

ModelArts 平台使用指导

华为云开发者 AI 青年班黑客松大赛 Baseline

文档版本 01
发布日期 2020-04-27

华为技术有限公司



版权所有 © 华为技术有限公司 2019。 保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明



HUAWEI和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为公司对本文档内容不做任何明示或默示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

华为技术有限公司

地址： 深圳市龙岗区坂田华为总部办公楼 邮编：518129

网址： <http://www.huawei.com>

客户服务邮箱： support@huawei.com

客户服务电话： 4008302118

目录

1 文档导读	4
2 获取竞赛数据集.....	5
2.1 在 OBS 创建桶及文件夹.....	5
2.2 在 ModelArts 全局配置中添加访问密钥.....	6
2.3 创建 Notebook 并通过代码拷贝竞赛数据集.....	8
3 在 Notebook 编写训练脚本	10
4 将生成的模型导入至模型管理.....	13
5 将模型部署为在线服务	15
6 提交模型判分.....	17
7 释放资源	19

1 文档导读

ModelArts 是面向 AI 开发者的一站式开发平台，您可以使用 ModelArts 完成赛题数据处理、训练建模、推理测试以及结果的提交。

本文档包含使用 ModelArts 完成赛题的全流程指导，通过本文档您可以快速了解 ModelArts 平台的使用，有助于您使用 ModelArts 来完成赛题。

您可以根据如下表格快速查找您需要了解的内容。

章节	说明
获取竞赛数据集	获取赛题数据集。
在 Notebook 编写训练脚本	在 ModelArts Notebook 编写模型训练脚本，进行模型的训练。
将生成的模型导入至模型管理	将训练生成的模型导入至模型管理。
将模型部署为批量服务	将模型部署为在线服务，查看预测结果。
提交模型判分	提交测试好的模型进行评分。
释放资源	为避免资源浪费或账号欠费，使用完计算资源后需及时释放资源。
附录	本 baseline 中的训练脚本、推理代码。

2 获取竞赛数据集

ModelArts 使用对象存储服务（OBS）存储数据和模型的备份，实现安全、高可靠和低成本的存储需求。

竞赛数据集已上传至华为云对象存储服务（OBS）桶中，您首先需要在 OBS 创建一个属于您的空桶；然后在 ModelArts 开发环境中创建 Notebook；最后在 Notebook 中执行两行简单的代码，将竞赛数据集拷贝至您创建的 OBS 桶中。

2.1 在 OBS 创建桶及文件夹

首先在 OBS 创建桶及文件夹，用于存放竞赛数据集、代码文件及模型训练输出的模型文件，具体操作如下。

1. 登录 OBS 管理控制台
<https://storage.huaweicloud.com/obs/?®ion=cn-north-1&locale=zh-cn#/obs/buckets>。
2. 单击页面右上方“创建桶”，在弹出的页面中选择区域、自定义桶名称，可参考下图完成创建桶参数配置。本 baseline 创建的桶名称为 obs-aifood-baseline。

The screenshot shows the 'Create Bucket' (创建桶) interface in the OBS console. The 'Region' (区域) is set to '华北-北京四'. The 'Bucket Name' (桶名称) is 'obs-aifood-baseline'. The 'Storage Class' (存储类别) is '标准存储'. The 'Access Policy' (桶策略) is '私有'. The 'Encryption' (默认加密) is '关闭'. The 'Data Expiration' (归档数据直接) is '关闭'. The 'Multi-AZ' (多AZ) is '关闭'. The 'Create Bucket' button is at the bottom right.

