

AutoDL_tutorial

本来是准备空闲的时候写的，但是我发现模型训练还要很久，而且我得等它结束了，然后把租的服务器关机啦，才能回去，不然就是纯烧钱！

首先附上租服务器的网站：<https://www.autodl.com/home>

其次我们正式开始教程。

一、注册账号

这个没有什么难度，然后记得申请学生认证！会有那么一丢丢的优惠！

二、官方教程



进入官网之后，有一个新手入门！点进去，是官方给的教程。其实官方给的教程是十分的详细的！还是很建议读一下的。当然之所以要了解这个服务器，是因为在工作实习中，基本上都是通过远程连接的方法去做的，很少会有一台服务器在你面前的情况。其次呢，服务器的配置主要就是两方面，一方面是环境配置，一方面是数据配置。其他的就相对简单些。当然这里还是推荐去阅读官方的教程，里面提到的一些应用工具，也是之前实习一直用到的，比如filezilla、scp、ssh等等很多。

三、选择服务器并租赁

在官网中，我们登录之后，右上角会有一个控制台，点击控制台，你会看到如下界面

AutoDL

算力市场 AI服务器 算法社区 私有云 帮助文档 更多

控制台 当前余额8795

容器实例

租用新实例

容器实例列表

文件存储 镜像 公开数据 费用 账号

容器实例

实例连续关机15天会释放实例，实例释放会导致数据清空且不可恢复，释放前实例在数据区。

实例ID / 名称

状态

规格详情

本地磁盘

健康状态

付费方式

释放时间/停机时间

SSH登录

快捷工具

操作

北京A区 / 0558机

47f511bde8-c50c49ab

设置名称

已关机

GPU充足

查看详情

系统盘 13.09%
数据盘 55.20%

正常

按量计费

14天23小时58分后释放
设置定时关机

SSH:
密码:

快速工具

开机 更多

北京B区 / 0088机

3cd440bf97-821f555c

设置名称

运行中

查看详情

系统盘 16.34%
数据盘 79.75%

正常

按量计费
余额不足24小时

关机15天后释放
设置定时关机

登录指令
ssh
密码:

快速工具

关机 更多

北京B区 / 0788机

215c4ea7ae-f7d6287a

设置名称

已关机

GPU充足

查看详情

系统盘 12.71%
数据盘 55.84%

正常

按量计费

14天18小时49分后释放
设置定时关机

SSH:
密码:

快速工具

开机 更多

左上方蓝色底色的租用新的实例！点击即可进入选择服务器啦。

容器实例 / 创建实例

计费方式: 按量计费 包日 包周 包月

创建完主机后仍然可以转换计费方式。如选择按量计费，价格发生变动以实例开机时的价格为准

选择地区: 西北B区 北京B区 北京A区 芜湖区 西南A区 佛山市 西北A区 内蒙A区 北京C区

GPU型号: 全部 RTX 4090 (572/2681) RTX 3090 (3/432) RTX 3080 Ti (25/392) L40 (1/13) RTX 3080x2 (28/96) RTX A4000 (0/24) RTX 3060 (0/32) GTX 1080 Ti (0/14) CPU (0/12)

GPU数量: 1 2 3 4 5 6 7 8 10 12

选择主机:

主机ID	算力型号/显存	空闲GPU	每GPU分配	CPU型号	硬盘	最高CUDA	价格(单卡)
430机	RTX 3090 24GB	1 / 12	CPU: 18核 内存: 60GB	AMD EPYC 9754	数据盘: 50GB 可扩展: 1820GB	12.2	¥1.66/时 ¥1.66/时
462机	RTX 3090 24GB	2 / 12	CPU: 18核 内存: 60GB	AMD EPYC 9754	数据盘: 50GB 可扩展: 1631GB	12.2	¥1.66/时 ¥1.66/时
439机	RTX 3090 24GB	0 / 12	CPU: 18核 内存: 60GB	AMD EPYC 9754	数据盘: 50GB 可扩展: 0GB	12.2	¥1.66/时 ¥1.66/时
435机	RTX 3090 24GB	0 / 12	CPU: 18核 内存: 60GB	AMD EPYC 9754	数据盘: 50GB 可扩展: 920GB	12.2	¥1.66/时 ¥1.66/时
431机	RTX 3090 24GB	0 / 12	CPU: 18核 内存: 60GB	AMD EPYC 9754	数据盘: 50GB 可扩展: 1400GB	12.2	¥1.66/时 ¥1.66/时

