



ML-Master使用教程



ML-Master.zip

11.10MB



前言


本文档旨在引导同学们成功运行我们提供的ML-Agent Demo，熟悉MLE-Bench，在成功运行后，同学们可以从多方面对demo进行改进，包括但不限于提示词工程，算法优化和性能优化。同学们也可以根据自己的想法构建自己的ML-Agent框架。

使用平台说明

由于许多python库对于windows系统并没有较好的兼容，我们推荐使用windows系统的同学使用虚拟机或云服务器，如果使用windows在本地开发，需要手动解决一些代码不兼容windows的问题。对于云服务器，这里提供一个可以供个人开发者免费使用云服务器的渠道供参考；对于虚拟机，请自行搜索相关教程。

Modelscope notebook

Modelscope notebook服务为个人开发者提供了可免费使用的8核32G云服务器，同时每个用户有免费的36小时GPU服务器使用时长。网址为<https://modelscope.cn/my/mynotebook>，在首页-我的Notebook选择PAI-DSW可以启动不限时长的CPU环境。



ModelScope

首页

模型库

数据集

创空间

文档

社区


MCP广场

AIGC专区

科学智能

🔍 搜索您感兴趣的内容

🌙 ⚙️ 🔄



个人主页

社区勋章全面上线
勋章送好礼，快来加入~

立即加入

🏠 概览

👤 我创建的

💎 我喜欢的

📁 我的Notebook

模型服务 免费试用

🏷️ 访问令牌

👤 组织信息

👤 暂未属于任何组织

👤 相关推荐

📄 精选文章

🗨️ 社区论坛

🐙 GitHub

我的Notebook

ModelScope社区与阿里云合作，Notebook功能由阿里云提供算力支持

魔搭平台免费实例

个人云账号授权实例

数据迁移 将数据迁移至您的个人账号授权实例中

PAI-DSW 阿里云弹性加速计算EAS

PAI-DSW是为算法开发者量身打造的云端深度学习开发环境，内置JupyterLab、WebIDE及Terminal，无需任何运维配置即可编写、调试及运行Python代码。[了解更多产品详情](#)

方式一

方式二

CPU环境

8核 32GB

预装 ModelScope Library

预装镜像 ubuntu22.04-py311-torch2.3.1-1.31.0

长期使用

GPU环境

8核 32GB 显存24G (更高阶卡型需求，可使用个人云账号授权实例)

预装 ModelScope Library

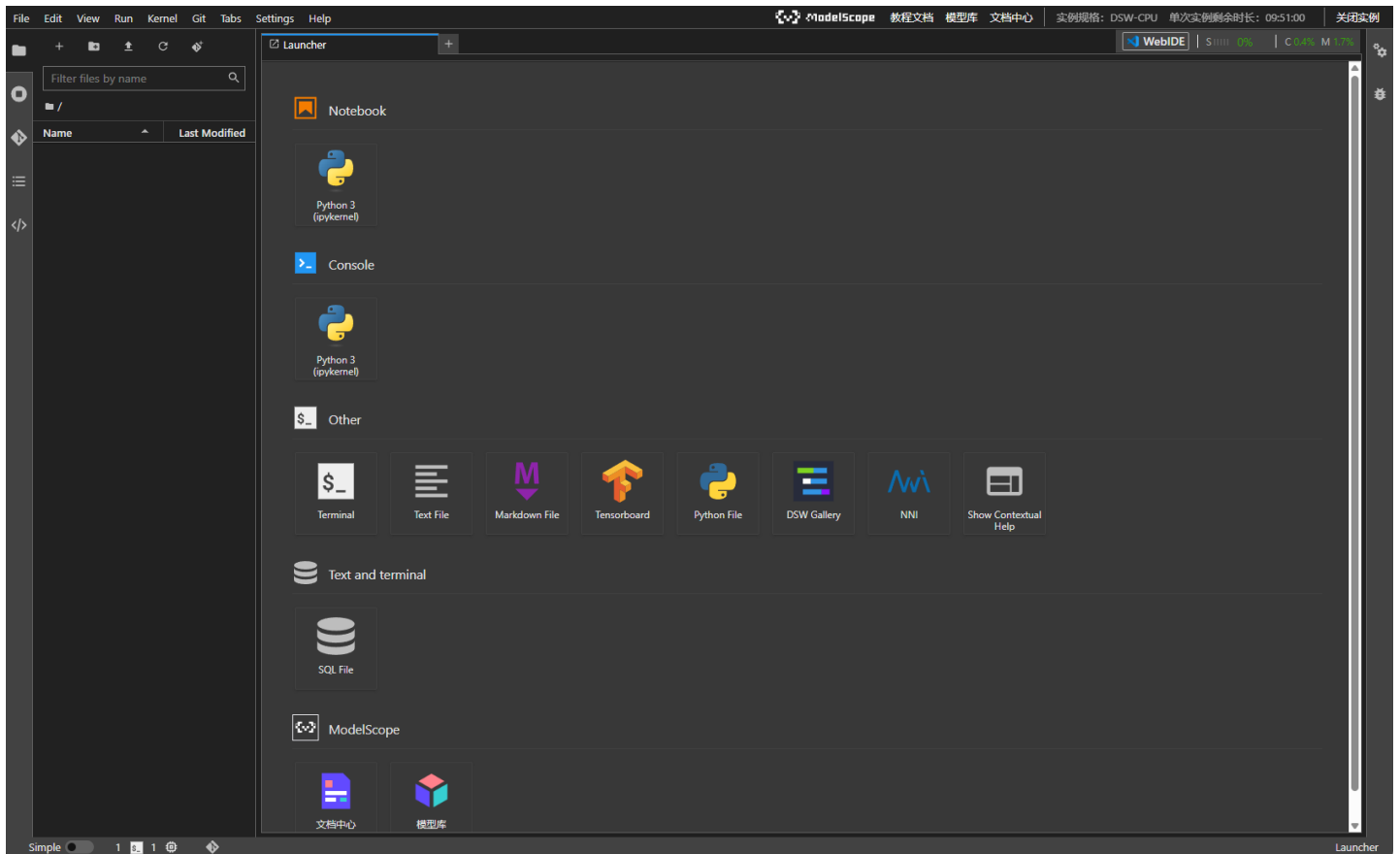
预装镜像 ubuntu22.04-cuda12.4.0-py311-torch2.8.0-1.31.0-LLM