注意

“区域”请选择“华北-北京四”。

OBS 桶名称全局唯一，不能与已有的任何桶名称重复。

建议“多 AZ”选择“关闭”，如选择开启会采用较高的计费标准。

3. 单击“立即创建”。
4. 在 OBS 桶列表单击桶名称，然后单击左侧导航栏中“对象”，单击页面中“新建文件夹”，自定义文件夹名称后，单击“确定”。新建文件夹详细操作可参见 https://support.huaweicloud.com/usermanual-obs/zh-cn_topic_0045829103.html。
 - 新建一个文件夹存放竞赛数据集；本 baseline 存放竞赛数据集路径为 obs-aifood-baseline/aifood。
 - 新建一个文件夹存放代码文件；本 baseline 存放代码文件路径为 obs-aifood-baseline/codes。
 - 新建一个文件夹存放训练输出文件，然后在该文件夹下新建一个存放模型文件的文件夹；本 baseline 存放模型文件路径为 obs-aifood-baseline/model_output。



2.2 在 ModelArts 全局配置中添加访问密钥

使用 ModelArts Notebook、训练作业、模型及服务时可能需要用到对象存储服务，因此在使用 ModelArts 之前，您需要添加访问密钥。

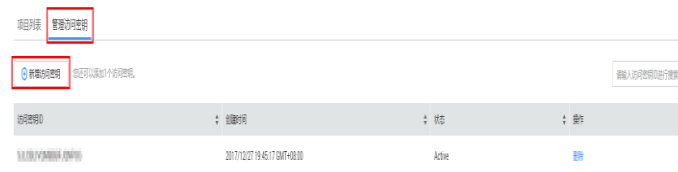
添加访问密钥前，您需要在“我的凭证”页面创建访问密钥，然后在 ModelArts 全局配置中添加访问密钥，具体操作如下。

创建访问密钥

1. 登录[华为云](https://console.huaweicloud.com/iam/#/myCredential)，打开“我的凭证”页面（您可直接单击 <https://console.huaweicloud.com/iam/#/myCredential>，进入“我的凭证”页面）。



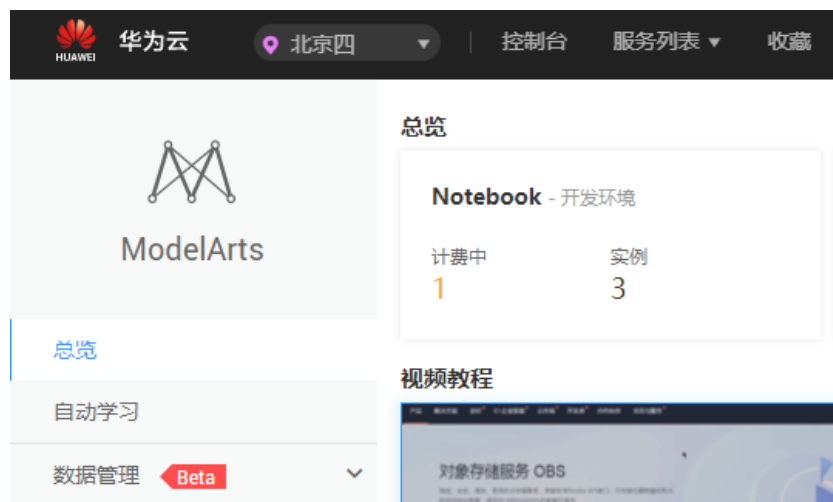
2. 在“我的凭证”页面，单击“管理访问密钥”页签下方的“新增访问密钥”。



3. 在弹出的“新增访问密钥”对话框，输入当前用户的登录密码，通过已验证手机或已验证邮箱进行验证，输入对应的验证码。
4. 单击“确定”，根据浏览器提示，保存密钥文件。密钥文件会直接保存到浏览器默认的下载文件夹中。
5. 打开下载下来的“credentials.csv”文件，即可获取到访问密钥（Access Key Id 和 Secret Access Key）。

添加访问密钥

1. 登录 ModelArts 管理控制台。
在管理控制台上方导航栏，单击“服务列表”，选择“EI 企业智能 > ModelArts”，进入 ModelArts 管理控制台，然后区域选择“华北-北京四”。



2. 在 ModelArts 左侧导航栏选择“全局配置”，单击页面中的“添加访问密钥”，在弹出的对话框中，填写获取的访问密钥。
 - “访问密钥（AK）”输入密钥文件中的 Access Key Id 字段内容，“私有访问密钥（SK）”输入密钥文件中 Secret Access Key 字段内容。
 - 请确保所填写的 AK、SK 为当前账号所获取的。
3. 单击“确认”，完成访问密钥的添加。

