

ModelArts

工具指南

文档版本 08
发布日期 2022-06-06



版权所有 © 华为技术有限公司 2022。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暗示的声明或保证。

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：<https://www.huawei.com>

客户服务邮箱：support@huawei.com

客户服务电话：4008302118

目录

1 PyCharm ToolKit 介绍	1
2 准备工作	3
2.1 下载并安装 ToolKit 工具	3
2.2 创建访问密钥（AK 和 SK）	5
2.3 使用访问密钥登录	6
3 PyCharm ToolKit（latest 版本）	8
3.1 训练模型	8
3.1.1 提交训练作业（新版训练）	8
3.1.2 停止训练作业	12
3.1.3 查看训练日志	12
4 PyCharm ToolKit（2.x 版本）	14
4.1 训练模型	14
4.1.1 提交训练作业（旧版训练）	14
4.1.2 查看训练作业详情	18
4.1.3 启动或停止训练作业	21
4.1.4 查看训练日志	21
4.1.5 提交不同名称的训练作业	23
4.2 部署上线	26
5 OBS 上传下载	30
6 常见问题	34
6.1 安装 ToolKit 工具时出现错误，如何处理？	34
6.2 为什么无法启动训练？	34
6.3 提交训练作业时，出现 xxx isn't existed in train_version 错误	35
6.4 提交训练作业报错	35
6.5 部署上线时，出现错误	36
6.6 如何查看 PyCharm ToolKit 的错误日志	36
A 修订记录	37

1 PyCharm ToolKit 介绍

由于AI开发者会使用PyCharm工具开发算法或模型，为方便快捷将本地代码提交到ModelArts的训练环境，ModelArts提供了一个PyCharm插件工具PyCharm ToolKit（插件下载请参见[下载并安装ToolKit工具](#)），协助用户完成代码上传、提交训练作业、将训练日志获取到本地展示等，用户只需要专注于本地的代码开发即可。

使用限制

- 当前仅支持PyCharm 2019.2及以上版本，包括社区版和专业版。
- 使用PyCharm ToolKit远程连接Notebook开发环境，仅限PyCharm专业版。
- 使用PyCharm ToolKit提交训练作业，社区版和专业版都支持，PyCharm ToolKit 2.x版本仅限于提交旧版训练作业，PyCharm ToolKit latest版本仅限提交新版训练作业。
- PyCharm ToolKit工具支持Windows、Linux或Mac版本的PyCharm。

支持的功能

表 1-1 ToolKit（latest）功能列表

支持的功能	说明	对应操作指导
SSH远程连接	支持SSH远程连接ModelArts的Notebook开发环境。	配置PyCharm ToolKit远程连接Notebook
训练模型	支持将本地开发的代码，快速提交至ModelArts并自动创建新版训练作业，在训练作业运行期间获取训练日志并展示到本地。	<ul style="list-style-type: none">• 提交训练作业（新版训练）• 停止训练作业• 查看训练日志
OBS上传下载	上传本地文件或文件夹至OBS，从OBS下载文件或文件夹到本地。	OBS上传下载

表 1-2 ToolKit（2.x 版本）功能列表

支持的功能	说明	对应操作指导
SSH远程连接	支持SSH远程连接ModelArts的Notebook开发环境。	配置PyCharm ToolKit远程连接Notebook
训练模型	支持将本地开发的代码，快速提交至ModelArts并自动创建旧版训练作业，在训练作业运行期间获取训练日志并展示到本地。	<ul style="list-style-type: none">• 提交训练作业（旧版训练）• 查看训练作业详情• 启动或停止训练作业• 查看训练日志• 提交不同名称的训练作业
部署上线	支持将训练好的模型快速部署上线为在线服务。	部署上线
OBS上传下载	上传本地文件或文件夹至OBS，从OBS下载文件或文件夹到本地。	OBS上传下载

2 准备工作

2.1 下载并安装 ToolKit 工具

ModelArts提供了PyCharm ToolKit工具的下载地址，在使用此工具之前，您需要通过此链接下载工具包，并根据如下操作指导完成在PyCharm中的安装配置。

前提条件

本地已安装2019.2及以上版本的PyCharm社区版或专业版。

- 使用PyCharm ToolKit远程连接Notebook开发环境，仅限PyCharm专业版。
- 使用PyCharm ToolKit提交训练作业，社区版和专业版都支持，但2.x版本仅限于提交旧版训练作业，latest版本仅限于提交新版训练作业。

下载 ToolKit 工具包

通过ToolKit工具的下载地址，获得工具包。

表 2-1 PyCharm ToolKit 工具包下载地址

局点	PyCharm ToolKit下载地址
华为云	中国站（latest）： https://modelarts-pycharm-plugin.obs.cn-north-1.myhuaweicloud.com/Pycharm-ToolKit-latest.zip 中国站（2.x）： https://modelarts-pycharm-plugin.obs.cn-north-1.myhuaweicloud.com/Pycharm-ToolKit-2.2.1.zip