数据盘: 免费50GB ☐ 需要扩容

实例规格:	GPU型号	CPU	内存	系统盘	数据盘
	RTX 3090 * 1卡	18核心	60GB	30GB	免费50GB SSD

镜像: [基础镜像](#) [社区镜像](#) [我的镜像](#) [社区镜像是什么?](#)

试试搜索Github项目名称, 如: YOLOv5、mmDetection。即刻拥有创建即用的算法运行环境

严禁使用WebUI等算法生成违禁图片, 一经发现立即封号!

open-mmlab/OpenPCDet/centerpoint / v1

创建完成后仍然可以更换其他镜像

优惠券: 请选择

日常费用: ¥0.00/日 配置费用: ¥1.66/时 费用明细

账户余额: ¥50.00

取消

立即创建

如上两张图, 首先要选择地区, 也就是官方把服务器按地区分开啦, 这里推荐选择服务器多的地区。因为服务器中的环境和数据仅能在相同地区的服务器之间进行拷贝! 服务器的选型就不用多说啦。注意下后面的数据扩容, 也比较好理解, 免费50G, 多出来的就要付费啦。

最重要的就是镜像, 你可以将这个理解为环境配置, 因为大家都是做深度学习的, 所以您可以通过基础镜像选择pytorch版本, cuda版本, ubuntu版本等。当然也可以选择社区镜像, 比如一些模型框架OpenPCDet, mmDet3D, 这些可能都有镜像, 这样就省去了配置环境的麻烦。

这里如果你不清楚什么是镜像, 简单理解, 就是将配置好的环境, 打包, 上传到了网上, 人们可以拉下来, 将镜像变成容器, 自己使用, 这块是docker的一些东西啦。但是租服务器这块, 很简单, 你只要选好镜像, 环境就直接给你配好啦。

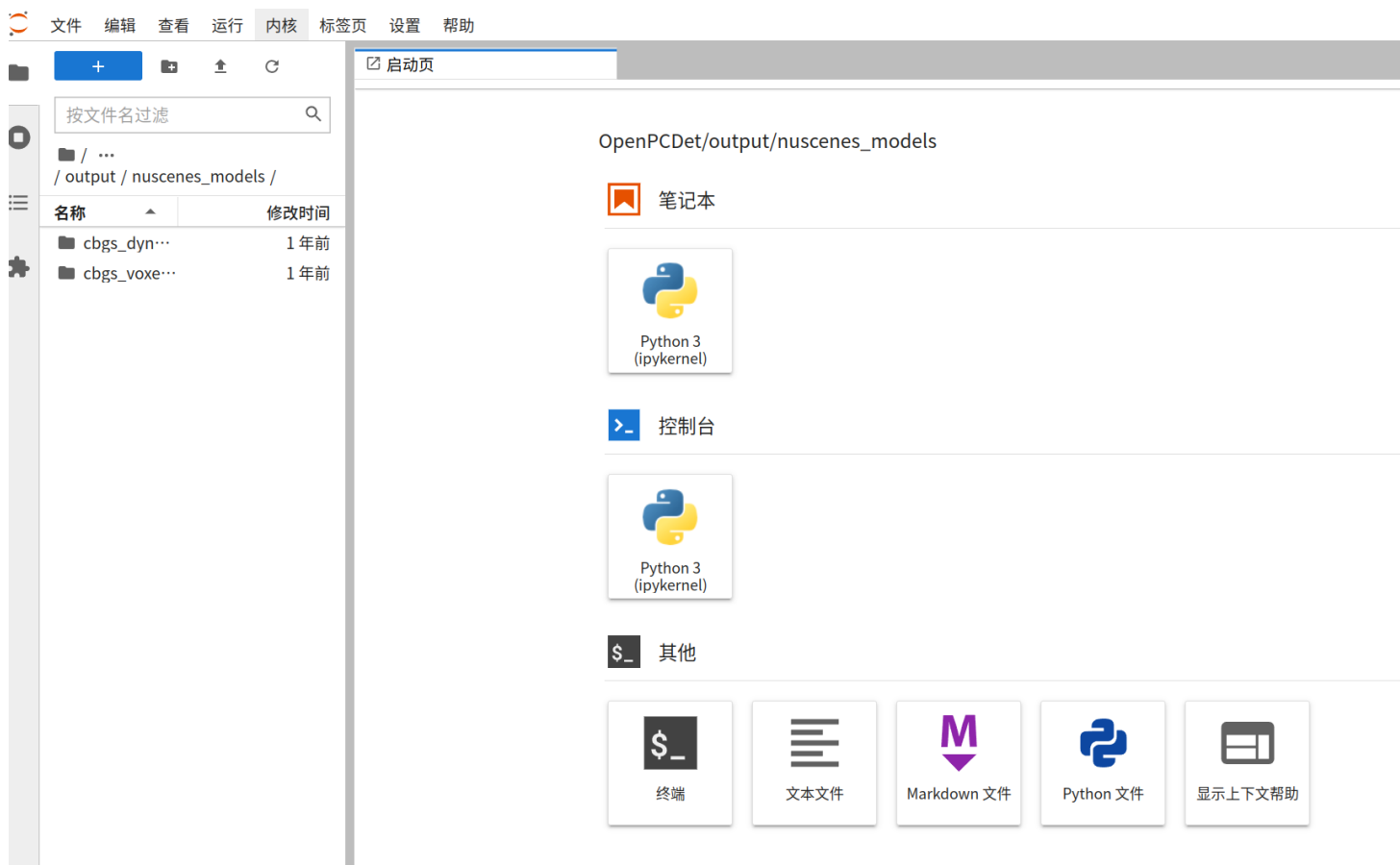
当然也可能你的环境比较特殊, 找不到社区镜像, 那么就只能选择跟你环境匹配的基础镜像来做。这样你拿到服务器之后, 就需要配环境。

