

# 复赛镜像提交说明

---

[提交说明](#)

[资源配置](#)

[镜像要求](#)

[镜像环境约定](#)

[平台介绍](#)

## 提交说明

复赛要求参赛选手提交解决方案docker镜像，推送至[阿里云容器镜像仓库](#)后，在天池提交页面中输入镜像地址、用户名和仓库密码。由比赛平台拉取镜像运行，调试阶段运行结束即可在成绩页面查询评测结果和日志，复赛阶段**每队每天可进行一次提交**，复赛开始阶段，平台方会尽量保证选手方案能够有效运行。

## 资源配置

- 处理器: 16核
- 内存: 32G
- GPU: NVIDIA A100 GPU卡 \* 1
- GPU显存: 40G
- cuda版本: 11.0 ~ 12.0

## 镜像要求

- 镜像需包含方案代码、LoRA模型文件（如果没有进行SFT则不需要LoRA模型文件）和运行代码需要的配套环境
- 封装入口脚本run.sh，请存放在/app目录下(如果镜像中没有需自行创建目录)
- 确保镜像正常运行，可使用命令：docker run your\_image 验证。
- 推送至阿里云容器镜像仓库，建议地域选择 上海。
- 使用第三方镜像时注意保证安装了curl, Dockerfile不能使用ENTRYPOINT。
- 镜像打包demo程序（包含镜像打包示例及官方提供基础镜像）：[镜像打包demo](#)

- 主办方运行选手docker镜像进行评测的demo脚本链接在：[run\\_solution\\_demo.sh](#)，选手可在本地验证所提交docker镜像是否能正常运行。

## 镜像环境约定

- 数据目录数据、赛题、基础模型，在评测过程中，会统一挂载到数据目录：`/tcdata`，选手无需在docker中打包相关文件；

```
1  # 相关数据集
2  /tcdata/bs_challenge_financial_14b_dataset
3  /tcdata/bs_challenge_financial_14b_dataset/dataset/博金杯比赛数据.db
4  /tcdata/bs_challenge_financial_14b_dataset/pdf/*
5  /tcdata/bs_challenge_financial_14b_dataset/pdf_txt_file/*
6
7  # 复赛题目
8  /tcdata/question_v2.json
9
10 # 基础模型
11 /tcdata/models
12 /tcdata/models/Qwen-7B
13 /tcdata/models/Qwen-7B-Chat
14 /tcdata/models/Qwen-7B-Chat-Int4
15 /tcdata/models/Qwen-7B-Chat-Int8
16 /tcdata/models/Tongyi-Finance-14B
17 /tcdata/models/Tongyi-Finance-14B-Chat
18 /tcdata/models/Tongyi-Finance-14B-Chat-Int4
```

- `bs_challenge_financial_14b_dataset`：modelscope下载数据集，目录结构保持一致；
- `question_v2.json`：复赛比赛题目，题目类型和数据分布与初赛保持一致；
- `models/*`：比赛相关模型
  - 包含通义千问7B模型和通义金融14B系列，总共7个模型；
  - 模型文件均从modelscope通过git lfs进行下载；
- 代码目录：`/app`，选手提供的镜像中需包含该目录：
  - 代码执行的入口文件：`/app/run.sh`
  - 代码执行的工作目录：`/app`
  - 结果文件写入的路径：`/app/submit_result.jsonl`
- 其他相关文件，如lora权重等，选手同样可以打包到 `/app` 目录；

备注：选手在本地调试过程中，请尽可能遵守约定的镜像环境设定，设置相同的挂载路径和文件组织方式，以保证选手方案能够正常运行；

## 平台介绍

- 开通阿里云容器镜像服务、构建并推送镜像，基础镜像地址等问题可参考镜像打包demo程序，如自行构建镜像请确保安装curl，打包镜像禁用ENTRYPOINT。
- 每个队伍在同一时间只能存在一个运行中的任务，上次提交的任务未运行结束再次提交会默认中断上次任务，提交时会有弹窗提示。
- 网络配置：容器内无可用网络，请注意所需软件环境直接安装进镜像，镜像中不要包含网络操作相关代码，否则镜像运行时会卡死或中断。