图 2-1 ModelArts 控制台上的训练管理

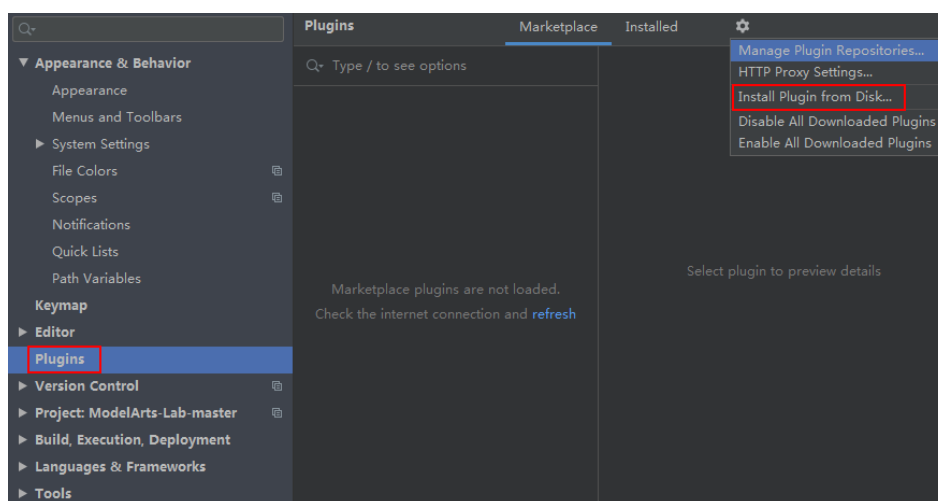


在 PyCharm 中安装 ToolKit 工具

请根据如下操作指导，将下载的ToolKit工具安装至PyCharm中。

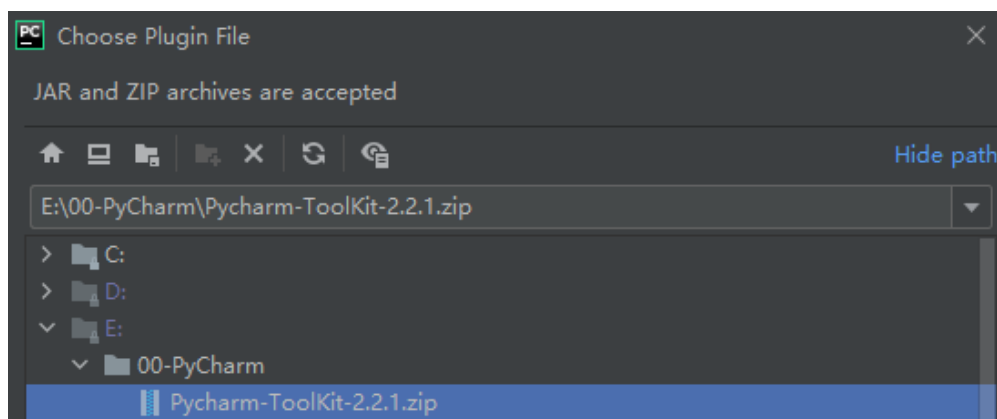
1. 打开本地PyCharm工具。
2. 在PyCharm工具中，选择菜单栏的“File > Settings”，弹出“Settings”对话框。
3. 在“Settings”对话框中，首先单击左侧导航栏中的“Plugins”，然后单击右侧的设置图标，选择“Install Plugin from Disk”，弹出文件选择对话框。

图 2-2 选择从本地安装插件



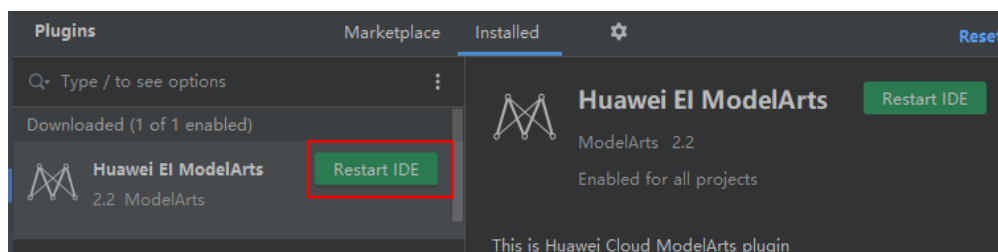
4. 在弹出的对话框中，从本地目录选择ToolKit的工具zip包，然后单击“OK”。

图 2-3 选择插件文件



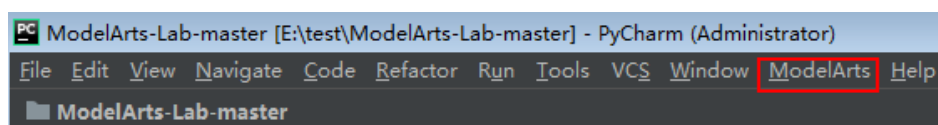
5. 单击“Restart IDE”重启PyCharm。在弹出的确认对话框中，单击“Restart”开始重启。

图 2-4 重启 PyCharm



6. 重启成功后，打开一个Project，当PyCharm工具栏出现“ModelArts”页签，表示ToolKit工具已安装完成。

图 2-5 安装成功



2.2 创建访问密钥（AK 和 SK）

本节介绍如何通过ModelArts管理控制台创建访问密钥（AK和SK），访问密钥（AK和SK）用于对请求加密签名，确保请求的机密性、完整性和请求双方身份的正确性。

创建 AK/SK

1. 登录华为云，在页面右上方单击“控制台”，进入华为云管理控制台。

图 2-6 控制台入口



2. 在控制台右上角的帐户名下方，单击“我的凭证”，进入“我的凭证”页面。

图 2-7 我的凭证



3. 在“我的凭证”页面，选择“访问密钥>新增访问密钥”，如图2-8所示。

图 2-8 单击新增访问密钥



4. 填写该密钥的描述说明，单击“确定”。根据提示单击“立即下载”，下载密钥。

图 2-9 新增访问密钥



5. 密钥文件会直接保存到浏览器默认的下载文件夹中。打开名称为“credentials.csv”的文件，即可查看访问密钥（Access Key Id和Secret Access Key）。