四、服务器的基本使用

这里简单叙述下服务器最基础的使用。在有服务器实例的安装图中, 可以看到一个服务器所包含的东西, 最主要的就是SSH登录的指令和快捷工具中的东西。

简单介绍下, SSH就是一个远程链接的东西, 看到上面有一个登录指令和密码, 你可以理解这个是一个账号和密码, 后面我们详细说这个的使用。

然后就是快捷工具中的一些东西, 这里最有用的就是JupyterLab, 点击即可进入电脑。



在这里你可以看到，我们可以打开jupyter notebook和终端之类的东西，也就是说，我们的终端操作均在这里进行即可。当然在左侧，你也可以看到文件的目录。上面的箭头可以用来上传文件，也就是本地文件上传到服务器中，传点小文件还可以，大文件不建议使用这个。其他的一些可以自行探索下。

快捷工具中还有一个AutoPanel，你可以理解这个是对服务器进行实时监控的东西。当然它最主要的功能是数据的上传！这个我们待会说。

最后就剩下，操作中的一些选项，关机开机不用说了吧，很简单，开机开始计费，关机停止计费。但是关机之后，你的显卡就是空闲的啦，别人可能就会占用你的显卡。到时候你就需要将整个实例迁移到别的服务器上去，虽然很方便。

在操作选项中的更多中，有很多选项，这里只说几个重要的：

- 1、无卡模式开机，好处就是一毛钱一个小时，坏处就是卡不再被你占用，可能被别人抢走，这个操作，一般用于数据上传环节。
- 2、更换镜像、保存镜像：选服务器的时候，我们选择的镜像还记的吧，更换就是你觉着一开始选的不好，重新整一个。保存镜像，我理解就是将现有的环境打包保存，后面方便你在其他服务器上使用。
- 3、克隆实例：这个真的好用，我一开始租的4090，因为上传数据，我整了好半天，开的无卡模式，等我开机之后，发现卡被别人占用了，然后我就用的这个，将整个实例克隆到别的服务器上，包括环境啦，数据啦。

五、数据上传

首先先介绍下服务器的文件配置

```
+-----AutoDL-----+
目录说明：


| 目录               | 名称  | 速度 | 说明                                    |
|------------------|-----|----|---------------------------------------|
| /                | 系统盘 | 快  | 实例关机数据不会丢失，可存放代码等。会随保存镜像一起保存。         |
| /root/autodl-tmp | 数据盘 | 快  | 实例关机数据不会丢失，可存放读写IO要求高的数据。但不会随保存镜像一起保存 |
| /root/autodl-nas | 网盘  | 慢  | 可以实现多实例间的文件同步共享，不受实例开关机和保存镜像的影响。      |


cat: /sys/fs/cgroup/cpu/cpu.cfs_quota_us: No such file or directory
cat: /sys/fs/cgroup/cpu/cpu.cfs_period_us: No such file or directory
bash: cfs_quota_us / cfs_period_us: division by 0 (error token is "cfs_period_us")
CPU : 核心
cat: /sys/fs/cgroup/memory/memory.limit_in_bytes: No such file or directory
内存: 0 GB
GPU : NVIDIA GeForce RTX 4090, 1
存储:
  系统盘 / : 17% 5.0G/30G
  数据盘 /root/autodl-tmp: 80% 40G/50G
  网盘 /root/autodl-nas:
+-----+