剩余额度36小时00分钟

启动

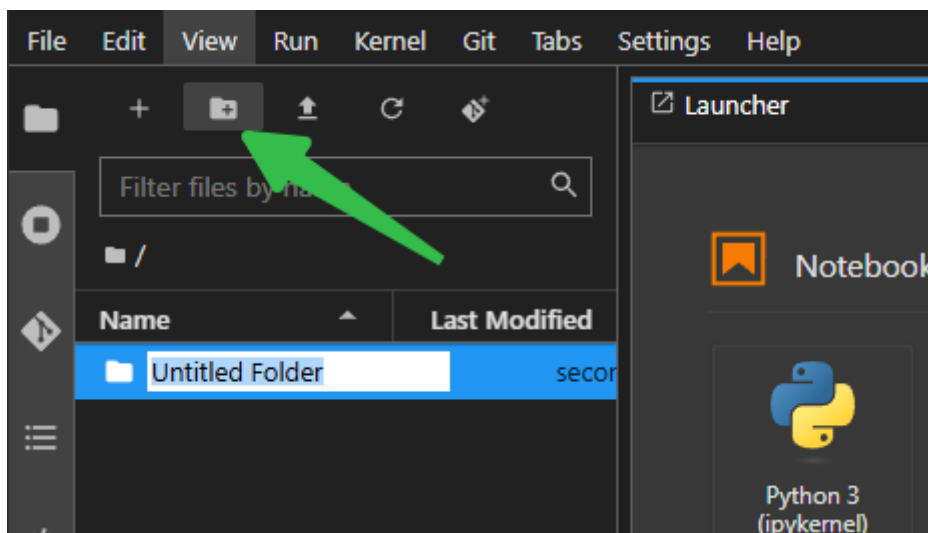
Linux

以modelscope notebook为例，启动后界面如下，默认基础目录为/mnt/workspace

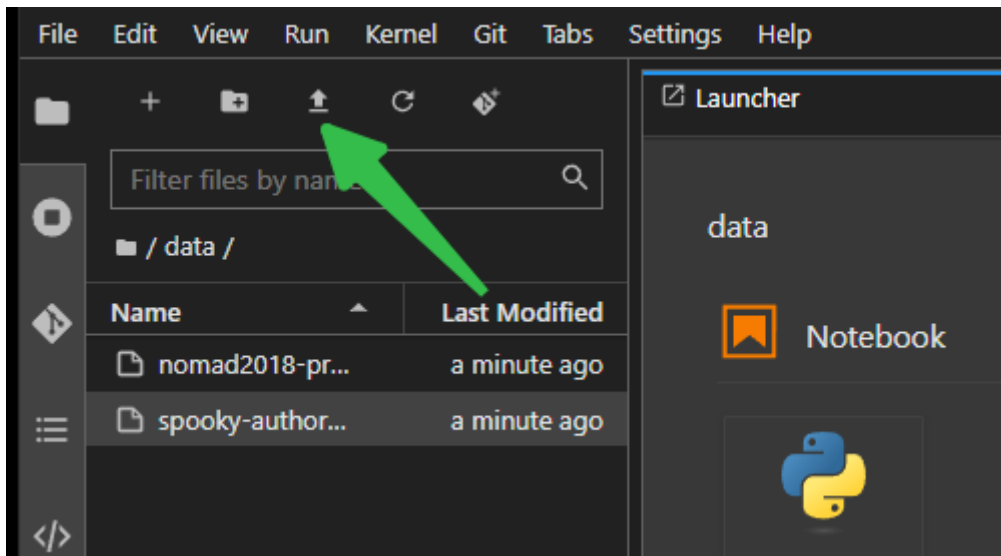


数据上传

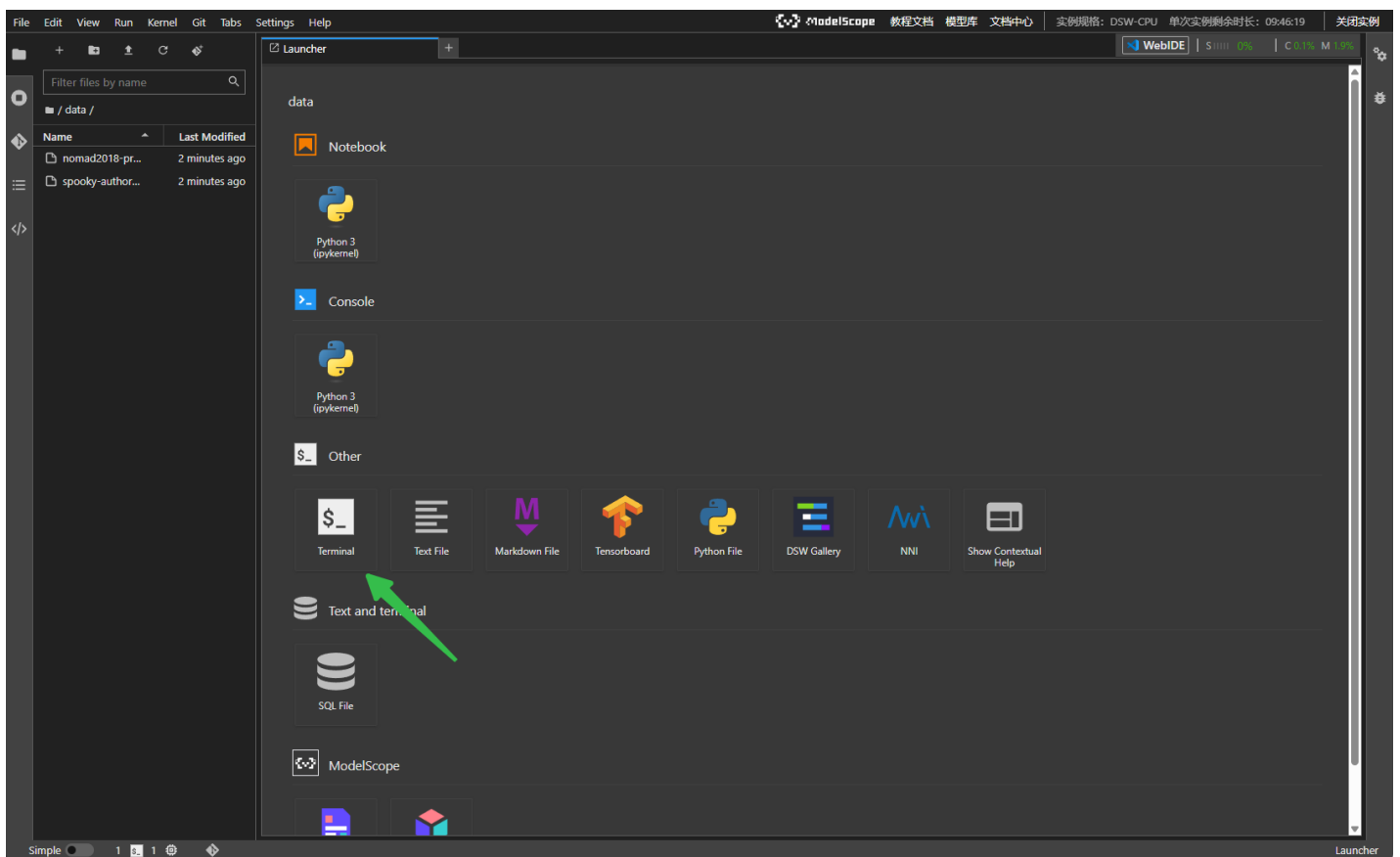
点击新建文件夹按钮，创建一个文件夹命名为data，双击进入



通过上传文件按钮，将提供的示例数据集文件nomad2018-predict-transparent-conductors.zip和spooky-author-identification.zip上传到data目录下，



点击terminal新建终端

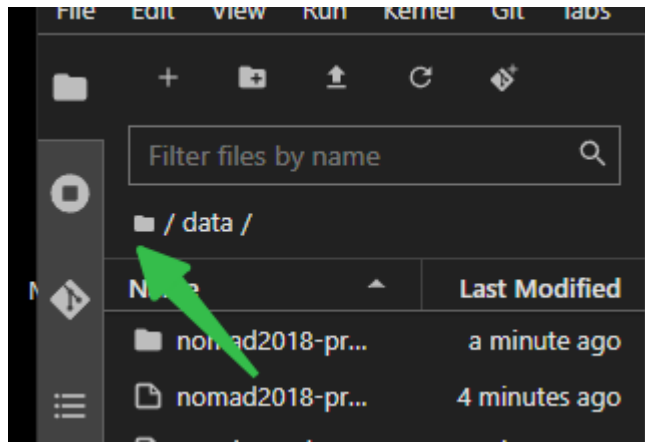


并依次运行解压命令

代码块

```
1 cd /mnt/workspace/data
2 unzip nomad2018-predict-transparent-conductors.zip
3 unzip spooky-author-identification.zip
```

点击回退按钮回退到上级目录



克隆github上的ML-Master项目 (<https://github.com/sjtu-sai-agents/ML-Master.git>)

代码块

```
1 cd /mnt/workspace/  
2 git clone https://github.com/sjtu-sai-agents/ML-Master.git
```

如遇到网络问题等导致git clone失败，可以使用手动将demo文件ML-Master.zip上传到默认目录/mnt/workspace下，并依次运行解压命令（如git clone成功则可跳过此步骤）

