2022年微信大数据挑战赛-环境使用说明

### 赛制说明

比赛分为两个阶段：

1. **预热阶段（7月1号至7月4号）**
2. **正式比赛阶段（7月5号至8月5号）**

#### 预热阶段：

预热阶段的目的是让选手熟悉比赛环境和比赛的提交流程。

该阶段仅为选手提供 CPU 计算资源，无 GPU 计算资源。此外，该阶段中不提供正式比赛的训练数据，仅有100条 demo 数据供选手调试代码。

该阶段开放外网，选手可在该阶段下载比赛过程中需要的 python 包，yum 包、开源的预训练模型等等。在比赛正式开始后，平台将禁止**除了腾讯源（**[mirrors.tencentyun.com](http://mirrors.tencentyun.com/)**）以外的外网访问权限**。

#### 正式阶段：

正式比赛阶段，大赛为选手提供18核80G V100\*2 （单卡32G）的 notebook 训练资源。

正式比赛阶段将限制外网访问权限。

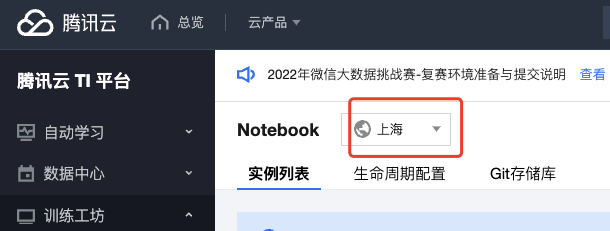
在正式比赛阶段，如还需下载某些未在预热阶段准备好的预训练模型。请联系组委会帮忙协助下载。

### 环境说明

#### 1. 进入 Notebook 实例

复赛的平台为腾讯云 Ti-ONE 的 notebook 环境。关于 Notebook 的介绍可以参阅官方的介绍文档：[腾讯云 TI 平台 TI-ONE Notebook 简介-操作指南-文档中心-腾讯云](https://cloud.tencent.com/document/product/851/74127)

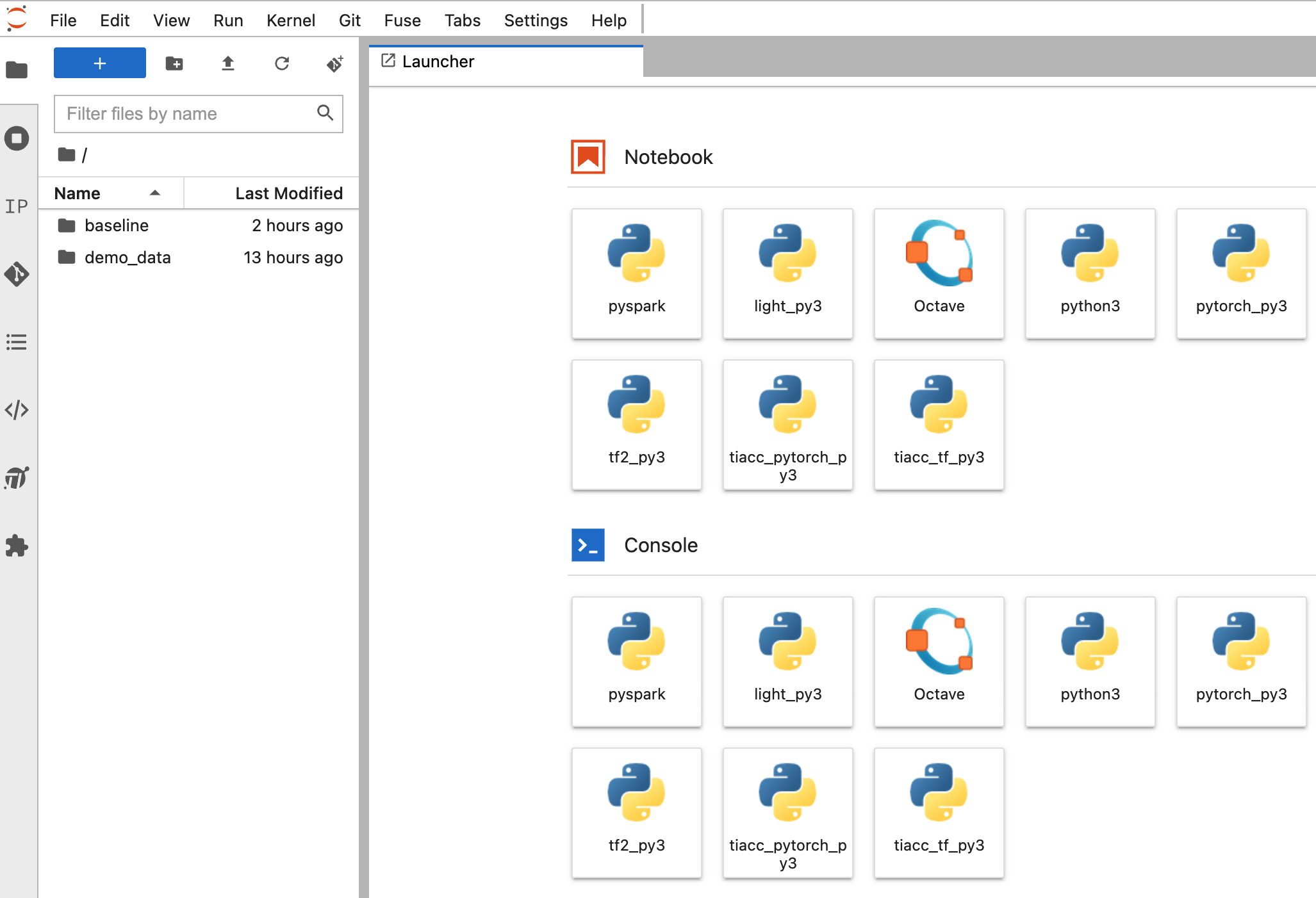
注意切换地域到上海



在登录 Ti-ONE 控制台（[登录 - 腾讯云](https://console.cloud.tencent.com/tione/v2/notebook/list?listTab=instance)）后，可以在左侧的【Notebook】中进入到notebook 实例列表页面。赛事官方已经为各队参赛选手准备好了 notebook 环境，无需（同时也禁止）选手自行创建。在此页面应该能够看见官方准备好的 notebook 实例。

