

PR #41082 完整报告

vllm-project/vllm

[Docs] Reorganize examples docs.

合并时间: 2026-05-08 14:23

原文链接: <http://prhub.com.cn/vllm-project/vllm/pull/41082>

PR 分析报告: [Docs] Reorganize examples docs. (PR #41082)

执行摘要

本 PR 根据 RFC #29362 对 [docs/examples/README.md](#) 进行了重构, 将原来按“离线推理 / 在线服务 / 其他”三分类的导航改为按使用场景 (basic、generate、pooling、features 等) 列出 13 个分类链接。变更仅涉及一个文档文件, +14/-4 行, 无代码逻辑改动。

功能与动机

RFC #29362 指出原来的 examples 目录按 [offline_inference](#) / [online_serving](#) / [others](#) 组织, 导致用户需要跨多个目录查找特定场景的示例。本 PR 将 README 中的导航更新为按场景分类, 使用户能更快定位所需示例。

实现拆解

1. 更新 [docs/examples/README.md](#): 将原来的三段式描述替换为 13 个带链接和说明的列表项。
 - 每个列表项左侧为场景名称链接, 右侧有简要描述。
 - 场景包括: basic、generate、pooling、speech_to_text、features、reasoning、tool_calling、applications、rl、deployment、ray_serving、disaggregated、observability。

无关键源码片段。本 PR 仅涉及文档文本, 无代码层面的变更。

评论区精华

无实质性讨论。

风险与影响

- 风险: 文档链接可能指向不存在的目录, 需确保对应目录已在 PR #41084 中创建。
- 影响: 仅影响用户查看 examples 文档时的导航体验, 无功能或性能影响。

关联脉络

- 本 PR 与 PR #41084 属于同一系列, PR #41084 负责实际移动 examples 目录中的文件, 本 PR 仅更新文档说明。
- 动机来自 RFC #29362。