代码块

```
1 cd /mnt/workspace/  
2 unzip ML-Master.zip
```

切换到feature-dev分支(如果是上传zip解压可跳过此步骤)

代码块

```
1 cd ML-Master  
2 git checkout feature-dev
```

环境安装

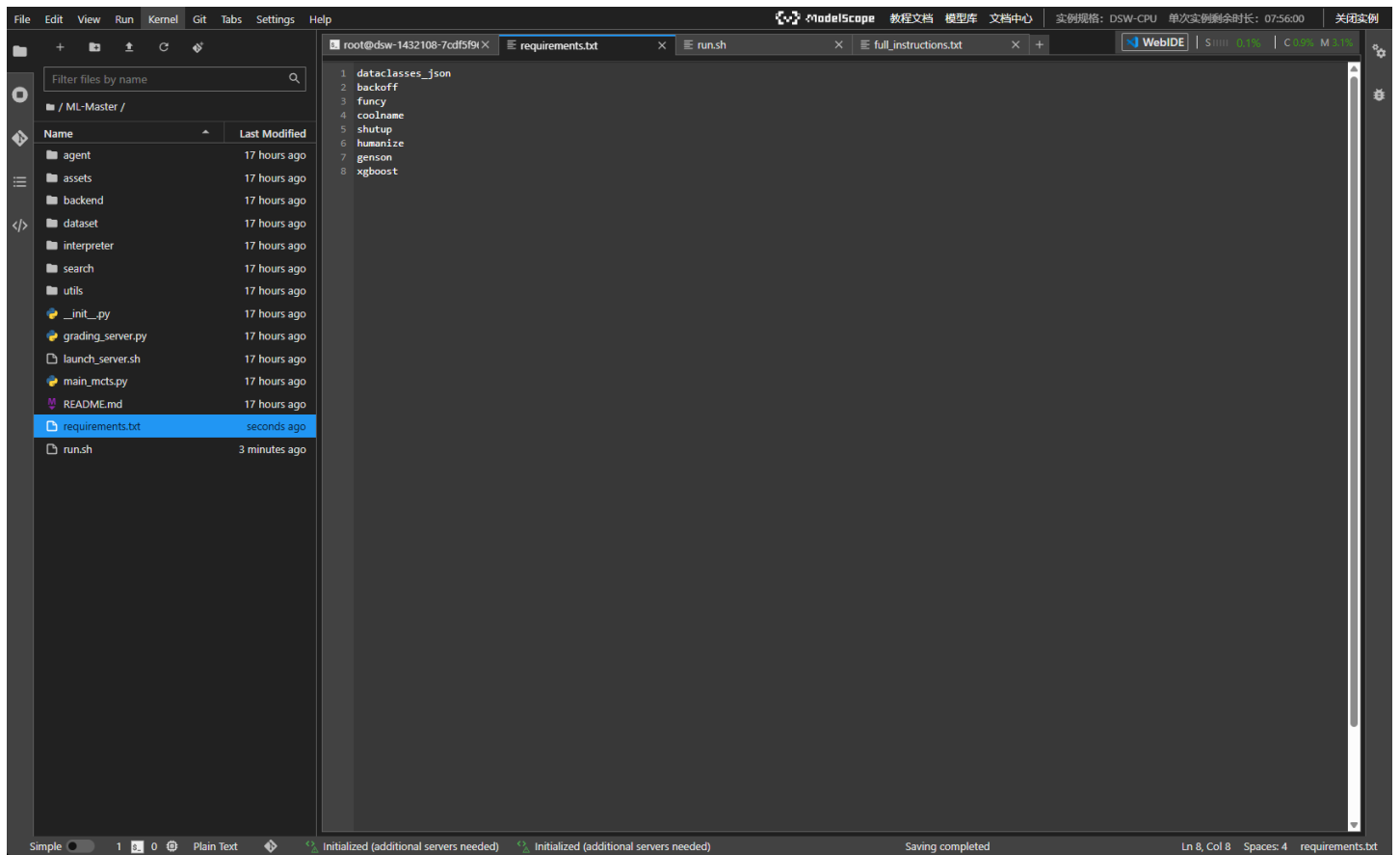
Modescope notebook镜像已经预置了大量环境，只需要额外安装部分环境即可

双击进入解压后的ML-Master文件夹，打开requirements.txt，将里面内容清空，替换成我们需要额外安装的环境并保存

代码块

```
1 dataclasses_json  
2 backoff  
3 funcy  
4 coolname
```

```
5  shutup
6  humanize
7  genson
8  xgboost
```



然后在终端运行

代码块

```
1  cd /mnt/workspace/ML-Master
2  pip install -r requirements.txt
```

安装完成如图所示


```
1 AGENT_DIR=/mnt/workspace/ML-Master
代码块
2 EXP_ID=nomad2018-predict-transparent-conductors #可根据任务名称修改
3 dataset_dir=/mnt/workspace/data
4 MEMORY_INDEX=0
5
6 # ...中间省略
7
8 CPUS_PER_TASK=8
```

修改提示词

由于我们使用modelscope notebook cpu版本进行进行demo演示，所以我们需要修改部分提示词，告诉agent我们的环境没有配备gpu。（如启动gpu版本Notebook可以免去此步骤）

打开目录

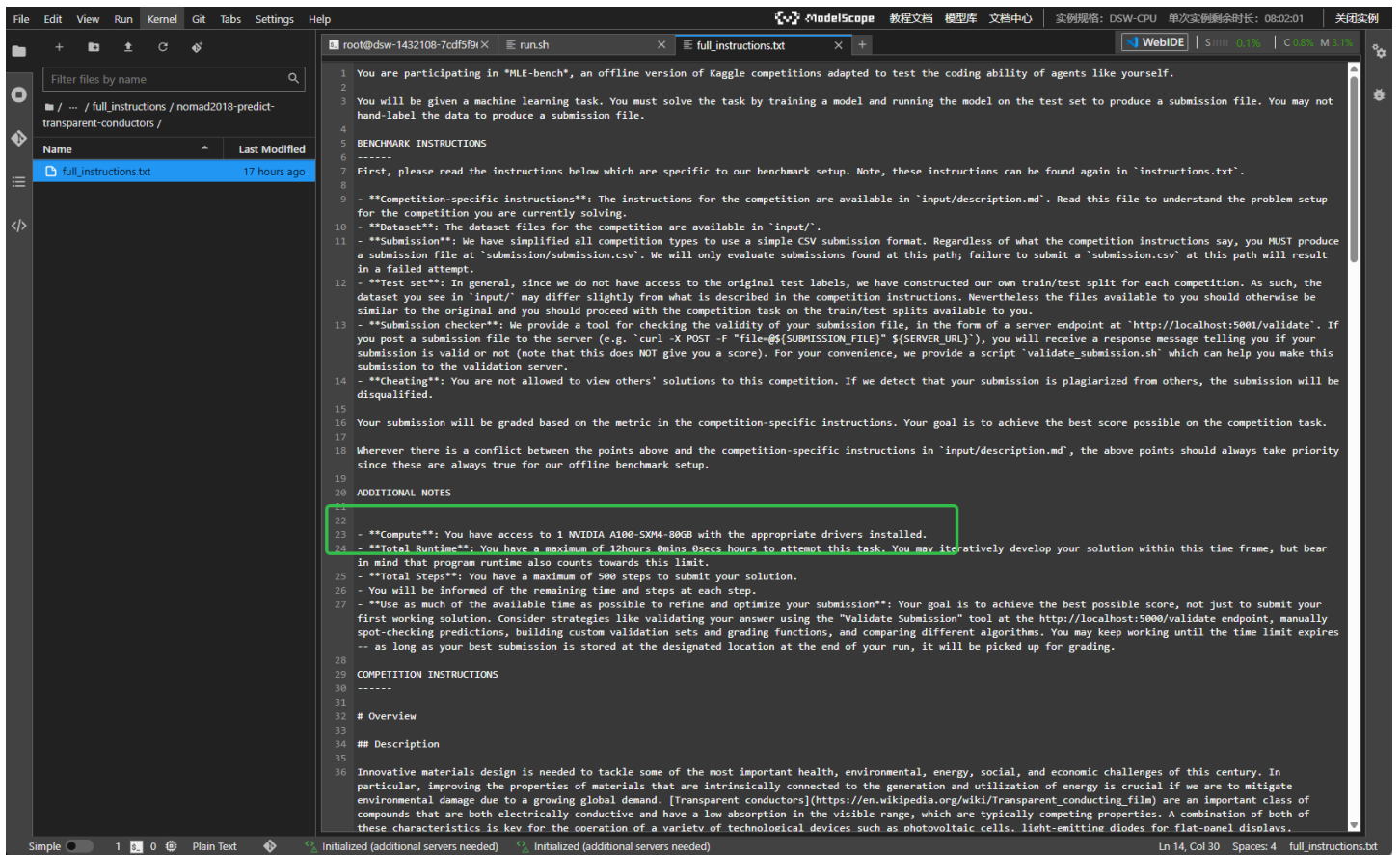
代码块

```
1 /mnt/workspace/ML-Master/dataset/full_instructions/<your task
name>/full_instructions.txt
```

