本作业分为两个难度，体验作业（低难度）和进阶作业（中等难度），推荐大家先完成体验作业，再尝试进阶作业

**体验作业：完成Evaluation流程**

体验作业提供：finetune完毕的tnews ckpt文件、tnews数据集、源代码

体验作业流程：

1. 申请华为云账号、申请ModelArts华为云昇腾集群服务公测

2. 从链接下载tnews ckpt文件、tnews数据集和源代码

3. 创建obs桶

4. 上传tnews ckpt文件、tnews数据集和源代码到obs桶

5. 提交ModelArts 训练作业

6. 等待、查看结果

**体验作业详细步骤**

1. 申请华为云账号、申请ModelArts华为云昇腾集群服务公测



<https://www.huaweicloud.com/>

注册完毕之后，进入到ModelArts界面，点击昇腾集群公测，完成申请流程



<https://www.huaweicloud.com/product/modelarts.html>

1. 完成申请之后，开始下载所需文件，包括数据集文件、ckpt文件和源代码文件

**tnews数据集：**

https://e-share.obs-website.cn-north-1.myhuaweicloud.com?token=

提取码:

123456

**bert源代码：**

<https://21days-bert.obs.cn-north-4.myhuaweicloud.com:443/bert.zip?AccessKeyId=M7KX8KLMT0ZL1P8QWXZ5&Expires=1634522579&Signature=LY3ZdN9gtoxx8ULDQRkkPuOp1nk%3D>

源代码下载完成后，在本地解压缩

**tnews ckpt:**

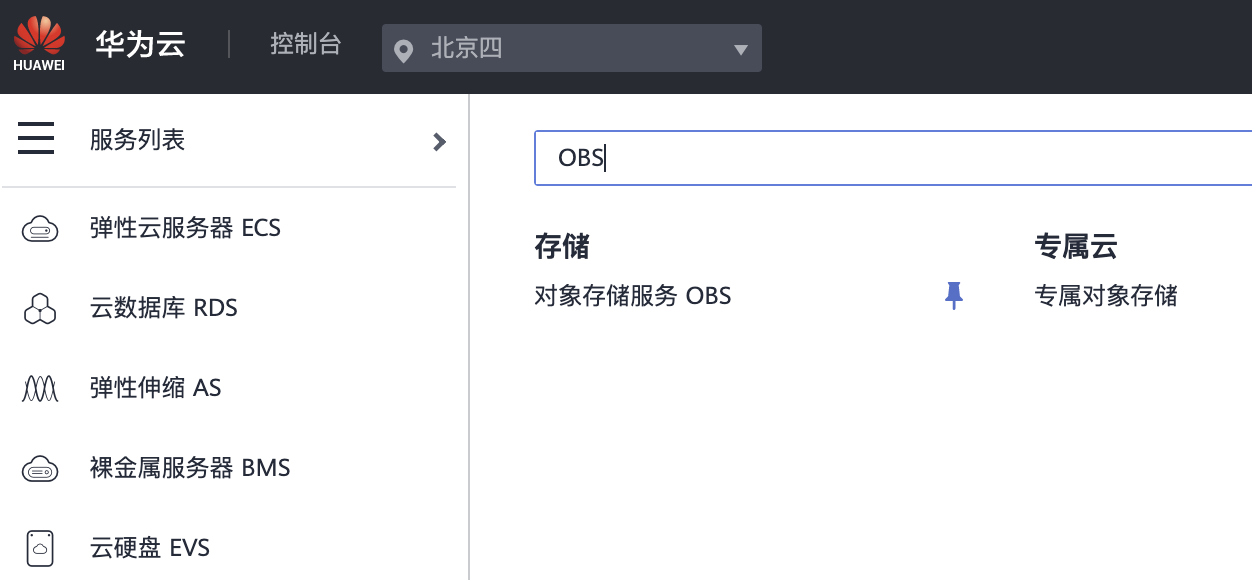
https://21days-bert.obs.myhuaweicloud.com:443/tnews-3\_3335.ckpt?AccessKeyId=M7KX8KLMT0ZL1P8QWXZ5&Expires=1634522122&Signature=d21XsEKkzByK04oyZJ2c9JMGNLk%3D

这个文件比较大，请大家耐心下载

1. 创建OBS桶，进入官网界面https://www.huaweicloud.com/product/obs.html，点击右上角的控制台按钮



在控制台界面中，在左上角选择北京四节点，同时在左边的弹出式菜单的搜索框中，搜索OBS，并选择“对象存储服务OBS”



