

PR #23684 完整报告

sgl-project/sglang

docs(DeepSeek-V4): note SGLANG_FIX_DSV4_BASE_MODEL_LOAD for base models

合并时间: 2026-04-25 11:27

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/23684>

执行摘要

该 PR 仅在 DeepSeek-V4 部署 cookbook 文档中追加了一小段说明，提示用户在加载 base model 时需要设置环境变量 `SGLANG_FIX_DSV4_BASE_MODEL_LOAD=1` 并使用最新代码。无代码变更，无风险，可直接合并。

功能与动机

DeepSeek-V4 文档缺少 base model 使用的相关指引，用户可能因未设置必要环境变量而导致加载失败。PR 在现有文档的“H200 image and checkpoint”小节增加了 `Base model usage` 小节，明确告知用户需要执行的操作。

实现拆解

1. 修改文件: docs_new/cookbook/autoregressive/DeepSeek/DeepSeek-V4.mdx
2. 在文件末尾附近（原文档第 147 行后）新增以下内容:

```
**Base model usage**
```

```
In order to use base models, please enable `SGLANG_FIX_DSV4_BASE_MODEL_LOAD=1` and use latest code, before the next round of testing matrix is finished.
```

该说明引用了环境变量 `SGLANG_FIX_DSV4_BASE_MODEL_LOAD`，并提醒用户在测试矩阵完善前使用最新代码以避免兼容性问题。

评论区精华

无 review 评论。仅有 Mintlify 预览部署机器人通知和 Gemini 配额提醒，无实质讨论。

风险与影响

- 风险: 无。仅文档变更，无代码逻辑改动。
- 影响: 极小。为使用 DeepSeek-V4 base model 的用户提供明确指引，改善文档完整性。

关联脉络

该 PR 是 DeepSeek-V4 文档系列更新的一部分，与 [#23634](#)、[#23622](#)、[#23628](#) 等 PR 共同完善 DeepSeek-V4 的部署指南。