

PR #6867 完整报告

verl-project/verl

[fully_async, doc] fix: ignore temperature config for teacher prompt_logprobs and warn when non-default value is set

合并时间: 2026-06-29 17:13

原文链接: <http://prhub.com.cn/verl-project/verl/pull/6867>

执行摘要

- 一句话: 修复 OPD 训练中 teacher 温度配置引发的崩溃
- 推荐动作: 值得快速合并, 是一个纯增益的低风险修复。可关注后续项目级日志配置清理 PR。建议阅读 `_get_teacher_sampling_params` 的注释以了解 Hydra 插值导致的配置传播问题。

功能与动机

在 OPD (在线策略蒸馏) 训练中, teacher 模型仅对已有 token 执行前向传播 (无采样), `temperature` 参数对 `prompt_logprobs` 无实际作用。默认配置通过 Hydra 插值 `${oc.select:actor_rollout_ref.rollout.temperature}` 从 student rollout 复制 `temperature`, 当 `rollout.temperature != 1.0` 时会导致 `NotImplementedError` 崩溃。

实现拆解

1. 在 `verl/experimental/teacher_loop/teacher_manager.py` 中, 将 `_get_teacher_sampling_params` 函数中的 `raise NotImplementedError` 替换为 `logger.warning`, 并始终返回 `temperature=1.0`。
2. 添加 `logging` 和 `os` 导入, 创建模块级 `logger`。
3. 在 `docs/algo/opd.md` 中补充文档, 说明 `temperature` 被强制为 1.0 的原因和机制。

关键文件:

- `verl/experimental/teacher_loop/teacher_manager.py` (模块 蒸馏; 类别 source; 类型 core-logic; 符号 `_get_teacher_sampling_params`): 核心修复文件: 将 `raise NotImplementedError` 改为 `warning`, 并强制 `temperature=1.0`。
- `docs/algo/opd.md` (模块 文档; 类别 docs; 类型 documentation): 文档更新, 解释 `temperature` 强制为 1.0 的原因和 Hydra 插值问题。

关键符号: `_get_teacher_sampling_params`

关键源码片段

`verl/experimental/teacher_loop/teacher_manager.py`

核心修复文件: 将 `raise NotImplementedError` 改为 `warning`, 并强制 `temperature=1.0`。

```
import logging
```

```

import os
from typing import Any, Optional
from uuid import uuid4

import torch
from omegaconf import DictConfig
from torch.nn import functional as F

from verl.utils.config import omega_conf_to_dataclass
from verl.workers.config import (
    DistillationConfig,
    DistillationLossConfig,
    DistillationTeacherModelConfig,
)
from verl.workers.rollout.llm_server import LLMServerClient

# 创建模块级 logger，遵循仓库惯例使用 __file__
logger = logging.getLogger(__file__)
logger.setLevel(os.getenv("VERL_LOGGING_LEVEL", "INFO"))

def _get_teacher_sampling_params(
    teacher_model_config: DistillationTeacherModelConfig,
    distillation_loss_config: DistillationLossConfig,
) -> dict[str, Any]:
    """Get sampling parameters for teacher model when computing log probabilities for
    distillation."""
    # Temperature has no effect on prompt_logprobs: the teacher performs a forward pass over
    # existing tokens (no sampling). Always use temperature=1.0 regardless of the config value.
    # The default distillation.yaml copies the student rollout temperature via Hydra interpolation
    # (temperature: ${oc.select:actor_rollout_ref.rollout.temperature}), which causes a spurious
    # crash when rollout.temperature != 1.0.
    if teacher_model_config.inference.temperature != 1.0:
        # 之前是 raise NotImplementedError，现改为警告并强制使用 1.0
        logger.warning(
            "Teacher inference temperature is set to %.1f, but temperature has no effect "
            "on prompt_logprobs (forward pass only). Using temperature=1.0.",
            teacher_model_config.inference.temperature,
        )
    num_logprobs = distillation_loss_config.topk if distillation_loss_config.loss_settings.use_topk
    else 0
    return {
        "max_tokens": 1,
        "temperature": 1.0, # 始终返回 1.0，忽略配置值
        "prompt_logprobs": num_logprobs,
    }

```

评论区精华

机器人审查建议将 `logging.getLogger(__file__)` 改为 `__name__` 以遵循标准日志层级配置。作者回应称当前代码库中有 53 个文件使用 `__file__`，项目级迁移应单独处理。最终版本保留了 `__file__`。

- logger 名称使用 `file__` 而非 `__name(style)`: 保留 `__file__`，不修改。

风险与影响

- 风险：风险极低：变更仅为将异常降级为警告并修正返回值，不影响 teacher 模型计算 `prompt_logprobs` 的正确性。但注意选择了 `__file__` 而非 `__name__`，可能破坏日志层级配置的继承（但此行为与仓库现有惯例一致）。
- 影响：影响范围小：仅影响 OPD 训练中 teacher 推理温度非 1.0 的场景。之前此类会崩溃，现在正常运行并打印警告。用户无需修改配置。
- 风险标记：低风险

关联脉络

- PR #5897 相关 teacher loop PR: PR body 中提及的相似功能 PR
- PR #6056 相关 teacher loop PR: PR body 中提及的相似功能 PR