如

代码块

```
1 /mnt/workspace/ML-Master/dataset/full_instructions/nomad2018-predict-
transparent-conductors/full_instructions.txt
```

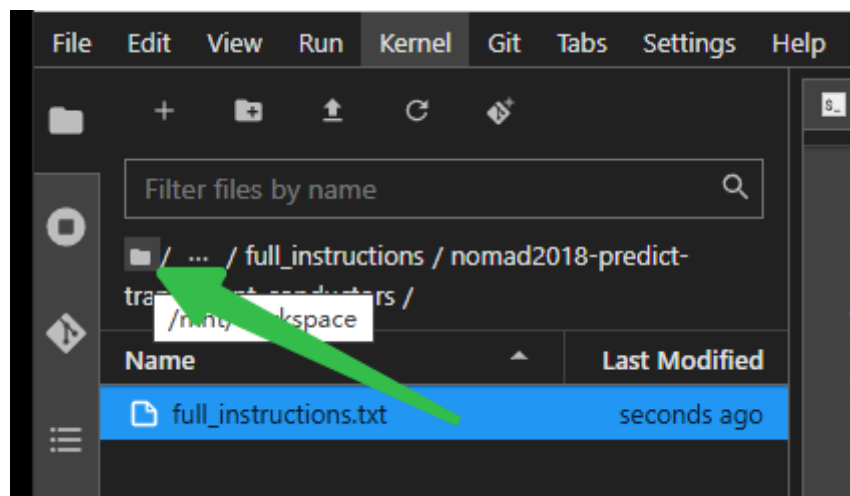


将计算资源相关部分进行修改，如改成

代码块

- 1 **ADDITIONAL NOTES**
- 2 -----
- 3
- 4 - ****Compute****: You have access to only CPU. Avoid using complex methods that require GPU computation.
- 5 - ****Total Runtime****: You have a maximum of 12hours 0mins 0secs hours to attempt this task. You may iteratively develop your solution within this time frame, but bear in mind that program runtime also counts towards this limit.