2.3 使用访问密钥登录

如果需要使得ToolKit与ModelArts完成连接，需使用当前帐号的访问密钥完成登录认证。

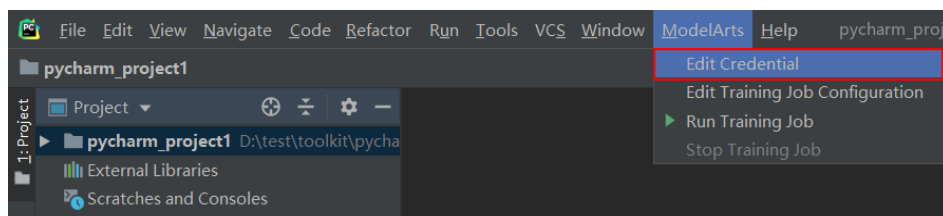
前提条件

- 已完成ToolKit工具安装。如果未安装，请参见[下载并安装ToolKit工具](#)。
- 已创建当前使用帐号的访问密钥，并获得对应的AK和SK。如果未创建，请参见[创建访问密钥（AK和SK）](#)。
- 在使用ToolKit之前，需要前往ModelArts控制台的全局设置页面，完成访问授权配置。如果ModelArts管理控制台未完成全局配置，登录ToolKit后，将无法访问并连接ModelArts。
- 确保当前帐号未欠费，且具备OBS和ModelArts的使用权限。

登录 ModelArts

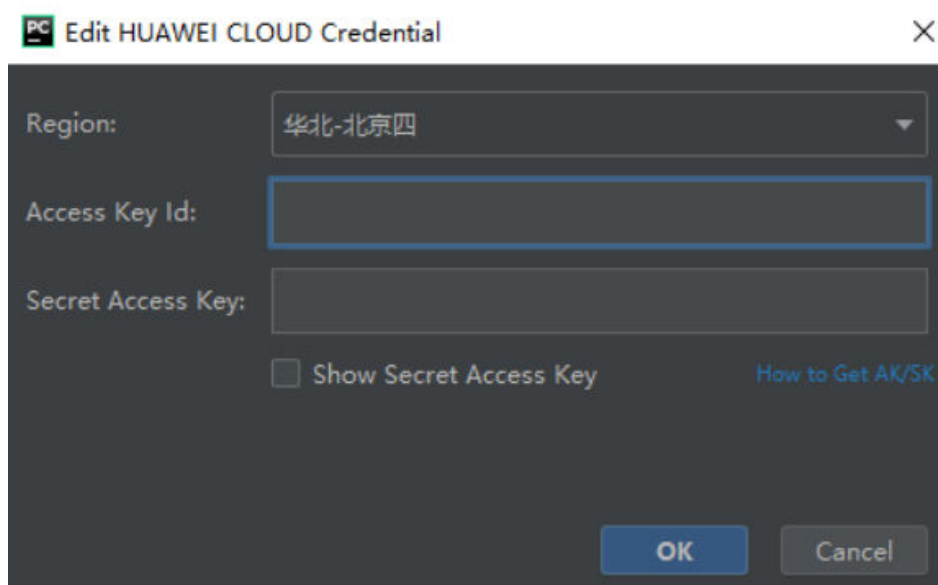
1. 打开已安装ToolKit工具的PyCharm，在菜单栏中选择“ModelArts > Edit Credential”。

图 2-10 Edit Credential



2. 在弹出的对话框中，选择您使用的ModelArts所在区域、填写AK、SK，然后单击“OK”完成登录。
 - “Region”：从下拉框中选择区域。目前支持“华北-北京四”、“华北-北京一”和“华东-上海一”区域，必须与ModelArts管理控制台在同一区域。
 - “Access Key ID”：填写访问密钥的AK。
 - “Secret Access Key”：填写访问密钥的SK。

图 2-11 填写区域和访问密钥



3. 查看认证结果。
在Event Log区域中，当提示如下类似信息时，表示访问密钥添加成功。

16:01Validate Credential Success: The HUAWEI CLOUDcredential is valid.

3 PyCharm ToolKit (latest 版本)

3.1 训练模型

3.1.1 提交训练作业（新版训练）

使用PyCharm ToolKit (latest版本) 工具，可以快速将本地开发的训练代码，提交至ModelArts侧进行训练。

前提条件

- 在本地PyCharm中已有训练代码工程。
- 已在OBS中创建桶和文件夹，用于存放数据集和训练输出模型。例如：创建命名为“test-modelarts”的桶，创建文件夹“dataset-mnist”和“mnist-output”。训练作业使用的数据已上传至OBS，且OBS与ModelArts在同一区域。
- 已配置credential，详细请参考[使用访问密钥登录](#)。
- 使用PyCharm ToolKit (latest版本) 提交训练作业，仅限于新版训练作业，不支持旧版训练作业。

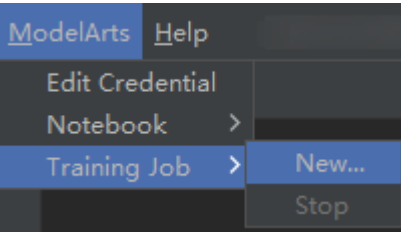
准备数据

- 训练代码工程案例请参考：在gitee的[ModelArts-Lab](#)工程中，单击“克隆/下载”，然后单击“下载ZIP”，下载工程。解压缩“ModelArts-Lab-master.zip”文件，然后在“\ModelArts-Lab-master\official_examples\Using_MXNet_to_Create_a_MNIST_Dataset_Recognition_Application\codes”目录中获取到训练代码文件“train_mnist.py”。
- 数据集案例请参考：单击[数据集下载链接](#)，将“Mnist-Data-Set”数据集下载至本地，然后解压zip包，将“Mnist-Data-Set”文件夹下的所有文件上传至“test-modelarts/dataset-mnist” OBS路径下。