```

这里说的很清楚，我们主要用到的文件夹就是/的系统盘，这里主要存放一些代码，/root/autodl-tmp/的数据盘，这里存放数据集，如果你有的话。因为服务器提供了很多公开的数据在/root/autodl-pub/中。

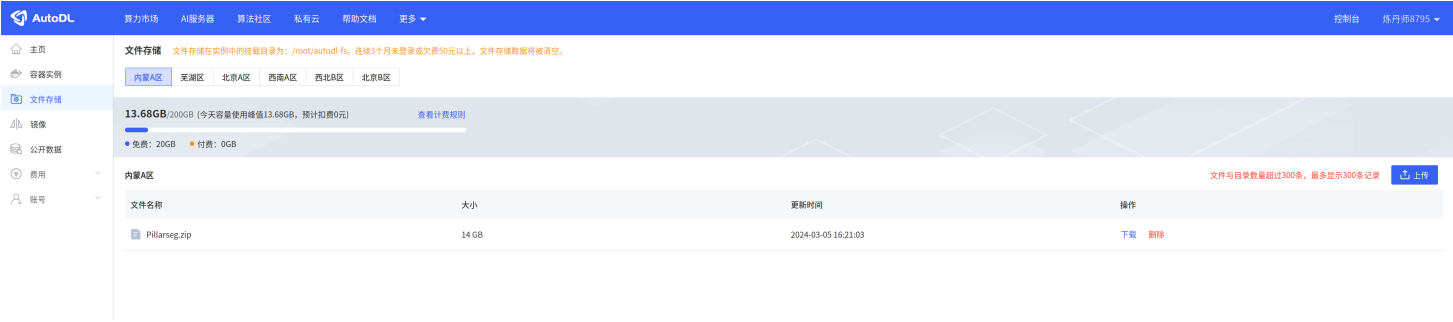
```
root@autodl-container-3cd440bf97-821f555c:~# ls
autodl-fs  autodl-nas  autodl-pub  autodl-tmp  miniconda3  OpenPCDet  Pillarseg  tf-logs
root@autodl-container-3cd440bf97-821f555c:~# cd autodl-pub/
root@autodl-container-3cd440bf97-821f555c:~/autodl-pub# du -sh *
1.1G  ADEChallengeData2016
15G   Aishell
895G  argoverse2.0-sensor
3.9G  BERT-Pretrain-Model
4.2G  CASIAWebFace
22G   CelebA
325M  cifar-10
162M  cifar-100
16K   CMLR
14G   CrowdHuman
1.2G  CUB200-2011
43G   CULane
8.5G  DIV2K
19G   DOTA
28M   DRIVE
11G   Flickr2K
72G   GOT10k
112M  horse2zebra.zip
86G   ILSVRC2015
158G  ImageNet
13G   ImageNet-mini
93G   KITTI_Depth_Completion.tar
287G  LaSOT
2.4G  ModelNet
5.5G  mot17
4.8G  MOT20
12G   mpii_human_pose
5.0G  mvtec_anomaly_detection.tar.xz
549G  nuScenes
1.4G  NUSWIDE
703G  Objects365
1.1G  RoBERTa-Pretrain-Model
15G   S3DIS
83G   SemanticKITTI
1.1T  TrackingNet
36G   TT100K
293M  vangogh2photo.zip
82G   Vimeo-90k
2.7G  VOCdevkit
root@autodl-container-3cd440bf97-821f555c:~/autodl-pub#
```

一开始我还满心欢喜，挺好的这里提供了Nuscenes数据集，但是我发现里面存的都是压缩文件，也就是说，如果你真的想用这个数据集的话，你需要把它拷贝到/root/auto-tmp/中，并解压，因为/root/autodl-pub/是仅可读的不可修改的。

ok，我们说下怎么上传本地数据集，这里需要用到阿里云网盘。这里仅推荐这一种方法，其他方法慢的要死！百度网盘因为授权麻烦，所以这里也不推荐。

首先注册阿里云，然后将本地数据集打包上传。然后在服务器的快捷工具栏中的AutoPanel打开，找到公网网盘，然后完成验证，即可下载数据集。这里数据集下载到/root/autodl-tmp/中。大的数据集仅推荐这一个方法。