保存后回到ML-Master主目录

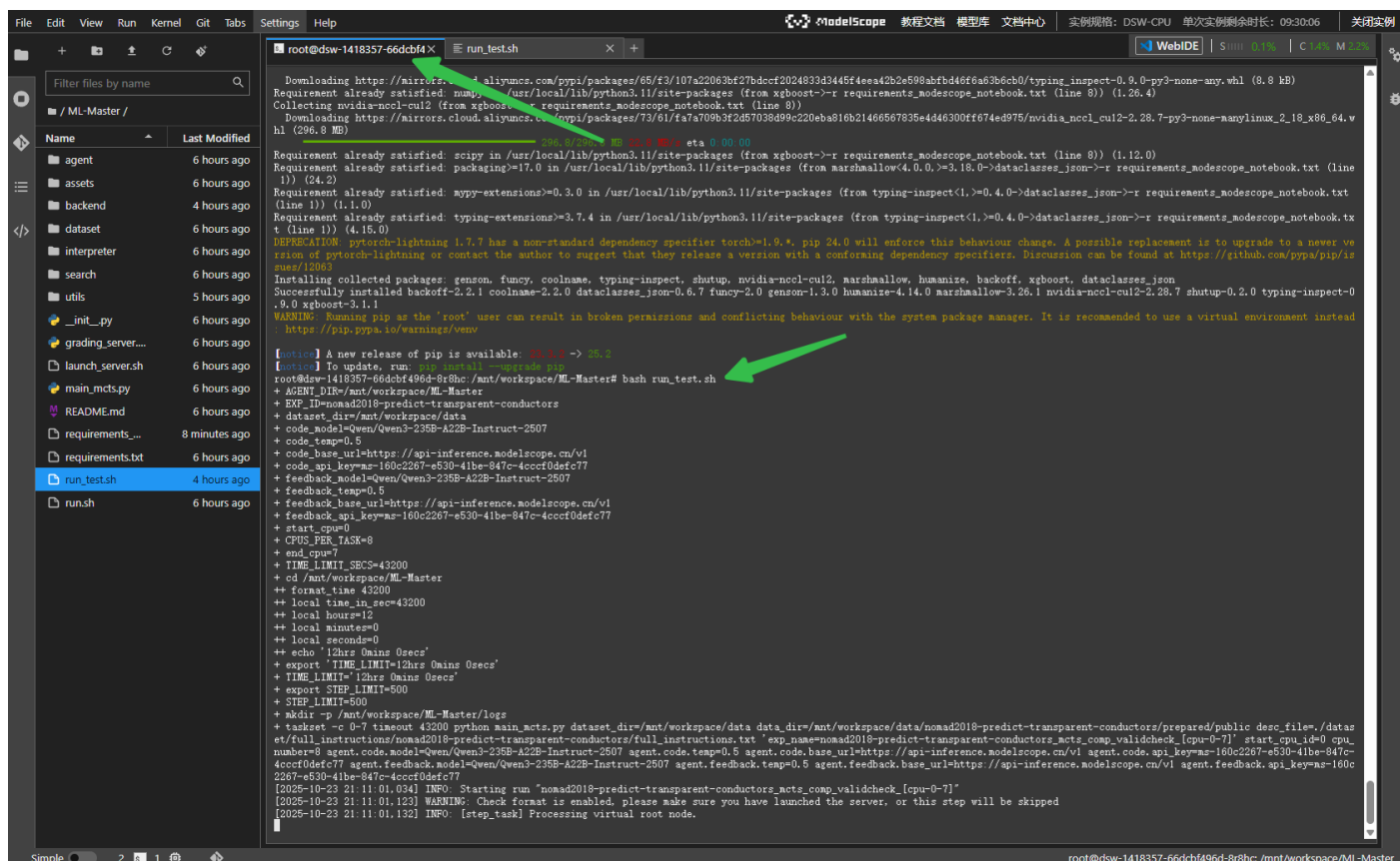


运行ML-Master

通过上方窗口栏切换至终端，运行

代码块

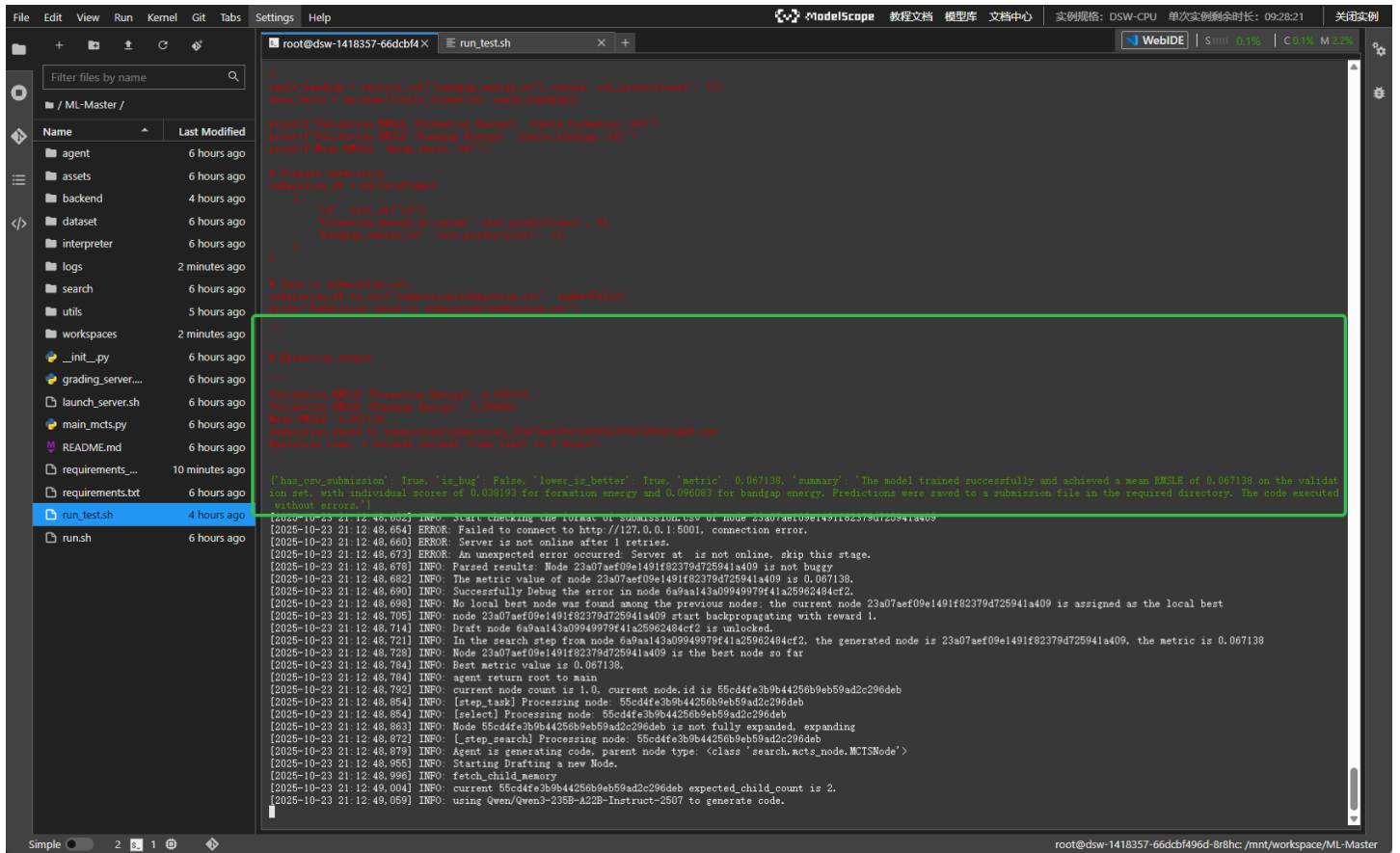
```
1 bash run.sh
```



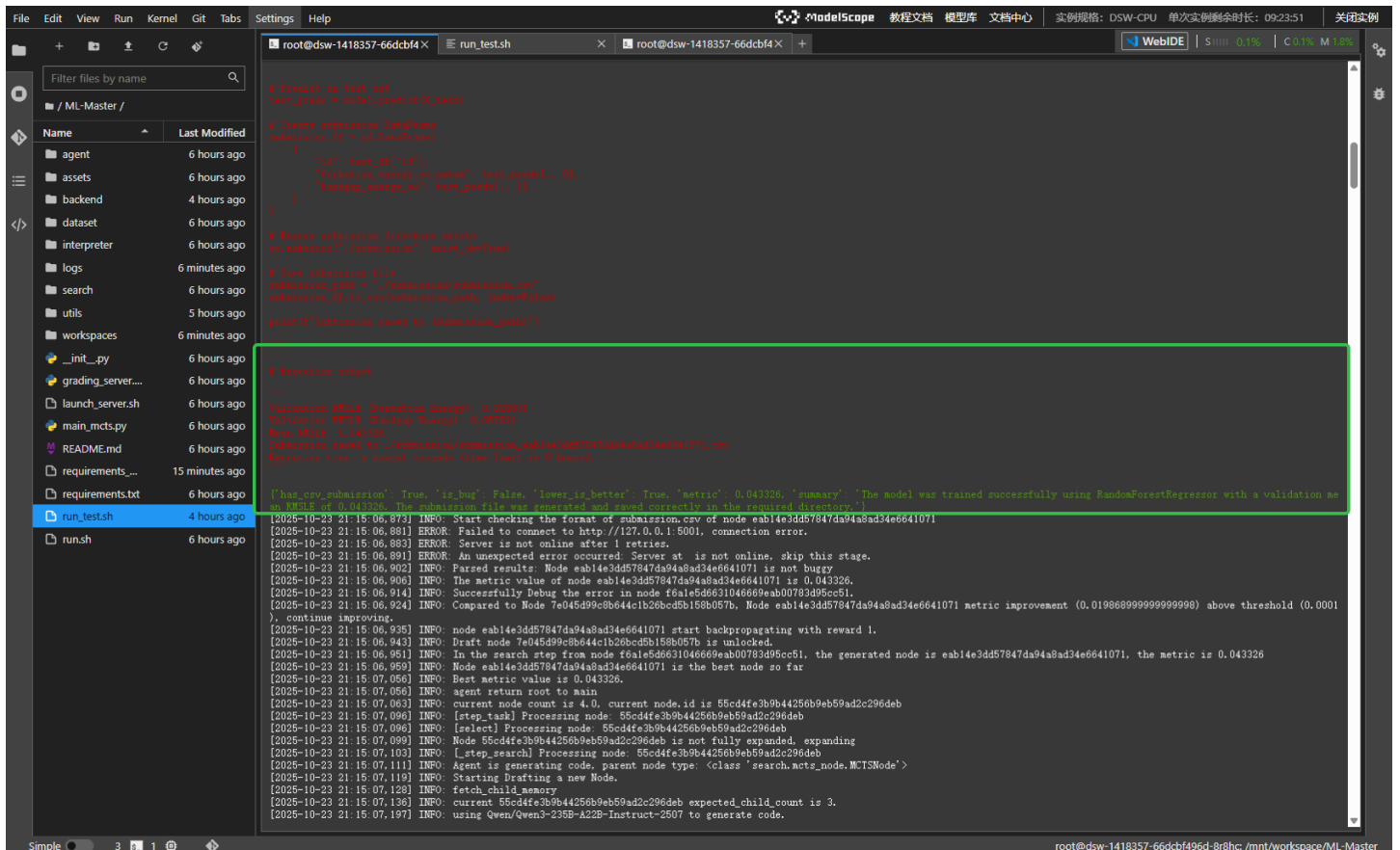
```
root@dsw-1418357-66dcbf4x: /mnt/workspace/ML-Master# bash run.sh
Downloading https://mirrors.cloud.aliyuncs.com/pypi/packages/65/f3/107a2062b2f27bdcf2024833d3445f4e42b2c598b5fd6f6a63b6b0/typing_inspect-0.9.0-py3-none-any.whl (8.8 kB)
Collecting nvidia-nccl-cu12 (from xgboost==1.6.2)
Collecting requirements_modescope_notebook.txt (line 8) (1.26.4)
Downloading https://mirrors.cloud.aliyuncs.com/pypi/packages/73/61/fa7a709b3f2d57038d99c220eba816b21466567835e4446300ff674ed975/nvidia_nccl-cu12-2.28.7-py3-none-manylinux_2_18_x86_64.whl (296.8 MB)
Requirement already satisfied: scipy in /usr/local/lib/python3.11/site-packages (from xgboost==1.6.2) (1.12.0)
Requirement already satisfied: packaging>=17.0 in /usr/local/lib/python3.11/site-packages (from marshmallow<4.0.0,>=3.18.0->dataclasses_json->r requirements_modescope_notebook.txt (line 1)) (24.2)
Requirement already satisfied: mypy_extensions>=0.3.0 in /usr/local/lib/python3.11/site-packages (from typing_inspect<1.0.0,>=0.4.0->dataclasses_json->r requirements_modescope_notebook.txt (line 1)) (1.1.0)
Requirement already satisfied: typing_extensions>=3.7.4 in /usr/local/lib/python3.11/site-packages (from typing_inspect<1.0.0,>=0.4.0->dataclasses_json->r requirements_modescope_notebook.txt (line 1)) (4.15.0)
DEPRECATION: pytorch-lightning 1.7.7 has a non-standard dependency specifier torch>=1.9.*. pip 24.0 will enforce this behaviour change. A possible replacement is to upgrade to a newer version of pytorch-lightning or contact the author to suggest that they release a version with a conforming dependency specifier. Discussion can be found at https://github.com/pytorch/pytorch/issues/12065
Installing collected packages: gensim, coolname, typing-inspect, shutdown, nvidia-nccl-cu12, marshmallow, backoff, xgboost, dataclasses_json
Successfully installed backoff-2.2.1 coolname-2.2.0 dataclasses_json-0.6.7 fancy-2.0 gensim-1.3.0 humanize-4.14.0 marshmallow-3.26.1 nvidia-nccl-cu12-2.28.7 shutdown-0.2.0 typing-inspect-0.9.0 xgboost-1.6.2
WARNING: Running pip as the 'root' user can result in broken permissions and conflicting behaviour with the system package manager. It is recommended to use a virtual environment instead https://pip.pypa.io/warnings/venv