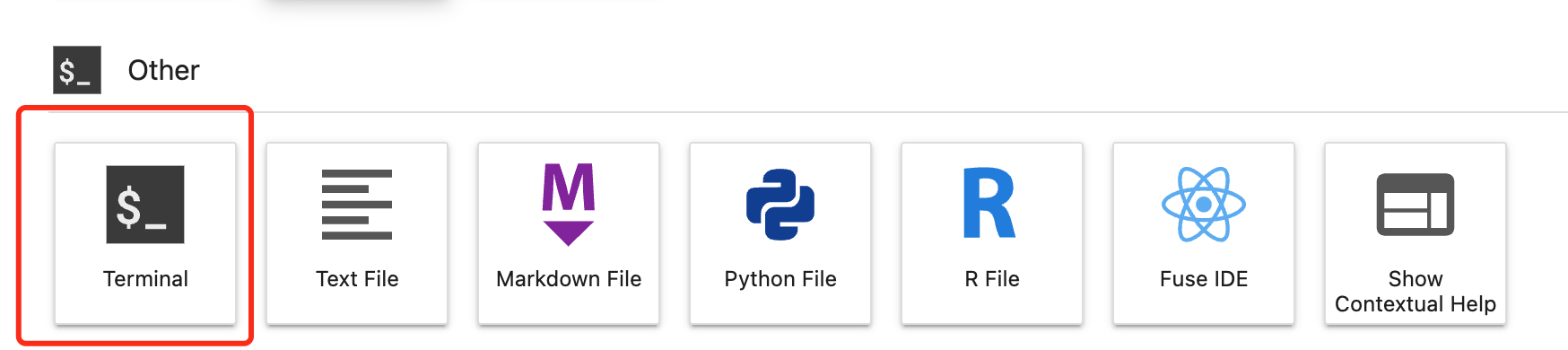
若实例的状态为【已停止】，请点击【启动】按钮；

若实例的状态为【运行中】，可以直接点击【打开】按钮进入到 notebook 中。页面如下所示：



#### 2. 运行环境设置

Notebook 中可以使用 Terminal 窗口来执行命令：



目前，已经为选手准备好了多个基础的运行环境。可以用如下命令进行查看：

conda env list

结果如下：

# conda environments:

#

base /opt/conda

JupyterSystemEnv \* /opt/conda/envs/JupyterSystemEnv

light\_py3 /opt/conda/envs/light\_py3

pyspark /opt/conda/envs/pyspark

python3 /opt/conda/envs/python3

pytorch\_py3 /opt/conda/envs/pytorch\_py3

tf2\_py3 /opt/conda/envs/tf2\_py3

tf\_py3 /opt/conda/envs/tf\_py3

tiacc\_pytorch\_py3 /opt/conda/envs/tiacc\_pytorch\_py3

tiacc\_tf\_py3 /opt/conda/envs/tiacc\_tf\_py3

如果已有环境能够满足所需的环境依赖，则不用进行环境安装。直接激活已有的环境（例如 source activate pytorch\_py3）或者打开某个环境的Jupyter Notebook即可，可以跳过后面的操作指引。

如果已有的conda环境无法满足需求，则需要创建一个新的conda环境。注意不要在平台提供的conda环境上进行安装，因为其位于/opt/conda/envs目录下，当Notebook实例重启时环境会丢失。具体操作步骤如下：

1. **使用conda create命令来创建所需的conda环境**。需要注意的是，只能使用--prefix模式指定安装路径来创建，而不能用--name的方式创建。因为后者创建的环境会默认保存在/opt/conda/envs目录下，当Notebook实例重启时环境会丢失。

conda create --prefix /home/tione/notebook/wbdc2022\_demo\_env -y --clone pytorch\_py3

1. **使用source activate激活新建的conda环境**

source activate /home/tione/notebook/wbdc2022\_demo\_env

1. **在激活的环境中，安装需要的包。** 大家在创建完环境之后用pip安装之前，最好先which pip 看看安装的是哪一个pip。只有pip路径是创建的conda环境才能安装到这个路径下，否则可能会安装到系统默认路径下，这样重启的时候就没有了。 或者在安装的时候也可以指定pip进行安装，例如 /home/tione/notebook/wbdc2022\_demo\_env/bin/pip install coscmd
2. **重启修复。**notebook 在停止、重启后，需要重新载入之前创建的配置。例如，**从预热阶段切换到正式阶段的时候**，notebook 就会进行重启。实例启动成功后，打开实例，运行如下命令，将之前创建的环境加到conda配置中。

conda config --add envs\_dirs /home/tione/notebook/env/wbdc2022\_demo\_env

（注：正式阶段与预热阶段的文件挂载点略有不同，在正式比赛阶段访问预热阶段的文件，需要加上 env/ 前缀。）

### 运行与调试

选手在 notebook terminal 中，可自行调试代码。其使用方法与常规 Linux 主机类似。

完成训练后，docker 提交可以参阅：

[2022年微信大数据挑战赛-容器镜像提交说明](https://doc.weixin.qq.com/doc/w3_AKAAOwaBACsk5e5dhkcQw2T4Ymbo9?scode=AJEAIQdfAAosVzZlUoAKAAOwaBACs)