配置训练作业参数

1. 在PyCharm中，打开训练代码工程和训练启动文件，然后在菜单栏中选择“ModelArts > Training Job > New...”。

图 3-1 选择作业配置



2. 在弹出的对话框中，设置训练作业相关参数，详细参数说明请参见[表3-1](#)。

表 3-1 训练作业配置参数说明

参数	说明
Job Name	训练作业的名称。
Job Description	训练作业的简要描述。
Algorithm Souce	训练算法来源，分为“常用框架”和“自定义镜像”两种，二者选一项即可。 常用框架指使用ModelArts训练管理中支持的常用AI引擎，当前支持的引擎列表请参见 训练管理支持的常用框架 。 如果您使用的AI引擎为支持列表之外的，建议使用自定义镜像的方式创建训练作业。
AI Engine	选择代码使用的AI引擎及其版本。支持的AI引擎与ModelArts管理控制台中训练作业支持的常用框架一致。
Boot File Path	训练启动文件，所选启动文件必须是当前PyCharm训练工程中的文件。
Code Directory	训练代码目录，系统会自动填写为训练启动文件所在的目录，用户可根据需要修改，所选目录必须是当前工程中的目录且包含启动文件。 当算法来源为自定义镜像，训练代码已预置在镜像中时，该参数可以为空。
Image Path (optional)	SWR镜像的URL地址，例如swr.cn-north-4.myhuaweicloud.com/image-org/image-name:version。关于自定义镜像的说明，请参见 自定义镜像介绍 。
Boot Command	启动本次训练作业的运行命令。例如“bash /home/work/run_train.sh python {python启动文件及参数}”。当“Algorithm Souce”选“Custom”时，显示此参数。 当用户输入的命令中不包含“--data_url”和“--train_url”参数时，工具在提交训练作业时会在命令后面自动添加这两个参数，分别对应存储训练数据的OBS路径和存放训练输出的OBS路径。
Data Obs Path	设置为存储训练数据的OBS路径，例如“/test-modelarts2/mnist/dataset-mnist/”，其中“test-modelarts”为桶名称。

参数	说明
Training Obs Path	设置OBS路径，该路径下会自动创建用于存放训练输出模型和训练日志的目录。
Running Parameters	运行参数。如果您的代码需要添加一些运行参数，可以在此处添加，多个运行参数使用英文分号隔开，例如 "key1=value1;key2=value2"。此参数也可以不设置，即保持为空。
Specifications	训练使用资源类型。目前支持公共资源池和专属资源池两种类型。 专属资源池规格以“Dedicated Resource Pool”标识。只有购买了专属资源池的用户才会显示专属资源池规格。
Compute Nodes	计算资源节点个数。数量设置为1时，表示单机运行；数量设置大于1时，表示后台的计算模式为分布式。
Available/Total Nodes	当“Specifications”选择专属资源池规格时，显示专属资源池的可用节点数和总节点数，用户选择“Compute Nodes”的个数不要超过可用节点数。

图 3-2 配置训练作业参数（公共资源池）

Job Name: MA-new-model-03-12-17-12

Job Description:

Frequently-used Custom

Algorithm Source: AI Engine: TensorFlow TF-1.15.0

Boot File Path: F:\DevTest\work\testtsts.py

Code Directory: F:\DevTest\work

Image Path(optional):

Training Obs Path: obs://[bucket]/code/

Data Obs Path: obs://[bucket]/data/dataset/

Specifications: CPU: 4vCPU

Compute Nodes: 1

Running Parameters:

Apply and Run Cancel Apply

图 3-3 配置训练作业参数（专属资源池）

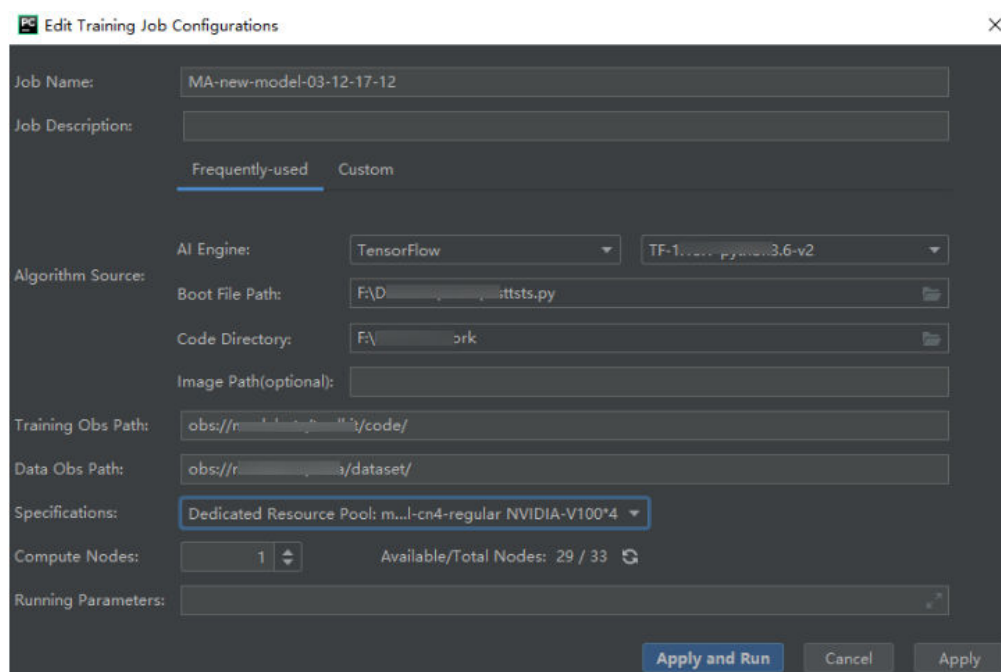
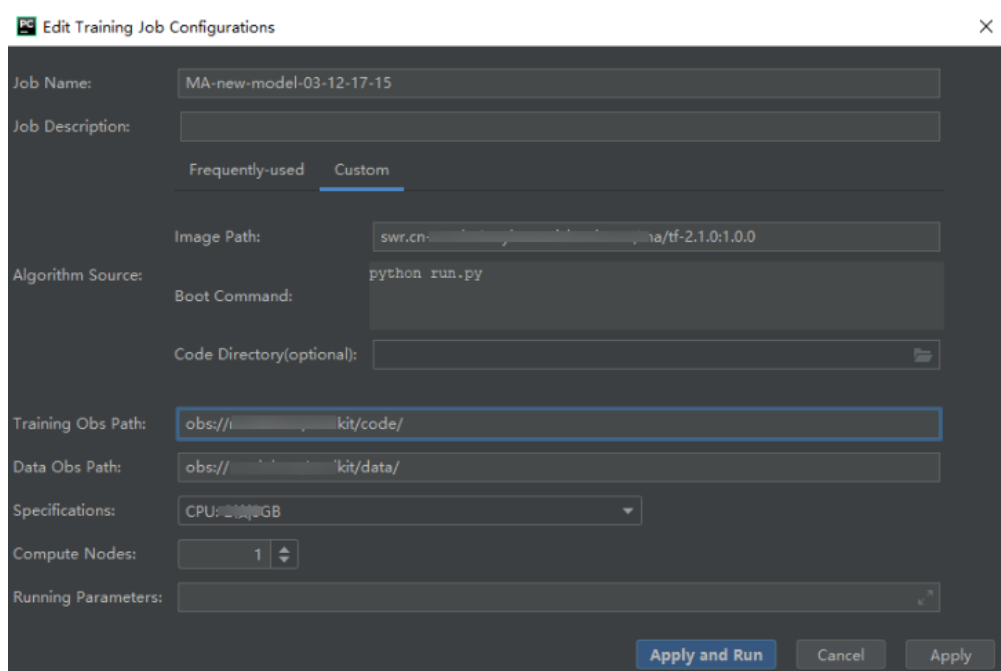


图 3-4 配置训练作业参数（自定义镜像）

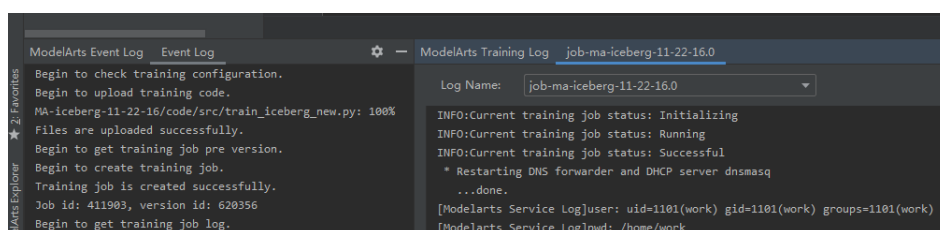


- 参数填写完成后，单击“Apply and Run”。即自动上传本地代码至云端并启动训练，在工具下方的Training Log区域，会实时展示训练作业运行情况。当训练日志中出现“Current training job status: Successful”类似信息时，表示训练作业运行成功。

说明

- 在单击“Apply and Run”按钮后，系统将自动开始执行训练作业。如果您想停止此作业，可以选择菜单栏中的“ModelArts > Training Job > Stop” 停止此作业。
- 如果单击“Apply”，不会直接启动运行，只是保存训练作业的设置，如果需要启动作业，可以单击“Apply and Run”。

图 3-5 训练日志展示样例



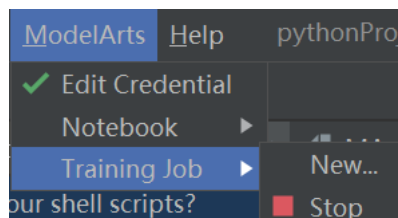
3.1.2 停止训练作业

当训练作业在运行过程中时，您可以执行停止作业的操作。

停止作业

当训练作业在运行过程中时，您可以在PyCharm菜单栏中，选择“ModelArts > Training Job > Stop” 停止此作业。

图 3-6 停止作业



3.1.3 查看训练日志

本章节介绍如何查看训练作业产生的日志。

在 OBS 中查看

提交训练作业时，系统将自动在您配置的OBS Path中，使用作业名称创建一个新的文件夹，用于存储训练输出的模型、日志和代码。

例如“train-job-01”作业，提交作业时会在“test-modelarts2”桶下创建一个命名为“train-job-01”的文件夹，且此文件夹下分别新建了三个文件夹“output”、“log”、“code”，分别用于存储输出模型、日志和训练代码。“output”文件夹还会根据您的训练作业版本再创建子文件夹，结构示例如下。

```
test-modelarts2
|---train-job-01
|   |---output
|   |---log
|   |---code
```


4 PyCharm ToolKit (2.x 版本)

4.1 训练模型

4.1.1 提交训练作业 (旧版训练)

