

PR #40266 完整报告

vllm-project/vllm

[Doc] Fix typos in token_embed pooling documentation

合并时间: 2026-04-20 10:27

原文链接: <http://prhub.com.cn/vllm-project/vllm/pull/40266>

执行摘要

本次 PR 修复了 `docs/models/pooling_models/token_embed.md` 文档中的两处小错误: 将 "outputs a embedding" 改为 "outputs an embedding" (语法修正), 以及将 "specify it via via" 改为 "specify it via" (删除重复词)。这是纯粹的文档维护, 不影响任何功能, 风险极低, 已合并。

功能与动机

根据 PR body, 作者发现文档中存在两处小错误:

- 语法错误: "outputs a embedding" 应改为 "outputs an embedding", 因为 "embedding" 以元音音素开头。
- 词语重复: "specify it via via" 中多余的 "via" 导致句子冗余。修正旨在提升文档的准确性和可读性, 属于常规维护。

实现拆解

变更仅涉及一个文件, 具体步骤如下:

1. 修正冠词使用: 在 `docs/models/pooling_models/token_embed.md` 中, 将句子 "while token embedding outputs a embedding for each token" 改为 "while token embedding outputs an embedding for each token"。
2. 删除重复词语: 在同一文件中, 将 "you need to manually specify it via via PoolerConfig(task=\"token_embed\")" 改为 "you need to manually specify it via PoolerConfig(task=\"token_embed\")"。
3. 无其他改动: 没有配套的源码、测试、配置或部署变更。

以下是修正后的关键文档片段 (注释说明变更点):

```
<!-- 修正前: The difference between the (sequence) embedding task and the token embedding task is that (sequence) embedding outputs one embedding for each sequence, while token embedding outputs a embedding for each token. ... When the default pooling task (embed) is not what you want, you need to manually specify it via via `PoolerConfig(task="token_embed")` offline or `--pooler-config.task token_embed` online. --> <!-- 修正后: The difference between the (sequence) embedding task and the token embedding task is that (sequence) embedding outputs one embedding for each sequence, while token embedding outputs an embedding for each token. // 修正: a -> an ... When the default
```

pooling task (embed) is not what you want, you need to manually specify it via
`PoolerConfig(task="token_embed")` offline or `--pooler-config.task token_embed`
online. // 修正: 删除重复的 via -->

评论区精华

review 中几乎没有技术讨论:

- gemini-code-assist[bot]评论: "This pull request corrects minor grammatical and typographical errors... I have no feedback to provide."
- noooop (维护者) 评论: "thanks for your fix!" 并批准了变更。讨论焦点仅在于确认修正内容, 无争议或未决问题。

风险与影响

- 技术风险: 无。这是纯文档修正, 不涉及代码逻辑、配置或 API, 因此没有回归、性能、安全或兼容性风险。
- 影响范围: 仅影响阅读 token_embed.md 文档的用户, 修正后文档更准确。对系统功能、团队开发或用户使用无任何影响。

关联脉络

从近期历史 PR 看, 本次 PR 是典型的文档维护 (如 PR #39782、#39079、#40189), 与其他涉及核心逻辑、性能优化或 bugfix 的 PR (如 #40283、#39083、#39120) 无直接关联。它独立于功能演进, 仅提升文档质量。