2.3 创建 Notebook 并通过代码拷贝竞赛数据集

本节介绍如何在 ModelArts 开发环境创建 Notebook，并在 Notebook 中通过两行代码将竞赛数据集拷贝至您的 OBS 桶中，具体操作如下。

1. 在 ModelArts 左侧导航栏选择“开发环境 > Notebook”，单击页面中“创建”。
2. 在创建 Notebook 页面填写参数，可参考下图填写配置参数。

注意

“存储配置”建议选择“对象存储服务”，并选择您创建的存放代码文件的 OBS 路径，这样您在 Notebook 编写的代码文件会自动保存在该路径下。本 baseline 存放代码的 OBS 路径为 `/obs-aifood-baseline/baseline/codes`。

The screenshot displays the 'Create Notebook' configuration interface in ModelArts. It includes sections for basic information (Plan Mode, Name, Auto Stop, Auto Stop Time, Description), environment settings (Work Environment, Resource Pool, Type, Instance), and storage configuration (Storage Configuration, Storage Location). The 'Storage Configuration' section is highlighted, showing the selection of 'Object Storage Service' and the path '/obs-aifood-baseline/codes/'. The bottom of the page shows the estimated cost of ¥7.60 per hour and a 'Next Step' button.

3. 单击“下一步”进行规格确认，确认无误后单击“立即创建”。
4. 返回 Notebook 列表，单击 Notebook 作业列表操作栏中“打开”，打开 Notebook。
5. 单击“Files”页签右上角“New”，选择“Pytorch-1.0.0”进入代码开发界面。

在代码输入栏中输入如下代码，即可从 OBS 公共桶中将竞赛数据拷贝至您创建的 OBS 桶中。**本 baseline 使用的是 Pytorch 框架，您也可以根据自身使用习惯选择其他框架。**

```
import moxing as mox

mox.file.copy_parallel('s3://obs-aifood-bj4/aifood',
's3://obs-aifood-baseline/aifood')

print('Copy procedure is completed !')
```

说明

将“obs-aifood-baseline/aifood”替换为您创建的存放竞赛数据集的 OBS 路径。

MoXing 是华为云 ModelArts 团队自研的分布式训练加速框架，并提供了一套文件对象 API，可以用来读写 ModelArts Notebook 容器本地文件及 OBS 文件。

竞赛数据拷贝成功后，可在您的 OBS 桶中查看，如下是 Baseline 的 OBS 桶拷贝成功的截图。



3 在 Notebook 编写训练脚本

获取竞赛数据之后，您可以在 ModelArts Notebook 中编写模型训练脚本；如果您训练的模型需要推理代码，也可以在 Notebook 中编写。

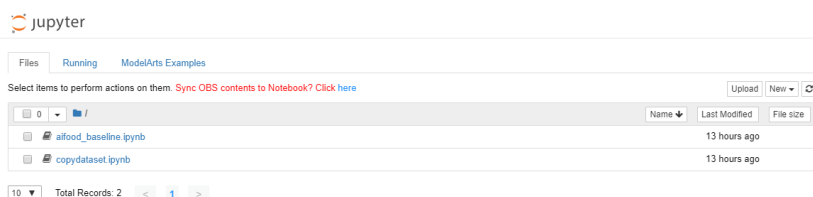
ModelArts 支持 TensorFlow、MXNet、PyTorch 等深度学习引擎，您可以基于自身使用习惯来选择；您还可以使用 `pip install` 在 ModelArts Notebook 中安装外部库。

编写模型训练脚本

打开 Notebook 后，单击“Files”页签右上角“New”，选择“Pytorch-1.0.0”进入代码开发界面，然后您可以进行模型训练脚本的编写和运行。

本 baseline 已经在 ModelArts Notebook 中编写并调试好模型训练脚本，即附录中的 `aifood_baseline.ipynb`，先下载至本地，然后将该训练脚本 Upload 至 Notebook，修改部分参数即可运行，具体操作如下。。