此时我们已经进入了OBS对象存储服务的界面，点击右上角“创建桶”按钮



按照以下模板创建桶



数据冗余存储策略选择“单AZ存储”

桶名称自己取一个不重复的名称即可

桶策略采用公共读

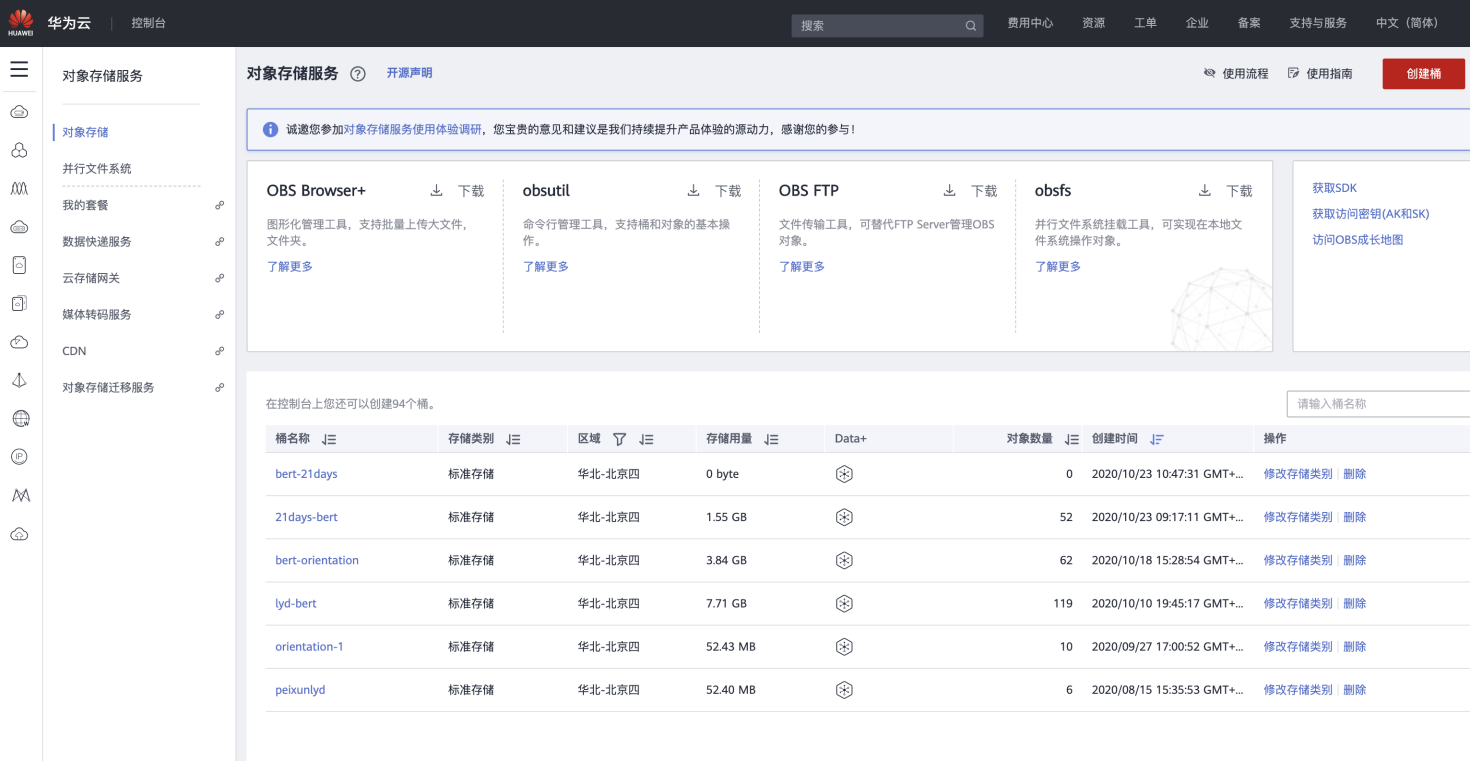
其他单策略使用默认单即可

点击创建按钮之后，OBS桶就创建完毕了

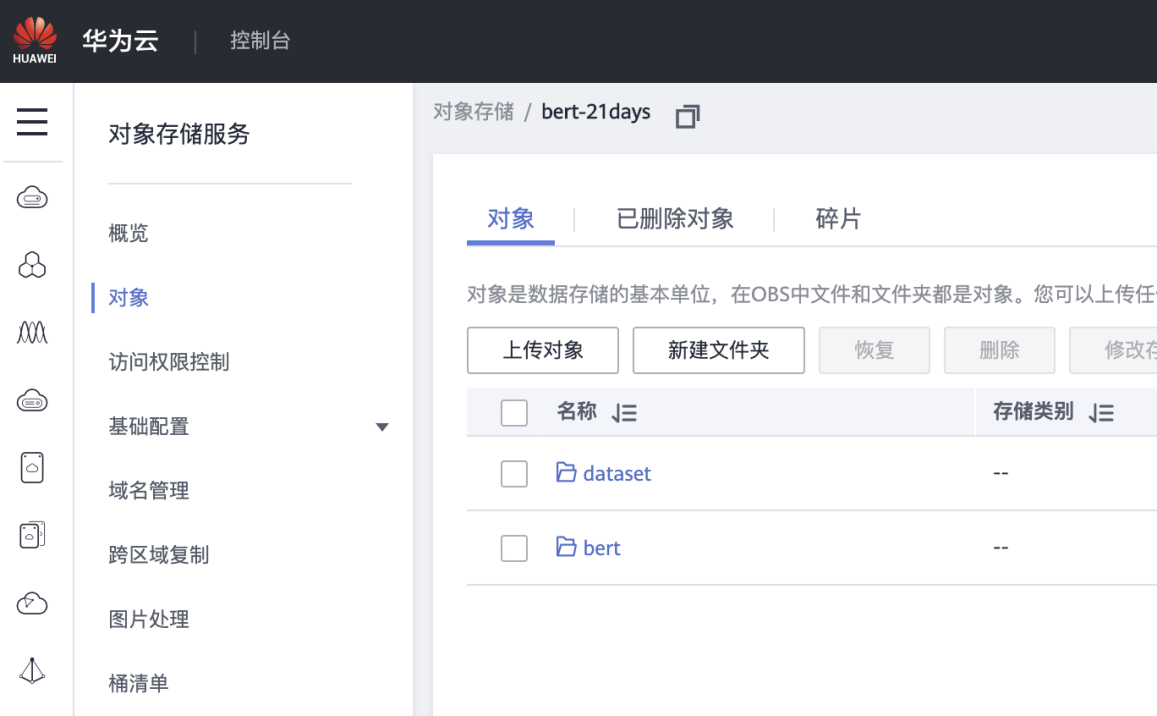
1. 上传文件