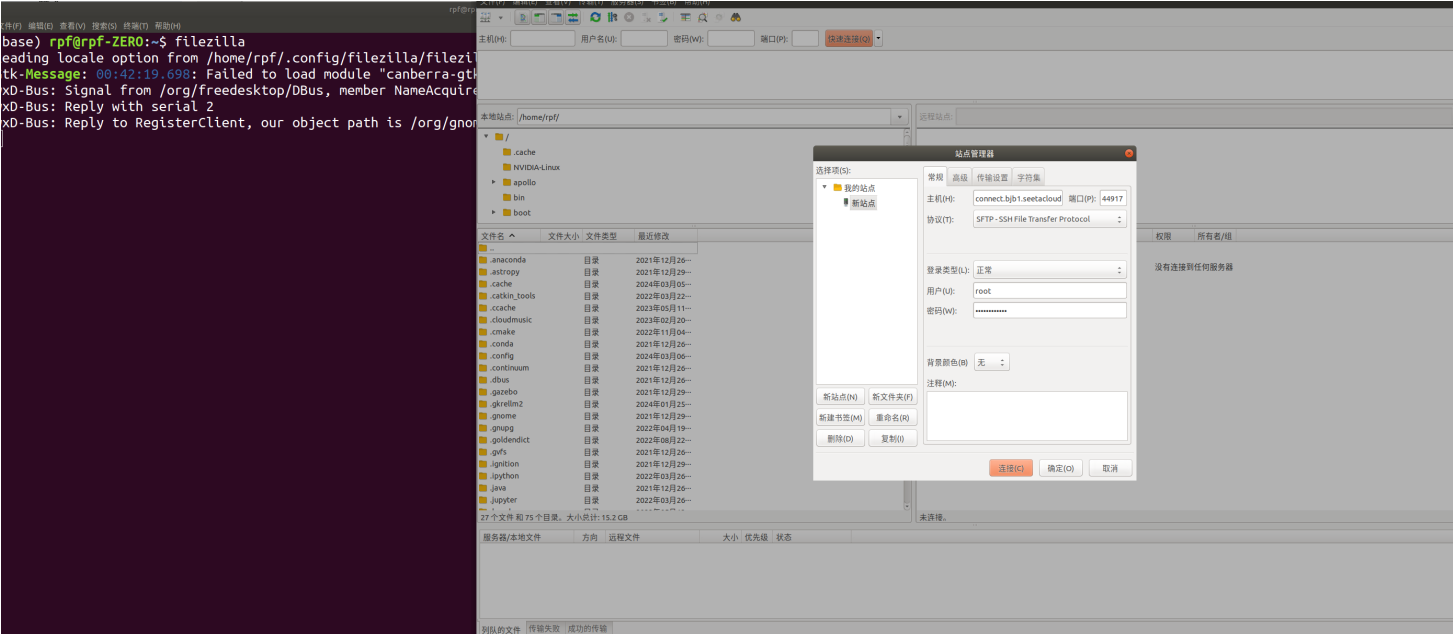
如果是小的文件，仅仅是代码之类的，可以选择的方法就比较多啦。



可以选择这里的文件存储。不可传文件夹。这里需要注意一点：你用的哪个区的服务器，就把数据上传到哪个区！别传错了，本来就传的慢，我等了半天，最后发现区传错了，人在绷不住的时候是真的绷不住。

第二种方法，就是刚才我们提到的在点击JupyterLab后，左上角的箭头可以上传文件。之前试过，这里也很慢！

第三种方法，可以通过filezilla或者XShell的工具进行上传，当然这两个工具，也可以将服务器的文件下载下来。我具体配置其实很简单，可以参考官方文档的教程。其中filezilla是在本地电脑安装的。如下图，需要的主机，端口，用户，密码，这些实际上都是服务器的SSH登录指令和密码。