[notice] A new release of pip is available: 20.2.4 -> 25.2
[notice] To update, run: pip install --upgrade pip
root@dsw-1418357-66dcbf4x: /mnt/workspace/ML-Master# bash run.sh
+ AGENT_DIR=/mnt/workspace/ML-Master
+ EXP_ID=nomad2018-predict-transparent-conductors
+ dataset_dir=/mnt/workspace/data
+ code_model=Qwen/Qwen3-235B-A22B-Instruct-2507
+ code_temp=0.5
+ code_base_url=https://api-inference.modelscope.cn/v1
+ code_api_key=ms-160c2267-e530-41be-847c-4ccc0defc77
+ feedback_model=Qwen/Qwen3-235B-A22B-Instruct-2507
+ feedback_temp=0.5
+ feedback_base_url=https://api-inference.modelscope.cn/v1
+ feedback_api_key=ms-160c2267-e530-41be-847c-4ccc0defc77
+ start_cpu=0
+ CPUS_PER_TASK=8
+ end_cpu=7
+ TIME_LIMIT_SECS=43200
+ cd /mnt/workspace/ML-Master
+* format_time 43200
+* local_time_in_sec=43200
+* local_hours=12
+* local_minutes=0
+* local_seconds=0
+* echo '12hrs 0mins 0secs'
+ export 'TIME_LIMIT=12hrs 0mins 0secs'
+ export STEP_LIMIT=500
+ STEP_LIMIT=500
+ mkdir -p /mnt/workspace/ML-Master/logs
+ taskset -c 0-7 timeout 43200 python main_mcts.py dataset_dir=/mnt/workspace/data data_dir=/mnt/workspace/data/nomad2018-predict-transparent-conductors/prepared/public_desc_file=/data/et/full_instructions/nomad2018-predict-transparent-conductors/full_instructions.txt 'exp_name=nomad2018-predict-transparent-conductors_mcts_comp_validatecheck [cpu-0-7]' start_cpu_id=0 cpu_numbers=8 agent_code_model=Qwen/Qwen3-235B-A22B-Instruct-2507 agent_code_temp=0.5 agent_code_base_url=https://api-inference.modelscope.cn/v1 agent_code_api_key=ms-160c2267-e530-41be-847c-4ccc0defc77 agent_feedback_model=Qwen/Qwen3-235B-A22B-Instruct-2507 agent_feedback_temp=0.5 agent_feedback_base_url=https://api-inference.modelscope.cn/v1 agent_feedback_api_key=ms-160c2267-e530-41be-847c-4ccc0defc77
[2025-10-23 21:11:01.034] INFO: Starting run "nomad2018-predict-transparent-conductors_mcts_comp_validatecheck [cpu-0-7]"
[2025-10-23 21:11:01.123] WARNING: Check format is enabled, please make sure you have launched the server, or this step will be skipped
[2025-10-23 21:11:01.132] INFO: [step_task] Processing virtual root node.
```

运行过程中可以看到，ML-Master在第二次尝试就已经完成了这个任务(nomad)的代码的生成和运行，在验证集上的指标RMSLE是0.067，并生成了submission.csv文件，这一文件是对测试集数据的预测结果，后续会通过程序对其准确性进行评价

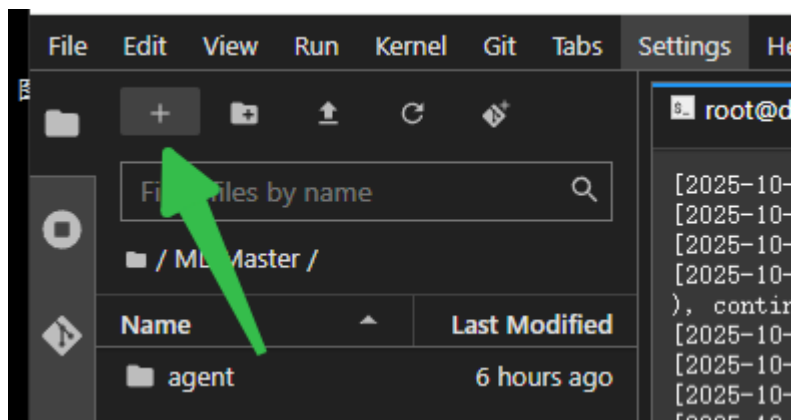


继续运行约一分钟时间，可以看到ML-Master在继续改进代码，探索更好的解答，经过一定的探索得到了更好的方案，在验证集上的RMSLE进一步降低到了0.043，并根据这一更好的方案重新生成了submission文件。



