

PR #40688 完整报告

vllm-project/vllm

[Deprecate] Deprecate LLM.reward offline api, use LLM.encode instead.

合并时间: 2026-04-24 13:37

原文链接: <http://prhub.com.cn/vllm-project/vllm/pull/40688>

执行摘要

- 一句话: 弃用 LLM.reward 离线 API, 推荐使用 LLM.encode 替代
- 推荐动作: 建议阅读 PR 的开发者关注文档更新和示例, 以理解 pooling 任务的正确使用; 同时监控下游代码是否使用 LLM.reward, 以便在 v0.23 移除前完成迁移。

功能与动机

因为 LLM.reward 离线 API 只支持 `pooling_task="token_classify"` 的令牌奖励模型, 但实际也支持序列奖励模型 (即 `pooling_task="classify"`), 导致接口混淆。参见 Issue #30312: 'Skywork Reward Model series not supported for llm.reward'。

实现拆解

1. 修改 `vllm/entrypoints/llm.py`: 在 reward 方法中添加 `logger.warning_once` 弃用警告, 并委托给 `self.encode` 固定 `pooling_task="token_classify"`; 更新 `_verify_pooling_task` 错误提示, 移除对 LLM.reward 的引用。
2. 更新测试基础设置: 修改 `tests/conftest.py` 中 `VLLMTest.reward` 方法, 直接调用 `self.llm.encode(prompts, pooling_task="token_classify")`。
3. 新增奖励模型示例: 添加 `examples/pooling/reward/sequence_reward_offline.py` 和 `sequence_reward_online.py` 展示序列奖励模型离线 / 在线用法; 重命名旧示例 `examples/basic/offline_inference/reward.py` → `examples/pooling/reward/token_reward_offline.py`, 并将 `llm.reward` 替换为 `llm.encode`; 重命名 `examples/pooling/pooling/pooling_online.py` → `examples/pooling/reward/token_reward_online.py`。
4. 更新文档: `docs/models/pooling_models/README.md` 重写离线 API 表格, 新增 Reward Usages 行; `docs/models/pooling_models/reward.md` 添加“Removed Features”节。
5. 清理测试文件: 删除 `tests/entrypoints/pooling/pooling/__init__.py`; 重命名测试文件以匹配新结构。

关键文件:

- `vllm/entrypoints/llm.py` (模块 前端入口; 类别 `source`; 类型 `core-logic`; 符号 `reward`, `_verify_pooling_task`): 核心修改: 添加弃用警告, 更新错误消息

- `examples/pooling/reward/sequence_reward_offline.py` (模块 示例代码; 类别 `source`; 类型 `core-logic`; 符号 `parse_args, main`) : 新增序列奖励模型离线示例, 展示新 API 用法
- `examples/pooling/reward/token_reward_offline.py` (模块 示例代码; 类别 `source`; 类型 `rename-or-move`; 符号 `parse_args, main`) : 重命名并更新使用 `encode`
- `examples/pooling/reward/sequence_reward_online.py` (模块 在线示例; 类别 `source`; 类型 `core-logic`; 符号 `post_http_request, parse_args, main`) : 新增序列奖励模型在线示例
- `tests/conftest.py` (模块 测试设施; 类别 `test`; 类型 `test-coverage`; 符号 `reward`) : 测试辅助方法更新
- `docs/models/pooling_models/README.md` (模块 用户文档; 类别 `docs`; 类型 `documentation`) : 更新 API 表格和示例引用
- `docs/models/pooling_models/reward.md` (模块 特性文档; 类别 `docs`; 类型 `documentation`) : 添加废弃特性记录和示例链接

关键符号: `reward, _verify_pooling_task, post_http_request, parse_args, main`

关键源码片段

[examples/pooling/reward/sequence_reward_offline.py](#)

新增序列奖励模型离线示例, 展示新 API 用法

```
def main(args: Namespace):
    # 样本提示
    prompts = [
        'Hello, my name is',
        'The president of the United States is',
        'The capital of France is',
        'The future of AI is',
    ]
    # 创建 LLM
    llm = LLM(**vars(args))
    # 关键变更: 使用 llm.encode 代替已废弃的 llm.reward, 并指定 pooling_task='classify'
    outputs = llm.encode(prompts, pooling_task='classify')
    # 打印奖励
    for prompt, output in zip(prompts, outputs):
        rewards = output.outputs.data
        print_embeddings(rewards.tolist(), prefix='Reward')
```

评论区精华

- `gemini-code-assist[bot]` 指出新表格中 `Classification` 和 `Embedding` 行列交换以及错误消息格式问题。
- `DarkLight1337` 建议使用 `"N/A"` 代替 `"nan"`, 并将弃用版本从 `v0.22` 改为 `v0.23`。
- 所有问题已在后续提交中修复。

- 文档表格列映射错误 (correctness): 已修正, 在后续提交中正确交换了列。
- 错误消息三重引号格式问题 (style): 开发者在后续提交中改为单字符串加换行。
- 弃用版本号调整 (correctness): 已修改为 v0.23。

风险与影响

- 风险:
 - 向后兼容性: LLM.reward 仍可工作但带警告, 不会破坏现有代码。
 - 文档正确性: Review 发现的表格错误已修正, 但仍需留意其他文档细节。
 - 测试覆盖: 序列奖励模型的测试依赖现有 pooling 测试框架, 可能不够充分, 但 Issue #30312 中提到的模型之前并未被测试覆盖。
- 影响:
 - 对用户: 使用 LLM.reward 的用户收到弃用警告, 需迁移到 LLM.encode(pooling_task="classify"/"token_classify")。
 - 对系统: 无性能或功能性影响。
 - 对团队: API 更加清晰, 减少混淆。
 - 风险标记: API 弃用, 向后兼容, 文档正确性

关联脉络

- 暂无明显关联 PR