这一步我们需要在刚刚创建的OBS桶内新建一些文件夹，并且上传我们刚刚下载好的文件

在OBS主页面点击进入到刚才创建出来的桶



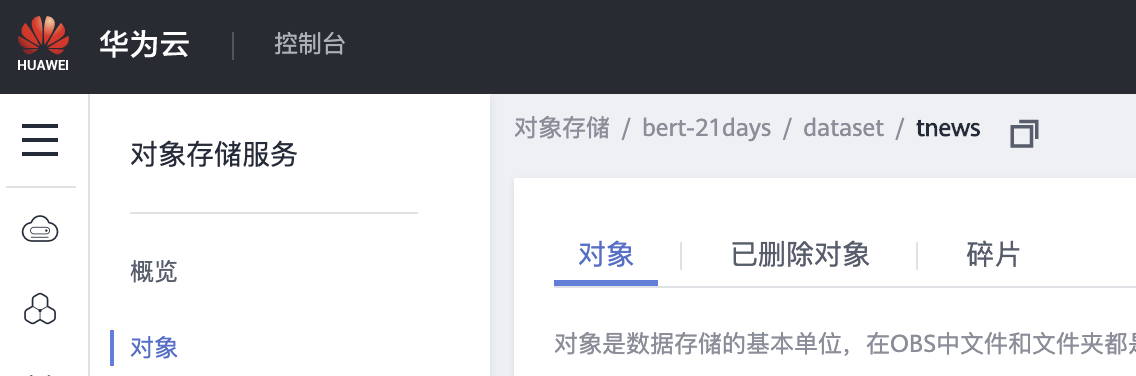
在左侧的弹出菜单中，点击“对象”，然后在右边点击新建文件夹,在根目录下新建两个文件夹：bert和dataset



点击进入dataset文件夹，再创建一个二级文件夹tnews，并且进入tnews文件夹，点击“上传对象”按钮开始数据集的上传，可以采用拖拽的方式，直接将文件拖拽进入下图的红框中