使用PyCharm ToolKit (2.x版本) 工具，可以快速将本地开发的训练代码，提交至ModelArts侧进行训练。

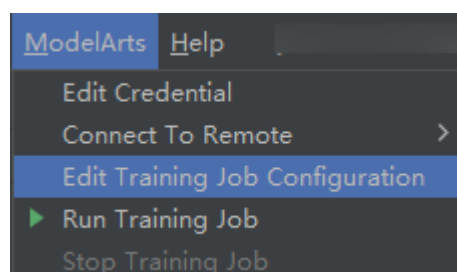
前提条件

- 在本地PyCharm中已有训练代码工程。
- 训练作业使用的数据已上传至OBS，且OBS与ModelArts在同一区域。
- 使用PyCharm ToolKit (2.x版本) 提交训练作业，仅限于旧版训练作业，暂不支持新版训练作业。

配置训练作业参数

1. 在PyCharm中，打开训练代码工程和训练启动文件，然后在菜单栏中选择“ModelArts > Edit Training Job Configuration”。

图 4-1 选择作业配置



2. 在弹出的对话框中，设置训练作业相关参数，详细参数说明请参见[表4-1](#)。

表 4-1 训练作业配置参数说明

参数	说明
Job Name	训练作业的名称。
Job Description	训练作业的简要描述。
Algorithm Souce	训练算法来源，分为“常用框架”和“自定义镜像”两种，二者选一项即可。 常用框架指使用ModelArts训练管理中支持的常用AI引擎，当前支持的引擎列表请参见 训练管理支持的常用框架 。 如果您使用的AI引擎为支持列表之外的，建议使用自定义镜像的方式创建训练作业。
AI Engine	选择代码使用的AI引擎及其版本。支持的AI引擎与ModelArts管理控制台中训练作业支持的常用框架一致。
Boot File Path	训练启动文件，所选启动文件必须是当前PyCharm训练工程中的文件。
Code Directory	训练代码目录，系统会自动填写为训练启动文件所在的目录，用户可根据需要修改，所选目录必须是当前工程中的目录且包含启动文件。 当算法来源为自定义镜像，训练代码已预置在镜像中时，该参数可以为空。
Image Path	SWR镜像的URL地址，例如swr.cn-north-4.myhuaweicloud.com/image-org/image-name:version。关于自定义镜像的说明，请参见 自定义镜像介绍 。
Boot Command	启动本次训练作业的运行命令。例如“bash /home/work/run_train.sh python {python启动文件及参数}” 当用户输入的命令中不包含“--data_url”和“--train_url”参数时，工具在提交训练作业时会在命令后面自动添加这两个参数，分别对应存储训练数据的OBS路径和存放训练输出的OBS路径。
OBS Path	设置OBS路径，该路径下会自动创建用于存放训练输出模型和训练日志的目录。
Data Path in OBS	设置为存储训练数据的OBS路径，例如“/test-modelarts2/mnist/dataset-mnist/”，其中“test-modelarts”为桶名称。
Specifications	训练使用资源类型。目前支持公共资源池和专属资源池两种类型。专属资源池规格以“Dedicated Resource Pool”标识。只有购买了专属资源池的用户才会显示专属资源池规格。专属资源池购买指导请参见 创建专属资源池 。
Compute Nodes	计算资源节点个数。数量设置为1时，表示单机运行；数量设置大于1时，表示后台的计算模式为分布式。

参数	说明
Available/Total Nodes	当“Specifications”选择专属资源池规格时，显示专属资源池的可用节点数和总节点数，用户选择“Compute Nodes”的个数不要超过可用节点数。
Running Parameters	运行参数。如果您的代码需要添加一些运行参数，可以在此处添加，多个运行参数使用英文分号隔开，例如"key1=value1;key2=value2"。此参数也可以不设置，即保持为空。

图 4-2 配置训练作业参数（公共资源池）

PC Edit Training Job Configurations

Job Name: MA-mnist-08-25-17-34

Job Description:

Frequently-used Custom

Algorithm Source:

AI Engine: TensorFlow TF-1.13.1-python3.6

Boot File Path: D:\Code\ModelArts\plugin-demo\mnist\train_mnist_tf.py

Code Directory: D:\Code\ModelArts\plugin-demo\mnist

OBS Path: test-modelarts2

Data Path in OBS: /test-modelarts2/mnist/dataset-mnist/

Specifications: GPU : 1*nvidia-v100 CPU : 8vCPUs 64GiB

Compute Nodes: 1

Running Parameters:

Apply and Run Cancel Apply

图 4-3 配置训练作业参数（专属资源池）

The screenshot shows the 'Edit Training Job Configurations' dialog box. The 'Job Name' is 'MA-mnist-08-20-14-44'. The 'Job Description' is empty. The 'Frequently-used' tab is selected. The 'Algorithm Source' section shows 'AI Engine' as 'TensorFlow' and 'TF-1.13.1-python3.6'. The 'Boot File Path' is 'D:\Code\ModelArts\plugin-demo\mnist\train_mnist_tf.py'. The 'Code Directory' is 'D:\Code\ModelArts\plugin-demo\mnist'. The 'OBS Path' is 'test-modelarts2'. The 'Data Path in OBS' is '/test-modelarts2/mnist/dataset-mnist/'. The 'Specifications' dropdown is set to 'Dedicated Resource Pool...00-pcie-32gb CPU:8U64GiB'. The 'Compute Nodes' is set to '1', and 'Available/Total Nodes' is '1 / 1'. The 'Running Parameters' field is empty. At the bottom right, there are buttons for 'Apply and Run', 'Cancel', and 'Apply'.