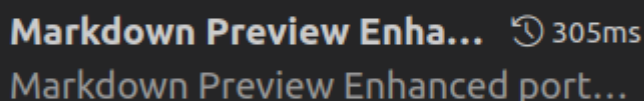


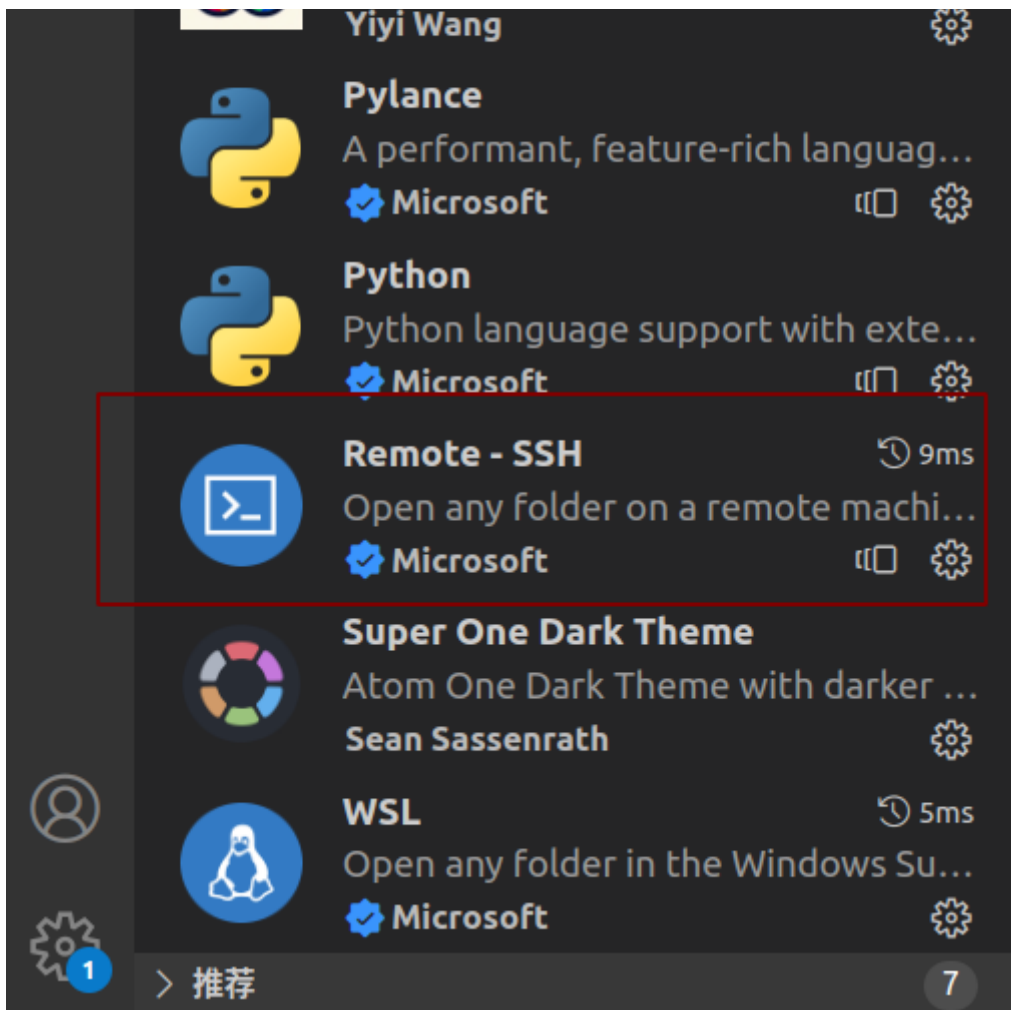
第四种方法，就是scp命令行的方法，这个不了解，可以去搜一下。官方也给了教程到这里，数据上传整体就结束啦！

VScode配置

数据有啦，代码有啦，环境大家自己配置哈，最后就差运行啦，那么选择什么编译器呢，官方教程给了很多，这里推荐的是vscode，因为远程连接很简单。pycharm的远程连接，需要专业版的，社区版不可以。

首先打开一个vscode，然后安装一个用于远程连接的插件Remote-SSH。其他插件你感兴趣也可以安装。在安装了remote-ssh之后，左侧就出现一个显示器的东西。





你可以点击那个显示器，也可以点击vscode窗口左下角的那个标志，大于号顶着小于号的标志，然后选择connet to Host，然后选择+Add New SSH Host，然后你就需要把SSH登录指令复制进去，然后会让你选择配置文件，这里不用选直接回车就成。最后连接输入密码即可。

然后我们就可以在vscode打开远程的代码文件夹啦。这里注意下哈，我们还需要给远程vscode安装本地插件，也就是左侧插件标志，点一下，会发现有一个远程插件！然后安装和本地一样就成，因为要装python插件，才能选择环境！

到这里，整体就结束啦~