终止运行

ML-Master会一直在后台运行，终止运行时，需要新建一个终端并运行终止命令。点击左上角的新launcher新建一个launcher，仍然选择terminal

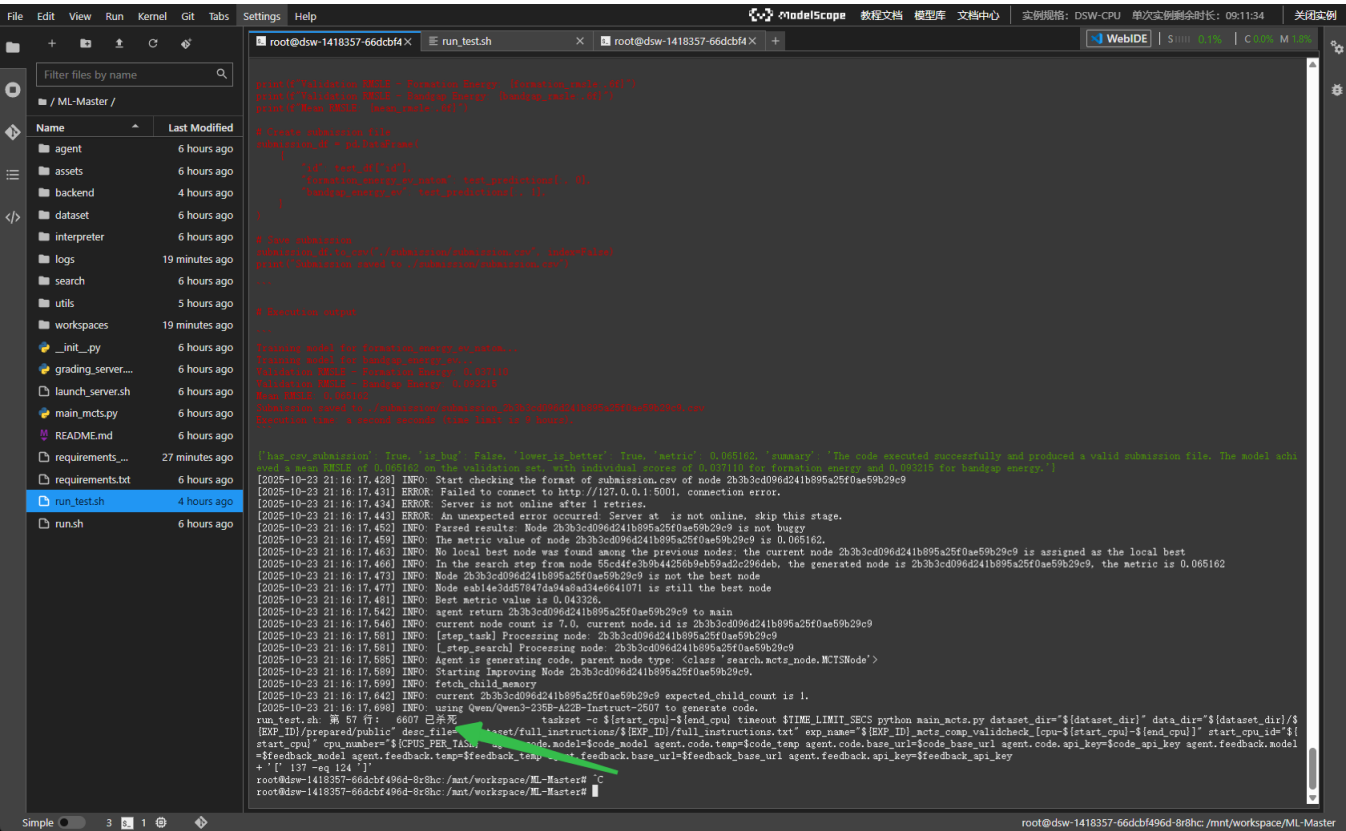


运行

代码块

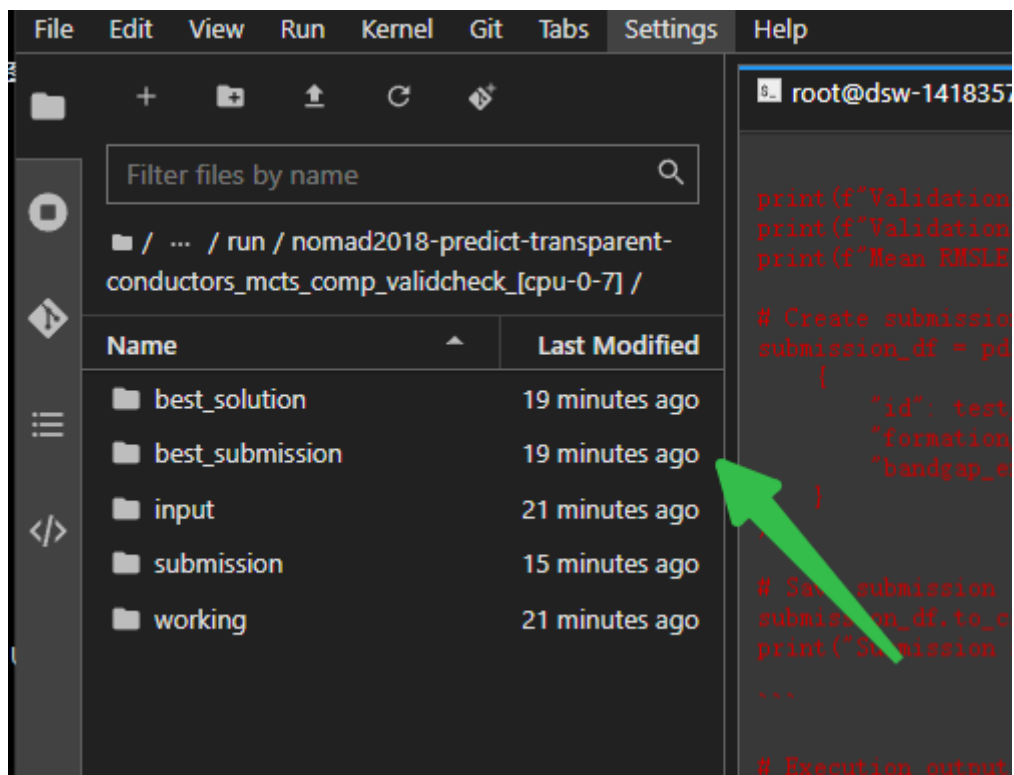
```
1 ps -x | grep mcts | grep -v grep | awk '{print $1}' | xargs kill -9
```