图 4-4 配置训练作业参数（自定义镜像）

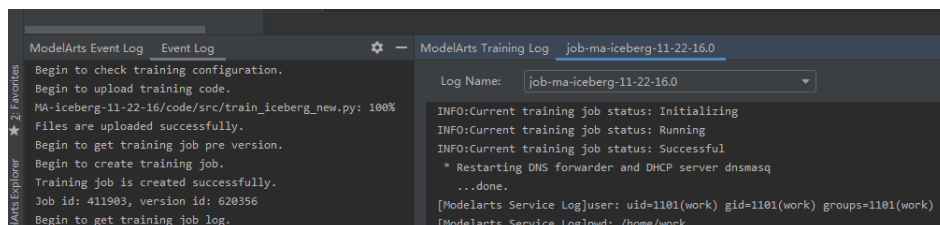
The screenshot shows the 'Edit Training Job Configurations' dialog box. The 'Job Name' is 'my-custom-image-job'. The 'Job Description' is empty. The 'Custom' tab is selected. The 'Image Path' is 'swr.cn-north-4.myhuaweicloud.com/modelarts-roma/train-tf-1_13_2:1.0.0'. The 'Algorithm Source' section shows 'Boot Command' as 'bash /home/work/run_train.sh /root/miniconda3/bin/python /home/work/user-job-dir/custom-image/mnist_softmax.py'. The 'Code Directory' is 'D:\Code\ModelArts\plugin-demo\mnist\custom-image'. The 'OBS Path' is 'test-modelarts2'. The 'Data Path in OBS' is '/test-modelarts2/mnist/dataset-mnist/'. The 'Specifications' dropdown is set to 'GPU : 1*nvidia-v100 CPU : 8vCPUs 64GiB'. The 'Compute Nodes' is set to '1'. The 'Running Parameters' field is empty. At the bottom right, there are buttons for 'Apply and Run', 'Cancel', and 'Apply'.

- 参数填写完成后，单击“Apply and Run”。即自动上传本地代码至云端并启动训练，在工具下方的Training Log区域，会实时展示训练作业运行情况。当训练日志中出现“Current training job status: Successful”类似信息时，表示训练作业运行成功。

说明

- 在单击“Apply and Run”按钮后，系统将自动开始执行训练作业。如果您想停止此作业，可以选择菜单栏中的“ModelArts > Stop Training Job”停止此作业。
- 如果单击“Apply”，不会直接启动运行，只是保存训练作业的设置，如果需要启动作业，可以选择菜单栏中的“ModelArts > Run Training Job”启动此作业。

图 4-5 训练日志展示样例



4.1.2 查看训练作业详情

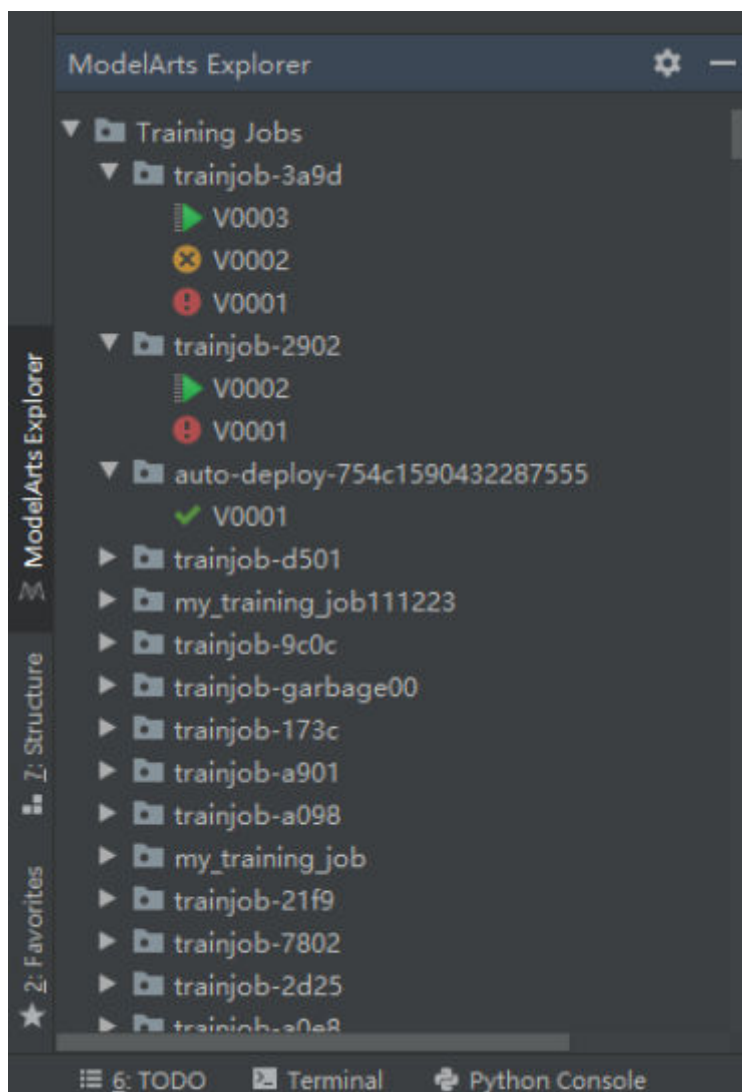
训练作业提交之后，您可以直接在PyCharm中查看此训练作业的详情。

在 PyCharm 中查看作业详情

训练作业运行成功后，您可以在PyCharm的左下边栏工具的“ModelArts Explorer”区域看到您提交的新训练作业。如果训练结束后未看到，可以尝试将“Training Job”折叠后再展开，等待页面刷新后，即可查看到您创建的作业。