1. 打开 Notebook 进入代码开发界面后，单击“Upload”，将 `aifood_baseline.ipynb` 上传至 Notebook。



2. 单击文件名称，进入代码编写/运行界面，按照 ipynb 文件内容修改参数，运行所有代码。

编写推理代码

如果您的模型需要推理代码，您可以在 ModelArts Notebook 中编写，编写完成后将推理代码保存并上传至 OBS 存放模型的 model 文件夹下。

本 baseline 已经编写好模型推理代码，您可以将附录中的推理代码 `customize_service.py` 上传至您创建的 OBS 桶中（存放模型的 `model` 文件夹下）。

如果您自己编写推理代码，具体操作可参考如下步骤。

1. 打开 Notebook，单击“Files”页签右上角“New”，选择“Python 3”进入代码开发界面，然后您可以进行推理代码的编写。

模型推理代码的编写可参考

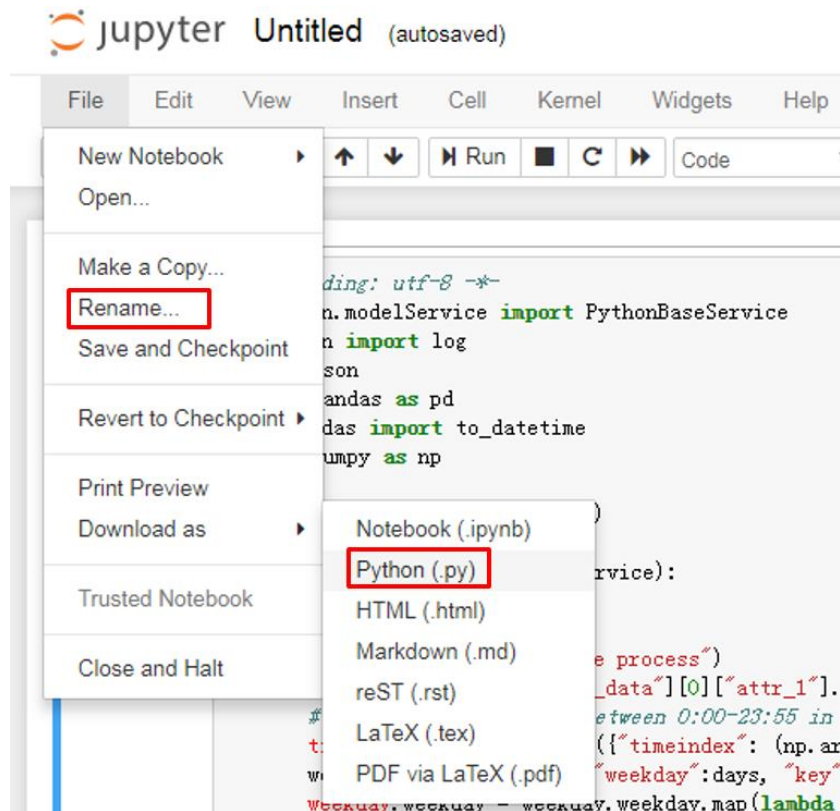
https://support.huaweicloud.com/engineers-modelarts/modelarts_23_0093.html。

2. 编写完成后，单击“File”页签中“Rename”，在弹出的对话框中修改文件名称为“`customize_service`”。

注意

推理代码名称固定为 `customize_service`，且为 `.py` 类型文件。

3. 单击“File”页签中“Download as > Python(.py)”，将文件下载至本地。



4. 将下载的 `customize_service.py` 文件上传至 OBS 存放模型的 `model` 文件夹下。

OBS 上传文件操作可参考

https://support.huaweicloud.com/usermanual-obs/zh-cn_topic_0045829661.html。

4 将生成的模型导入至模型管理

模型训练完成且编写好模型推理代码后，您可以将生成的模型导入至模型管理。

导入模型前，需要将附件中本 baseline 模型配置文件、推理代码，以及美食图片标签文件 labels_10c.txt 上传至 model 文件夹下。本 baseline 是上传至/obs-aifood-baseline/model_output/model 目录下。



