

PR #36950 完整报告

sgl-project/sglang

[Docs] Restore the AIME25 label so GLM-5.3 FP8 and BF16 scores render again

合并时间: 2026-08-29 11:38

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/36950>

执行摘要

本 PR 修复了 GLM-5.3 文档页面中 AIME25 标签缺失的问题, 使 FP8 和 BF16 的 AIME25 得分能够正常渲染。变更仅涉及一个配置文件, 无代码逻辑改动, 风险极低。

功能与动机

PR 标题及提交信息明确指出, 需要恢复 AIME25 标签以让 GLM-5.3 FP8 和 BF16 得分在文档中重新渲染。此前该标签被移除或遗漏, 导致文档中对应的基准分数无法展示。

实现拆解

1. 在 docs/src/snippets/configs/zai-org/glm-5.3.jsx 文件的 benchmarkCommands.accuracy 对象中新增 aime25_pct 命令块, 内容与 aime26_pct 类似, 但运行 sgl-eval run aime25。
2. 在文件底部的 accuracyLabels 数组中插入 ["aime25_pct", "AIME25", "%"] 条目, 并调整原有 AIME26 与 GSM8K 条目的顺序以保持一致性。
3. 无测试或配置改动, 因为这是纯文档展示配置修复。

docs/src/snippets/configs/zai-org/glm-5.3.jsx

该文件是文档展示的配置来源, 本次修复的核心改动位于此, 新增 AIME25 命令块和标签。

```
// 在 benchmarkCommands.accuracy 对象中新增 AIME25 命令块
// 该命令用于运行 AIME25 评估, 并输出到指定目录
aime25_pct:
`# To install sgl-eval: pip install git+https://github.com/sgl-project/sgl-eval
sgl-eval run aime25 \
  --model {{MODEL_NAME}} --api-key <api-key> \
  --n-repeats 16 --max-tokens 64000 \
  --temperature 1.0 --top-p 0.95 --thinking \
  --out-dir /sgl-workspace/logs \
  --base-url http://{{CURL_HOST}}:{{CURL_PORT}}/v1`,

// 在 accuracyLabels 数组中注册 AIME25 标签, 使文档正确渲染对应分数
accuracyLabels: [
  ["aime26_pct", "AIME26", "%"],
  ["aime25_pct", "AIME25", "%"],
  ["gsm8k_pct", "GSM8K (1-shot)", "%"],
],
```

评论区精华

无显著 review 讨论，Mintlify 自动预览已就绪，审核人 JustinTong0323 批准通过。

风险与影响

仅修改文档配置文件，不影响任何运行时逻辑。风险极低，唯一潜在问题是命令块中的参数（如 `--n-repeats`）是否与 AIME25 评估需求匹配，但未见问题。影响范围仅限 GLM-5.3 文档页面，修复了 AIME25 分数不显示的问题，提升文档可用性。

关联脉络

与近期文档更新 PR（如 #36476）同属文档维护，但内容独立。此修复表明 GLM-5.3 基准展示的维护仍在持续。