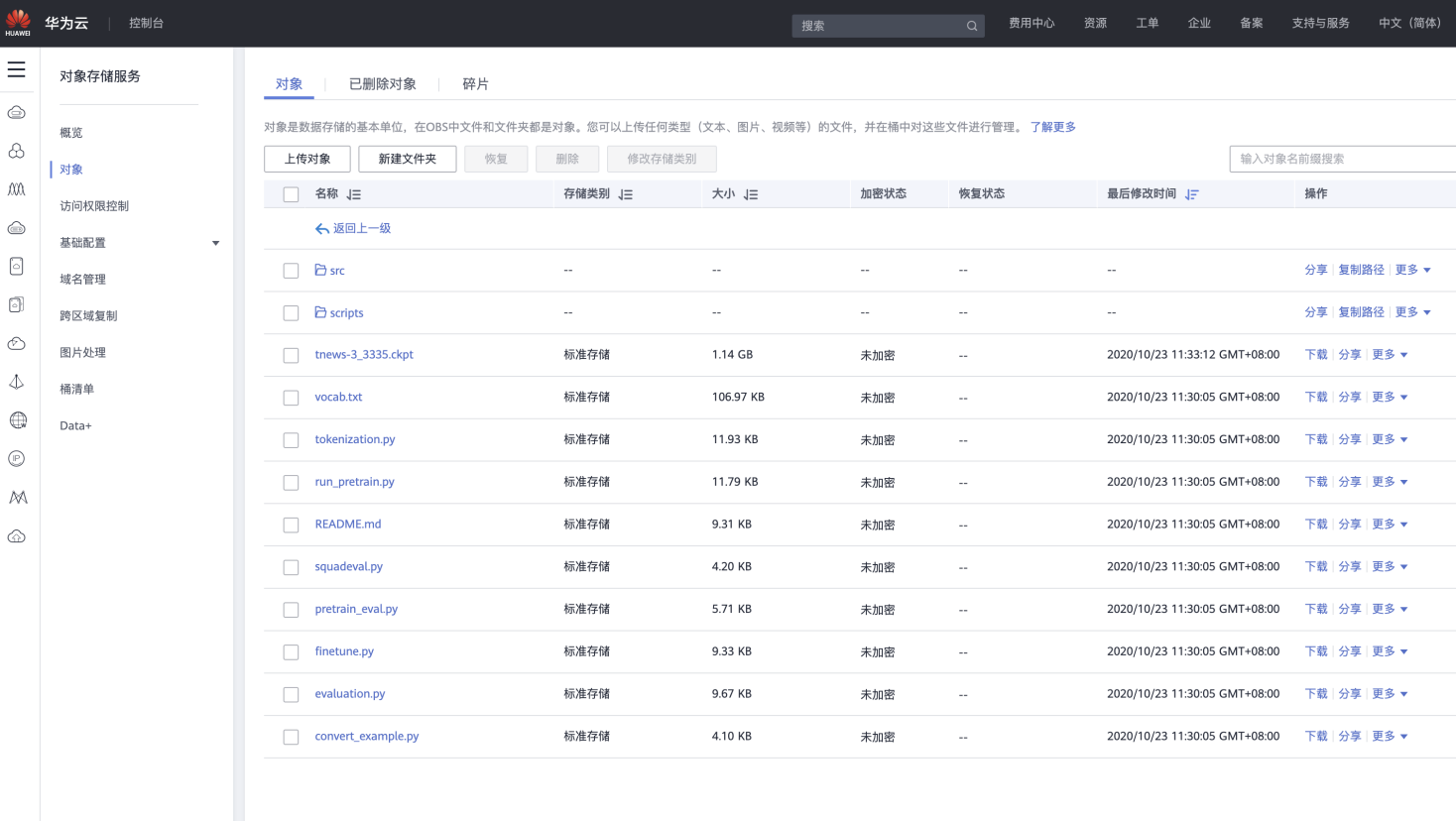


在完成了数据集的上传之后，回到桶的根目录，并且点击进入bert文件夹，同样的方法，开始上传解压缩之后的源代码文件，这里推荐使用拖拽的方式，把解压之后的文件和文件夹全选，然后拖拽进入上图的框中



