户, 让读者认为 K2.6 不支持 P2P 权重传输, 需要同步文档消除信息差。

实现拆解

- 定位变更入口: docs/advanced/p2p-weight-transfer.md 的 Supported Model Architectures 表格。
- 修改 DeepseekV3ForCausalLM / DeepSeek V3 对应行的模型列: 从 Kimi-K2 (1T) * 改为 Kimi-K2 (1T) *, Kimi-K2.6 (1T)。
- 因为 K2.6 与 K2 共权重映射路径, 无需改动映射代码、示例脚本或测试; 表格行保持同一 DeepseekV3ForCausalLM 条目, 只是补全模型名列表。
- 配套情况: 无测试、无配置、无部署改动; 由 1 个 commit 完成, Zhichenzzz 以空 body approve 后合并。

关键源码片段

本次变更是纯文档表格更新, 没有源码可展示。关键表格片段如下:

```
<!-- docs/advanced/p2p-weight-transfer.md 的 Supported Model Architectures 表格摘录 -->
| 权重映射类 | 架构 | 支持模型 |
| --- | --- | --- |
| `Glm4MoeLiteForCausalLM` | GLM4-MoE | GLM-4.7-9B-Flash |
| `DeepseekV3ForCausalLM` | DeepSeek V3p2 | GLM-5 (744B-A40B) |
| `DeepseekV3ForCausalLM` | DeepSeek V3 | Kimi-K2 (1T) \*, Kimi-K2.6 (1T) |
<!-- 注: Kimi-K2.6 与 Kimi-K2 共用同一 `DeepseekV3ForCausalLM`
权重映射路径, 本次仅补充文档模型列表。 -->
```

评论区精华

本 PR 没有任何 review 评论，reviewer Zhichenzzz 直接 APPROVED（空 body）。整个讨论为空，说明变更无争议、无技术交锋。

风险与影响

- 风险：文档型变更，无运行时风险。唯一需留意的是文档与真实 P2P 支持清单的一致性——如果未来模型映射路径变化，表格可能再次漂移。另外 * 注记在表格中保留，避免破坏原有排版语义。
- 影响：影响面限定在访问 P2P 文档的用户，可正确获知 Kimi-K2.6 支持 P2P 权重传输；对训练、服务、dashboard 等模块无影响，团队工作量可忽略。

关联脉络

本 PR 属于模型支持矩阵文档维护链路：历史 PR #2356 新增 `scripts/run_kimi_k2.py`，为 Kimi-K2 提供 Python 启动脚本；PR #2354 清理 DeepSeek 相关启动脚本与文档。它们与本次 Kimi-K2.6 文档补充共同构成“模型支持清单 + 启动方式 + 文档”的一致性维护脉络。若后续维护可以引入对模型清单的自动化校验，避免此类一行文档修改反复出现。