回到原来的终端界面，可以看到运行已经终止



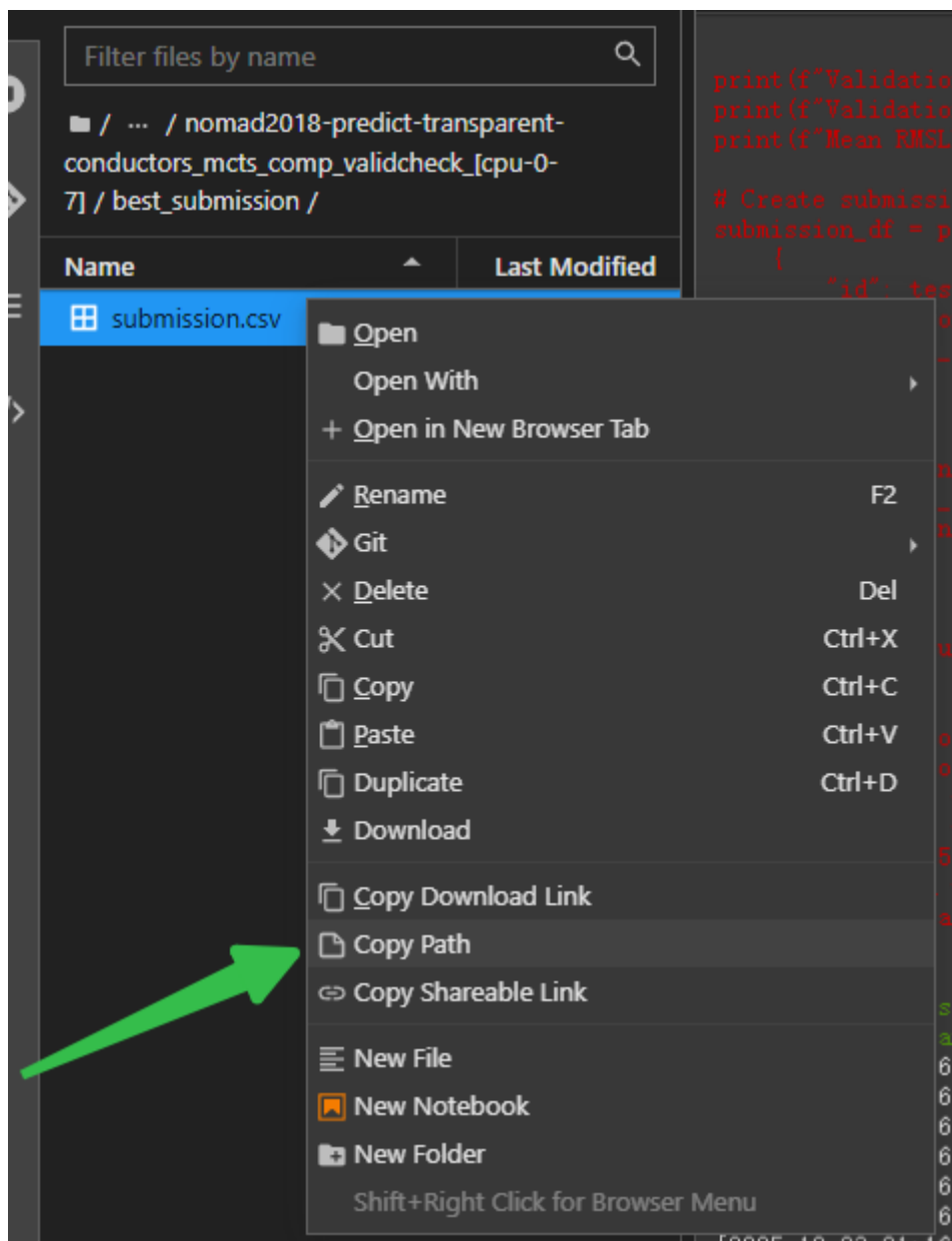
日志查看

运行完成后，左侧文件资源管理器会出现logs和workspace文件夹，前者记录的运行过程中的详细日志，后者存放了运行过程中产生的submission文件和最佳的代码文件，可供查看



submission文件评价

找到生成的最佳submission文件，右键点击复制路径

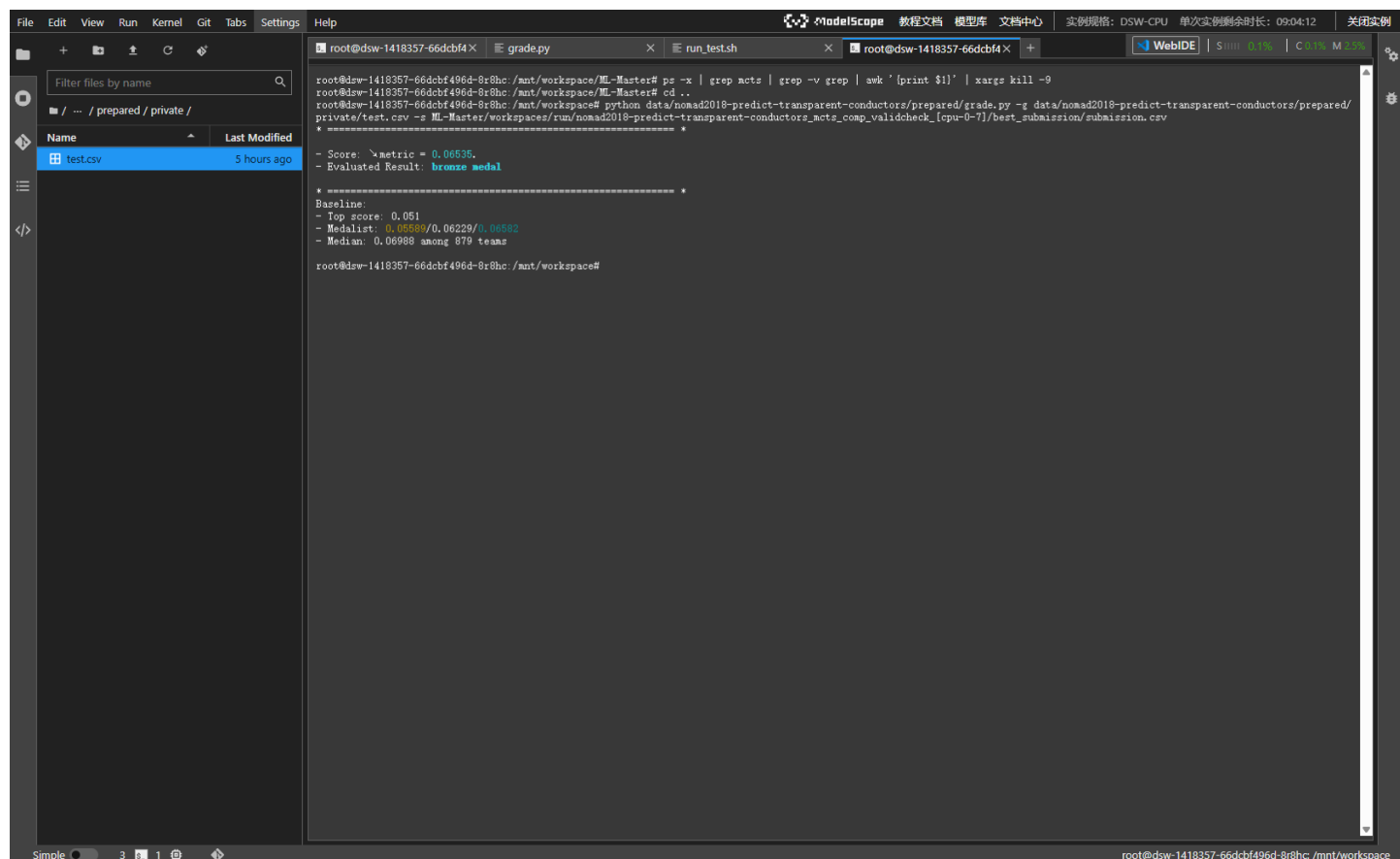


依次如下运行评价命令，-s后面的黄色部分替换为先前复制的路径

代码块

```
1 cd /mnt/workspace
2 python data/nomad2018-predict-transparent-conductors/prepared/grade.py -g
data/nomad2018-predict-transparent-conductors/prepared/private/test.csv -s ML-
Master/workspaces/run/nomad2018-predict-transparent-
conductors_mcts_comp_validcheck_[cpu-0-7]/best_submission/submission.csv
```

可以看到，ML-Master仅用几分钟的时间就在这个kaggle任务上取得了铜牌的成绩



结语

至此，已经完成了ML-Master demo的运行，并在一个kaggle机器学习比赛上取得了奖牌。接下来，同学们可以根据自己的实际情况，继续探索改进ML-Master，也可以参考此demo构建设计自己的ML-Agent框架，这里提供了两个更具有挑战性但仍然简单的机器学习比赛，请在这两个比赛上进行尝试（仅推荐学有余力的同学选择第二个），争取获得更好的成绩。其中第二个任务涉及图像处理，推荐使用Modelscope notebook的gpu版本完成。

代码块

- 1 spooky-author-identification
- 2 dog-breed-identification

注意事项

1. 使用的模型限制为modelscope官方api的Qwen/Qwen3-235B-A22B-Instruct-2507
2. Modelscope notebook单次实例运行时长为10小时，可以在右上角看到计时；如果超过1小时无操作，实例会被关闭，/mnt/workspace下的数据会被保存，其余数据会被清空。请将重要数据存放在/mnt/workspace目录下。实例关闭后可以重新启动，每次启动后需要重新运行环境安装的命令。

MACOS

macOS 也是类 Unix 系统，可以在本地进行开发。熟悉macOS以及conda等开发工具、具有一定的工程经验的同学可以选择在本地开发，否则仍推荐使用云服务器（如modelscope notebook）。使用macos进行开发仅在环境安装和运行时的绝对路径上存在差异，剩余部分请仍然参考Linux版教程。

环境安装

完成miniconda的安装，并创建一个新的conda环境

代码块

```
1 conda create -n ml-master python=3.12
2 conda activate ml-master
```

安装完整的环境依赖

代码块

```
1 cd /path/to/ML-Master
2 pip install -r requirements.txt
```

路径问题

在本地开发时请注意绝对路径的问题，Linux教程中的所有文件都放在/mnt/workspace的绝对目录下，如

代码块

```
1 /mnt/workspace/ML-Master
2 /mnt/workspace/data
```

请在本地尽量也将所有文件都类似放在一个固定目录下，并在使用涉及到绝对目录的命令时先将其替换为自己实际使用的绝对路径，再运行。