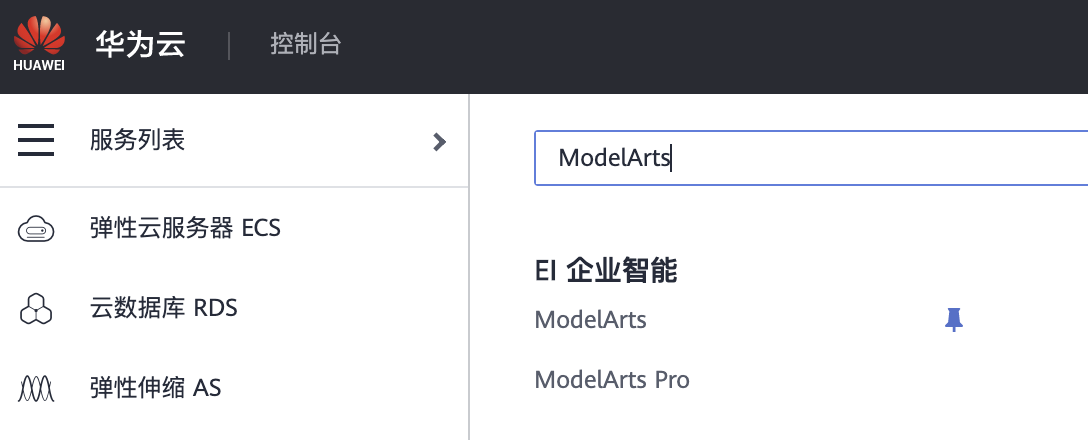
最后上传tnews ckpt文件，这里我们直接把tnews ckpt文件上传到bert目录即代码根目录下，由于tnews ckpt这个文件比较大，超过1GB，所以上传时间会略微长一些

上传全部完成之后，截图如下：



至此，上传部分结束

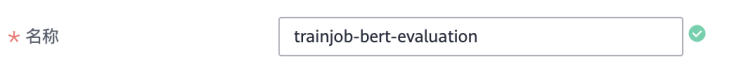
1. 提交ModelArts训练作业，类似我们进入OBS的方式，在左侧服务列表的弹出菜单中，在搜索框中输入modelarts，并且选择下面的“ModelArts”



进入ModelArts界面之后，点击左侧“训练管理”，在下拉菜单中点击“训练作业”，进入训练作业界面后，点击中上方的“创建”按钮



在创建训练任务的界面中，名称同样的，自己取一个不重复的名称即可



下面的参数配置中，以下几个参数需要配置：

“算法来源” 选择“常用框架”，

“AI引擎”选择“Ascend-Powered-Engine”和“MindSpore-0.5-python3.7”

“代码目录”选择之前在obs建立的bert目录即代码根目录

“启动文件”选择代码根目录下的evaluation.py

“数据来源”选择“数据存储位置”

在“数据存储位置”中之前在obs中建立的dataset文件夹下的tnews文件夹

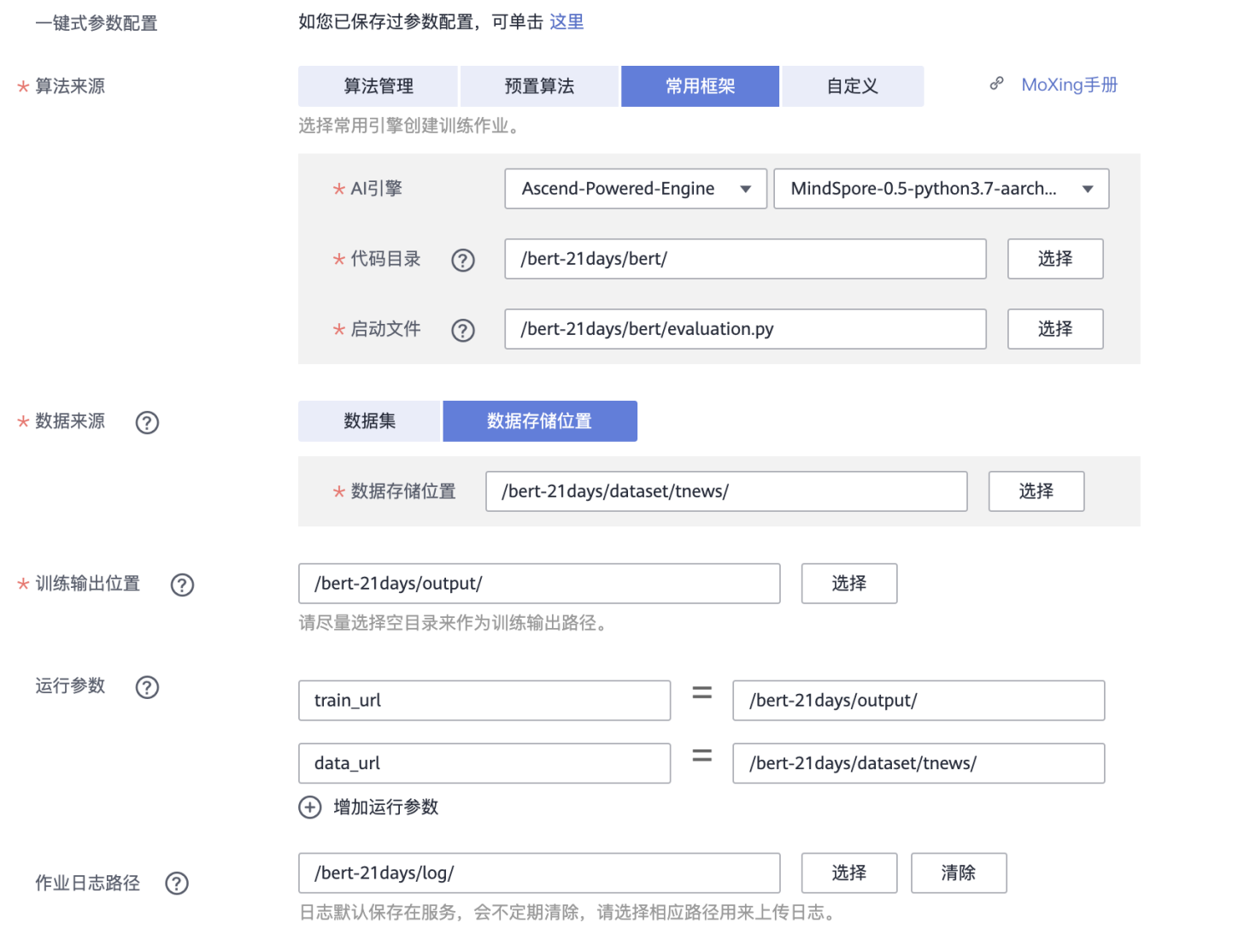
“训练输出位置”可以选择在obs根目录下新建一个文件夹“output”专门用来存储输出