导入模型具体操作如下。

1. 在 ModelArts 左侧导航栏选择“模型管理”，单击页面中“导入”。
2. 在导入模型页面填写名称，选择元模型来源；其中元模型来源如从 OBS 中选择，请选择 model 文件夹上一级目录；本 baseline model 文件夹上一级目录为 obs-aifood-baseline/model_output。

5 将模型部署为在线服务

导入模型完成后，接下来您可以将模型部署为在线服务。

部署为批量服务具体步骤如下。

1. 在 ModelArts 左侧导航栏中选择“部署上线 > 在线服务”，然后单击页面中的“部署”。
2. 在部署页面填写参数。

The screenshot shows the deployment configuration interface in ModelArts. It includes fields for service name, auto-stop settings, description, resource pool selection, and model configuration.

*** 名称** service-aifood

是否自动停止 ☒ 开启该选项后，在线服务的运行时间将在您选择的时间点后，自动停止，同时服务计费停止。

☒ 1小时后 ☐ 2小时后 ☐ 4小时后 ☐ 6小时后 ☐ 自定义

描述

0/100

*** 资源池** ☒ 公共资源池 ☐ 专属资源池

*** 选择模型及配置**

模型	model-aifood	0.0.1	分流 (%)	<input type="text" value="100"/>
计算节点规格	CPU: 2 核 8 GiB	计算节点个数	<input type="text" value="1"/>	
环境变量	增加环境变量			

3. 单击“下一步”，参数确认无误后，单击“提交”。

创建成功后，返回在线服务列表，您可以在在线服务列表中查看部署进度；当状态变为“运行中”后，单击服务名称，进入详情页面，如下图所示。

在线服务 > service-aifood

名称	service-aifood	服务ID	fd0b1b55-649d-4c1d-898d-3374250d241d
状态	运行中 (53 分钟后停止)	来源	我的部署
调用失败次数/总次数	0 / 2 详情	网络配置	未设置
描述	- 编辑	数据采集	未开启 编辑

[调用指南](#) [预测](#) [配置更新记录](#) [数据采集](#) [监控信息](#) [事件](#) [日志](#) [共享](#) [溯源图](#)

API接口地址 <https://08a2ca2d3fd042e781138954d81360b1.apigw.cn>

模型: model-aifood 0.0.1

参数配置

POST /	
输入参数	
名称	类型
images	file
输出参数	
名称	类型
logits	object

在“预测”页签添加图片进行测试，如下图所示。

名称service-aifood

状态运行中 (53 分钟后停止)

调用失败次数/总次数0 / 2 [详情](#)

描述- [编辑](#)

服务IDfd0b1b55-649d-4c1d-898d-3374250d241d

来源我的部署

网络配置未设置

数据采集未开启 [编辑](#)

[调用指南](#) [预测](#) [配置更新记录](#) [数据采集](#) [监控信息](#) [事件](#) [日志](#) [共享](#) [溯源图](#)

请求路径: / [选择预测图片文件](#) [上传](#) [添加标签](#) [清除标签](#)

预测图片预览



预测结果显示

✓ 预测成功

```
1 {
2   "logits": {
3     "三鞭汁": 0.006,
4     "小蜜梨": 0.976,
5     "土豆泥": 0.004,
6     "小米粥": 0,
7     "松茸鱼": 0.004,
8     "辣白菜": 0.001,
9     "玉米粥": 0.007,
10    "甜酸鱼": 0.001,
11    "芒果蛋糕": 0,
12    "鸡蛋布丁": 0
13  }
14 }
```


