

PR #6927 完整报告

verl-project/verl

[megatron] fix: guard ModelType.encoder_and_decoder for Megatron-Core >= 0.18 compatibility

合并时间: 2026-07-06 09:47

原文链接: <http://prhub.com.cn/verl-project/verl/pull/6927>

执行摘要

- 一句话: 兼容 Megatron-Core 0.18 移除的枚举成员
- 推荐动作: 值得快速集成。变更简洁, 逻辑清晰, 兼容性好。适合作为 Megatron-Core 版本升级的先行补丁。

功能与动机

随着 Megatron-Core 升级到 0.18, `ModelType` 枚举中移除了 `encoder_and_decoder` 成员, 导致 `get_model()` 抛出 `AttributeError`。该修复是使用 Megatron-Core 新特性 (如 `output_processor` hook) 的前提条件。

实现拆解

1. 修改导入条件判断: 将 `verl/utils/megatron_utils.py` 中三处直接引用 `ModelType.encoder_and_decoder` 替换为 `getattr(ModelType, "encoder_and_decoder", None)`。- 第 72 行: `interleaved schedule` 的断言条件。- 第 92 行: 非 `interleaved` 路径的断言条件。- 第 95 行: `if model_type == ModelType.encoder_and_decoder` 条件判断。
2. 保持逻辑不变: `getattr` 在旧版本中返回真实的枚举成员, 行为与之前一致; 在新版本中返回 `None`, 使得 `model_type` (始终为 `encoder_or_decoder`) 不等于 `None`, 断言通过且死代码分支 (`encoder_and_decoder` 处理逻辑) 被跳过。
3. 无额外测试: 变更简单, 仅 5 行新增和 3 行删除。现有测试套件可覆盖兼容性。

关键文件:

- `verl/utils/megatron_utils.py` (模块 工具函数; 类别 `source`; 类型 `core-logic`; 符号 `get_model`): 唯一修改的文件, 核心函数 `get_model()` 中三处直接引用 `ModelType.encoder_and_decoder` 改为 `getattr` 安全调用。

关键符号: `get_model`

关键源码片段

`verl/utils/megatron_utils.py`

唯一修改的文件, 核心函数 `get_model()` 中三处直接引用 `ModelType.encoder_and_decoder` 改为 `getattr` 安全调用。

```
# verl/utils/megatron_utils.py ( 关键片段 )

def get_model(
    model_provider_func,
    model_type=ModelType.encoder_or_decoder,
    wrap_with_ddp=True,
    use_distributed_optimizer=True,
    transformer_config=None,
    override_ddp_config=None,
):
    """Build the model."""
    # 使用 getattr 安全引用, 兼容 Megatron-Core 0.18+ (移除了 encoder_and_decoder)
    _unsupported = getattr(ModelType, "encoder_and_decoder", None)

    # interleaved schedule 分支
    if (
        mpu.get_pipeline_model_parallel_world_size() > 1
        and mpu.get_virtual_pipeline_model_parallel_world_size() is not None
    ):
        assert model_type != _unsupported, (
            "Interleaved schedule not supported for model with both encoder and decoder"
        )
        # ... 后续模型构建逻辑
    else:
        # 非 interleaved 分支
        assert model_type != _unsupported, (
            "Model type encoder_and_decoder is not supported"
        )
        if model_type == _unsupported:
            # 此分支在 Megatron-Core >= 0.18 不会进入
            # ... encoder_and_decoder 特殊处理
        else:
            model = model_provider_func(pre_process=pre_process, post_process=post_process)
            model.model_type = model_type
    # ...
```

评论区精华

无讨论, PR 仅有一条来自 `gemini-code-assist[bot]` 的自动评论, 无交互。maintainer `wuxibin89` 直接批准。

- 暂无高价值评论线程

风险与影响

- 风险: 风险极低。使用 `getattr` 安全降级, 新旧版本行为一致。唯一潜在风险是如果未来 Megatron-Core 重新引入同名枚举但语义不同, 但可能性极小。
- 影响: 直接影响所有使用 Megatron-Core ≥ 0.18 的用户, 使其能够正常运行 `veRL`。对旧版本用户无影响。为后续引入 `output_processor hook` 等新功能铺平道路。

- 风险标记：低风险兼容性补丁

关联脉络

- PR #6901 [megatron, perf] feat: pad BSHD micro-batches to mini-batch max seq_len: 同属 Megatron 相关改动，体现对 Megatron 后端的持续优化。
- PR #6898 [megatron, doc] docs: point Megatron Lite examples to verl-project/Megatron-LM: 同为 Megatron 模块的维护工作。