“作业日志路径”跟“训练输出位置”类似，可以在obs根目录下创建一个“log”文件夹用来存放日志输出

“资源池”选择“公共资源池”，“类型”选择“Ascend”，“规格”选择“Ascend: 1 \* Ascend 910 CPU: 24核 96GiB”

其他参数选择默认即可

最终参数设置结果如下图所示：



然后就可以提交任务了

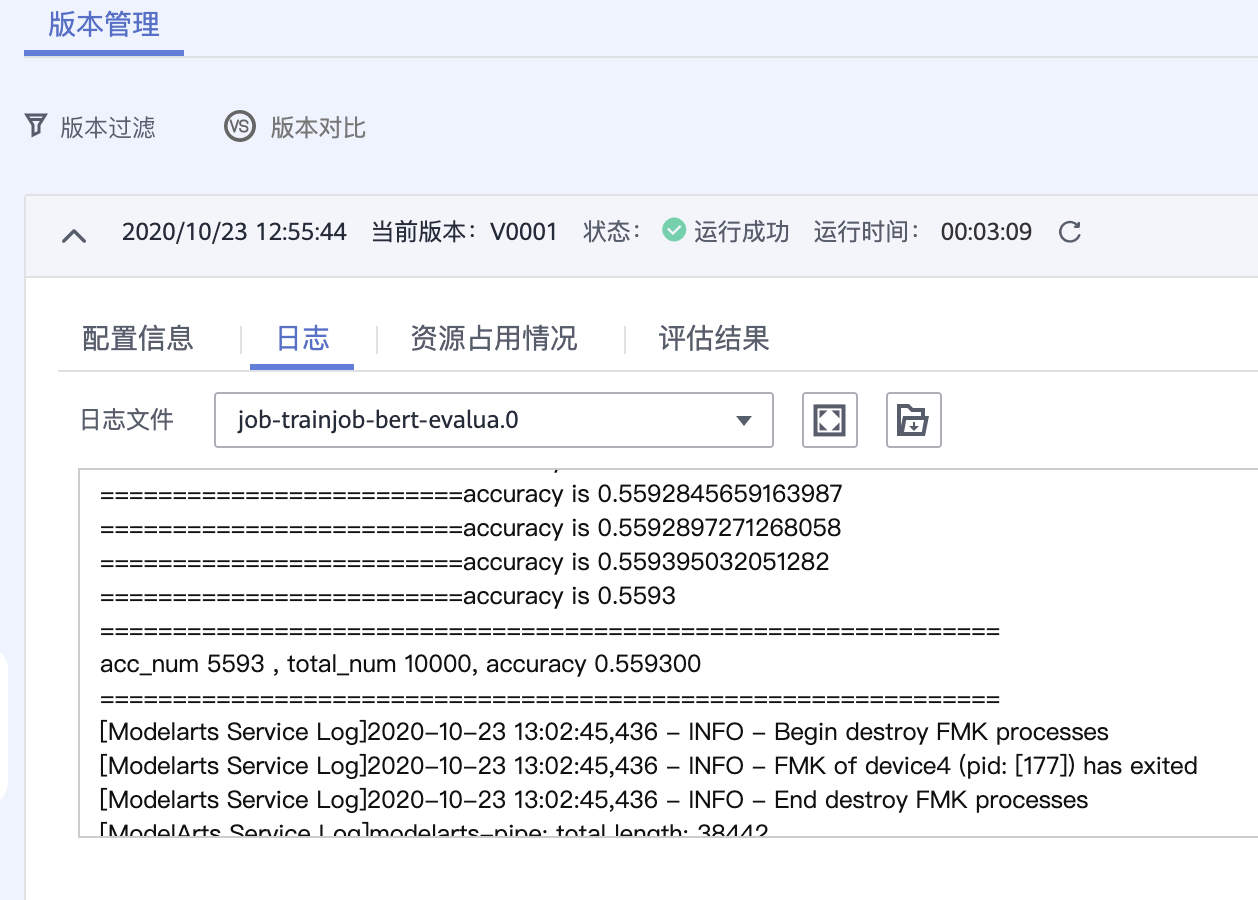
1. 等待、查看结果