6 提交模型判分

模型部署测试后，您可以将模型提交判分，具体操作如下。

注意：由于每支队伍每天有提交次数限制，为避免浪费提交次数，请在模型提交前务必要部署测试模型，确认模型无误后再提交判分。

步骤一：在ModelArts上提交模型

1. 在 ModelArts 左侧导航栏中选择“模型管理”，单击模型名称前方箭头；然后单击页面右侧操作栏中的“发布”，单击“参赛发布”。



2. 在弹出的“参赛模型提交”对话框中，选择比赛项目、比赛阶段，确认模型名称正确后，然后单击“确定”。

×

参赛模型提交

* 比赛项目

选择您参与的比赛

* 比赛阶段

选择比赛阶段

[查看比赛主页](#)

提交限额

0 / 5 (每日限额)

模型名称

确认模型名称是否正确

模型ID

e9d3fbaf-64d9-45d5-9735-5bf46c6d6f38

模型版本

0.0.5

温馨提示：您提交的参赛模型将由ModelHub提供调度和判分服务

确定

取消

确定后，即成功提交模型判分。在如下界面中可点击“现在加入”，也可以点击“以后再说”或直接点击右上角关掉该对话框。

×

参赛模型提交

✓

您的参赛模型提交申请已成功受理

小M推荐您加入AI市场，与其他参赛小伙伴，一起分享算法经验。甚至还可以分享和交易自己开发的模型。这是一个有趣的社区，将会包含数据集、案例集、模型算法等丰富的资源。现在加入，即可获取参赛模型提交结果，**判分结果的邮件实时通知**哦！

我是ModelHub,你的小M，我为AI市场代言

现在加入

以后再说

步骤二：在竞赛平台上查看分数

在大赛平台上“提交作品”页签中查看判分分数。其中“提交作品”页面需报名比赛后才会显示。

18

7 释放资源

使用完 ModelArts 相关计算资源后，请务必及时释放资源，否则会造成资源浪费甚至账号欠费。

ModelArts 以下模块会涉及云资源的消耗并会按需或按周期产生费用：自动学习训练和部署，Notebook 实例，训练作业，TensorBoard，在线服务。使用完需要“停止”自动学习部署、Notebook、TensorBoard、在线服务。

停止所有需消耗云资源的作业后，可在 ModelArts 总览页面确认所有计费项都为“0”。



The screenshot shows the ModelArts console interface. On the left is a navigation menu with options like '总览' (Overview), '自动学习' (AutoML), '数据标注' (Data Annotation), '数据管理' (Data Management), '开发环境' (Development Environment), '训练作业' (Training Jobs), '模型管理' (Model Management), and '部署上线' (Deployment). The main area is titled '总览' (Overview) and displays usage statistics for different services. Red boxes highlight the '计费中' (Billing) status for several items, indicating zero usage.

自动学习		物体检测	
图像分类	0 计费中	0 计费中	
1 项目		2 项目	
● eeeML-butterfly		● eeeML-yunbao	

AI全流程开发				
数据管理	开发环境	训练作业	模型管理	部署上线
0 计费中	0 计费中	0 计费中	0 计费中	0 计费中
4 数据集	9 实例	7/1 作业服务/TensorBoard	19 模型	18 服务
dataset-flowers	● notebook-训练	● trainjob-训练	model-训练	● service-训练

ModelArts 详细计费说明可参见：

https://support.huaweicloud.com/price-modelarts/modelarts_07_0002.html。

注意

公有云的按需使用是采用后付费模式，即先使用再付费，一个小时会出一次话单，比如：2019.6.24 04:00~05:00 这段时间使用的费用，在 05:00 时刻采集完成后，才会出话单，进行扣费。

8 附录

如下 baseline files.zip 文件包含了本指导中涉及的代码文件，文件说明见下表。



baseline files.zip

文件	说明
copydataset.ipynb	拷贝竞赛数据集代码文件
aifood_baseline.ipynb	模型训练代码文件（Upload 至 ModelArts Notebook）
readme.docx	介绍如何在 ModelArts Notebook 中运行代码
customize_service.py	模型推理代码（导入模型前，需上传至 model 目录下）
config.json	模型配置文件（导入模型前，需上传至 model 目录下）
labels_10c.txt	部署预测时需要的标签文档（导入模型前，需上传至 model 目录下）