

PR #2075 完整报告

radixark/miles

session: apply the trained LoRA adapter to session-server rollouts

合并时间: 2026-08-03 13:32

原文链接: <http://prhub.com.cn/radixark/miles/pull/2075>

执行摘要

- 一句话: session 转发补 lora_path, LoRA 轨迹不再 off-policy
- 推荐动作: 改动小、风险低、价值明确, 建议合入。值得关注的有两点: 一是与 compute_request_payload 的跨路径一致性做法; 二是 LORA_ADAPTER_NAME 的注册与命名机制。后续建议为 chat_completions 的请求组装补充一条单测, 覆盖 is_lora_enabled 开 / 关两种分支, 避免回归。

功能与动机

PR body 指出: session server 将 /v1/chat/completions 请求直接转发给 SGLang 且从不设置 lora_path, 而 SGLang 只在请求点名 adapter 时才应用 LoRA 权重。结果是训练中的 adapter 永远不会影响它被评分的轨迹——生成停留在冻结的基础策略上, rollout 漂移出策略, 训练行为无法回灌数据采集。提交信息强调这是 silently off-policy rather than loudly broken。

实现拆解

1. 新增导入: 在 miles/rollout/session/core.py 顶部从 miles.utils.lora 引入 LORA_ADAPTER_NAME 和 is_lora_enabled, 复用现有 LoRA 开关与 adapter 命名。
2. 注入 lora_path: 在 chat_completions 的 Phase 1 请求准备段 (设置 no_stop_trim 之后、chat_template_kwargs 处理之前) 插入条件: if is_lora_enabled(self.args): request_body["lora_path"] = LORA_ADAPTER_NAME。请求体随后被序列化为 proxy_body 转发给引擎。
3. 行为对齐与收尾: 该写法与既有非 session 生成路径 compute_request_payload 一致; 第二个 commit 仅删除捆绑示例启动器, 不影响功能。未新增自动测试, 作者以 4 节点 GLM-5.2 LoRA agentic 任务实测验证。

关键文件:

- miles/rollout/session/core.py (模块 会话服务; 类别 source; 类型 dependency-wiring; 符号 chat_completions, is_lora_enabled, LORA_ADAPTER_NAME) : session server 请求转发核心逻辑; 改动直接修正 LoRA 训练下 rollout 由基础权重服务的 off-policy 根因, 并与非 session 生成路径行为对齐。

关键符号: chat_completions

关键源码片段

miles/rollout/session/core.py

session server 请求转发核心逻辑；改动直接修正 LoRA 训练下 rollout 由基础权重服务的 off-policy 根因，并与非 session 生成路径行为对齐。

```
# miles/rollout/session/core.py —— SessionCore.chat_completions 关键片段
# 本次改动：LoRA 场景下为转发请求显式指定 adapter，避免引擎只用基础权重服务。

from miles.utils.lora import LORA_ADAPTER_NAME, is_lora_enabled

# --- Phase 1: 在 session 锁内准备请求体（后续序列化为 proxy_body 转发给引擎） ---
async with session.lock:
    # ... 已有逻辑：解析 body、fake streaming、强制 logprobs / return_meta_info ...
    # 只有显式声明 adapter，SGLang 才会应用 LoRA 权重；否则生成停留在冻结的
    # 基础策略上，训练中的 adapter 对轨迹零影响，即静默 off-policy。
    if is_lora_enabled(self.args):
        request_body["lora_path"] = LORA_ADAPTER_NAME

    # 请求级 chat_template_kwargs 覆盖渲染，随后准备 TITO input_ids
    request_messages = request_body.get("messages", [])
    prompt_token_ids = session.prepare_pretokenized(
        request_messages,
        tools=request_body.get("tools"),
        tito_tokenizer=tito_tokenizer,
    )
    request_body["input_ids"] = prompt_token_ids
    proxy_body = json.dumps(request_body).encode()
# --- 锁释放：Phase 2 无锁转发 ---
```

评论区精华

本次 review 无实质交锋：唯一 reviewer [guapisolo](#) 直接给出 LGTM，没有 review 评论。唯一 issue 评论来自 [gemini-code-assist\[bot\]](#)，声明其代码审查服务已停用，不构成技术讨论。提交信息对“静默 off-policy”风险的定性是本次值得关注的表述：这是一个不显式报错、却会让 LoRA 训练数据静默劣化的缺陷。

- 整体评审意见 (other): 无未解决疑虑，改动已合并。

风险与影响

- 风险：
 1. 请求级覆盖：注入为直接赋值而非 setdefault，若上层 agent 显式传入其他 lora_path 会被强制替换为 LORA_ADAPTER_NAME；在 session 服务 + LoRA 训练场景下符合预期，但对透传语义有一定侵入。
 2. 测试缺失：无自动测试配套，回归依赖手动验证；is_lora_enabled 的全局语义若变化，CI 无法立即捕获。

3. 兼容性: lora_path 依赖 SGLang ChatCompletionRequest 支持该字段, 升级 SGLang 时需留意。非 LoRA 路径因条件不成立完全不受影响。 - 影响: 影响范围集中在 --use-session-server 与 LoRA 训练交集的 agentic 场景: 生成从基础权重采样切换为 adapter 采样, 训练数据与策略同步, 消除静默 off-policy; 对不启用 LoRA 的运行无任何行为变化。影响程度中等, 是训练有效性的正向修复。 - 风险标记: 核心路径变更, 缺少测试覆盖, 请求级 lora_path 覆盖

关联脉络

- PR #1794 feat(multi-lora): enable and validate MoE expert adapters: 同一 LoRA 功能线, 本 PR 让 session 转发路径也能应用训练中的 adapter。
- PR #1759 (2/2) refactor(session): assemble training samples on the session server; records never leave it: session server 是训练样本组装方, 与本 PR 共同决定 LoRA 训练数据的来源与策略一致性。
- PR #1916 (1/2) refactor(rollout): drop --generate-multi-samples and its per-turn sample semantics: 近期 session/rollout 样本语义统一演进, 本 PR 是同一链路上补齐 LoRA adapter 应用。