本身的evaluation任务时间是很短的，但如果同时提交任务但人数很多，可能会导致大部分人必须排队等待，等待的时间可长可短，所以到了这一步，大家也可以做一些其他但事情，过一段时间回到ModelArts网页上检查一下作业状态即可。

如果显示“运行成功”，即代表Evaluation运行作业已经完成，我们可以，可以点击作业名称，进入到作业内部详情，再点击“日志”，查看输出



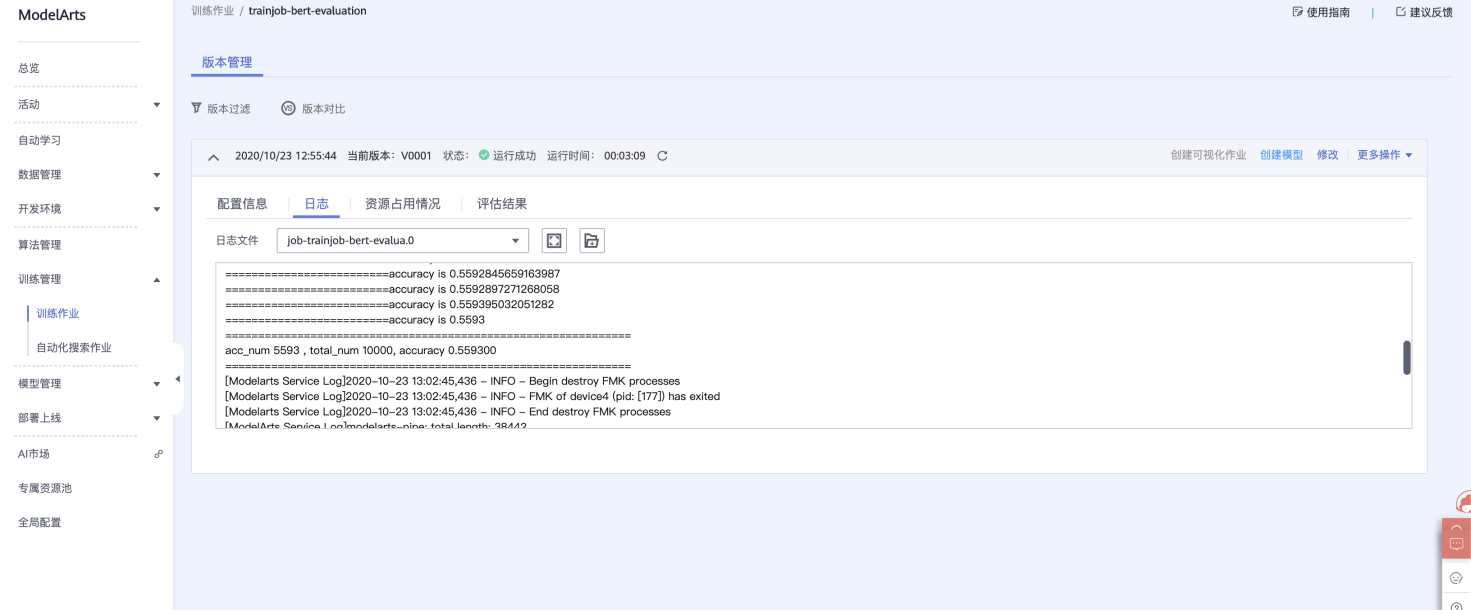
在日志里上下拖动滑动条，找到accuracy这一项，最后但这一条accuracy就是整个验证集但正确率



这样把正确率但截图上传完成了这次的体验作业

附加题：如果我们并不是想对整个验证集做推理，而只是想对一句话（一句中文短新闻）做推理，应该怎么做呢？

其实1～4步都是一样的，只有第5步略微不一样，我们可以回到提交过的作业界面，点击右上方的“修改”按钮，修改我们的参数，重新提交一次：



绝大多数参数都一样，只有一个地方有区别，在运行参数里面，我们点击下面的加号按钮，添加一行参数：

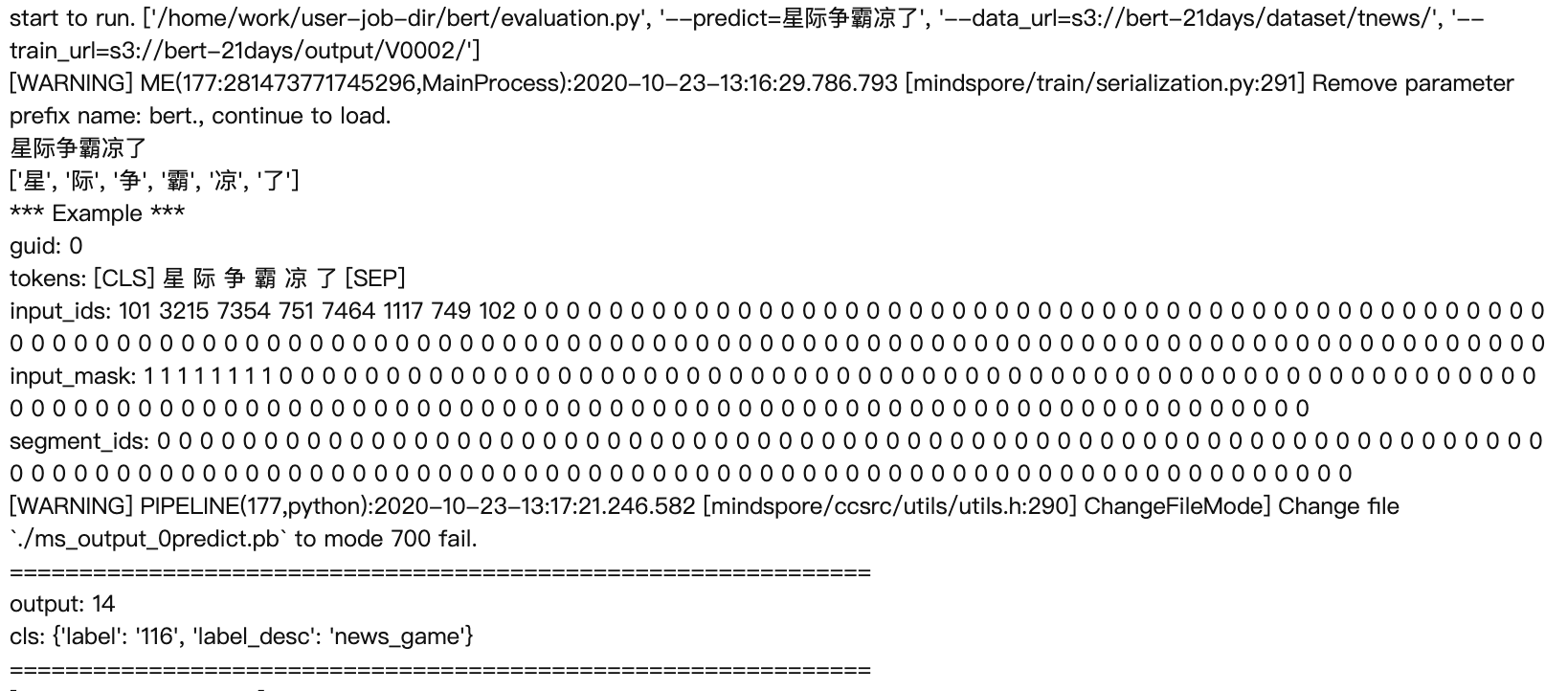


在左边的框中，输入“predict”，而在右边的参数框中，大家可以自由发挥输入一句中文的短新闻，我这里输入的是“星际争霸凉了”：



其他参数保持不变，提交作业

同样等待作业完成之后，查看日志：



可以看到分类显示“新闻\_游戏”，到这一步，单句新闻的分类也完成了，大家可以截图上传作业