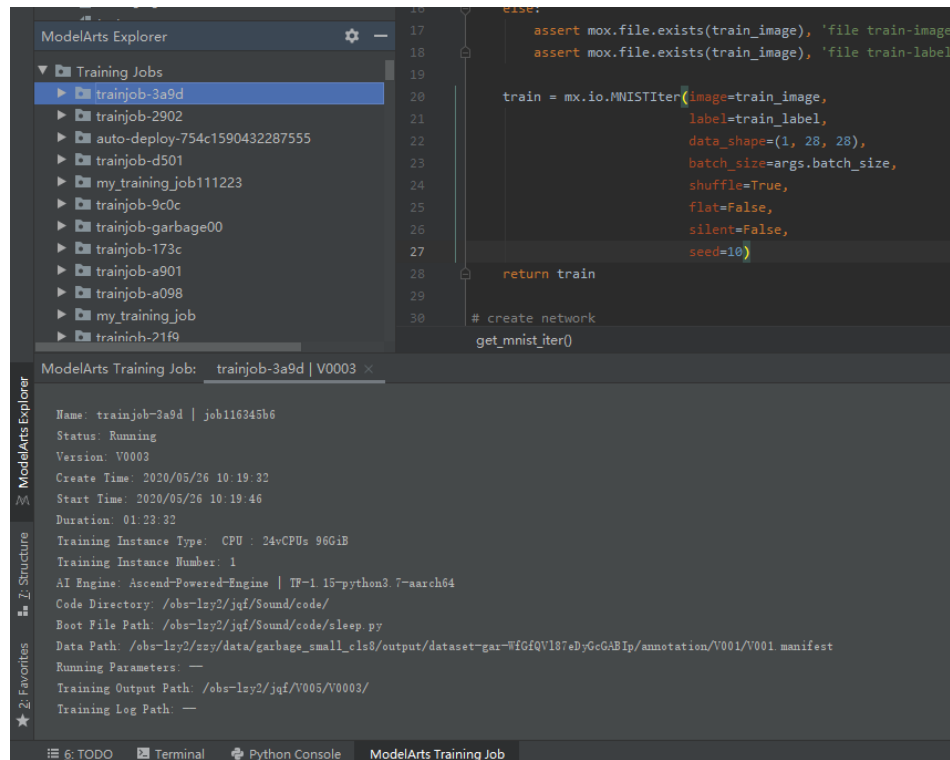
- 作业版本号前的图表代表作业的当前状态。
 - 绿色三角：运行中。
 - 绿色对勾：运行成功。
 - 黄色：已取消。
 - 红色：运行失败。

图 4-6 查看提交的训练作业



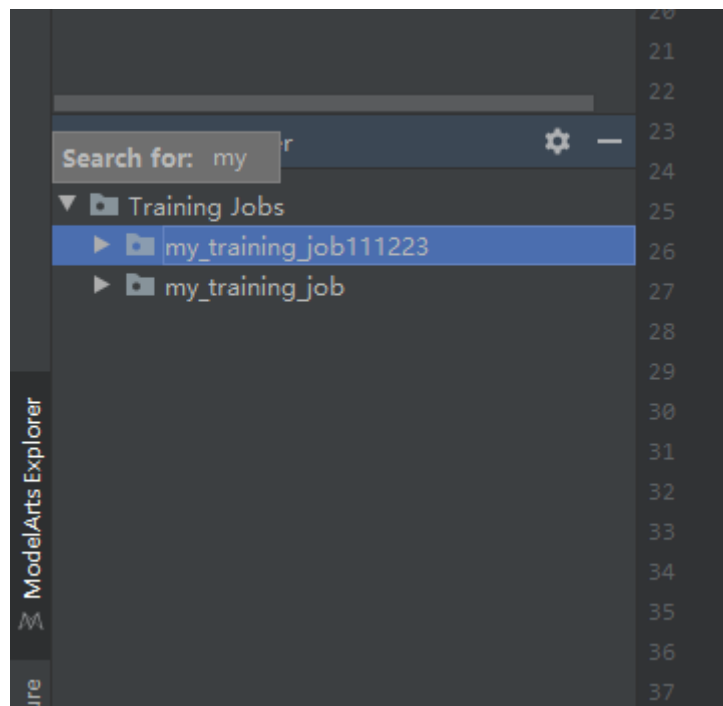
- 双击训练作业版本号V000x，可以查看作业详情。

图 4-7 查看训练作业详情



- ModelArts Explorer窗口支持快速搜索作业名称，鼠标单击窗口内的任何一个区域，直接输入作业名称即可实现搜索，作业列表只显示名称匹配的作业。按“Esc”键可以退出搜索模式。

图 4-8 搜索作业



4.1.3 启动或停止训练作业

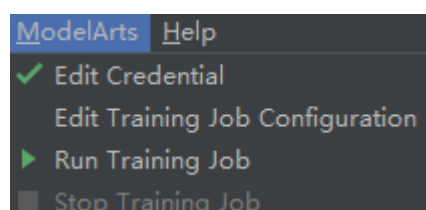
当训练作业在运行过程中时，您可以执行停止作业的操作。

针对完成配置但未启动、或者停止中的训练作业，您也可以执行启动作业操作。如果您需要更改配置后重新启动运行，建议您先停止作业，完成配置修改后，再执行启动操作。

启动作业

针对完成配置但未启动、或者停止中的训练作业，您可以在PyCharm菜单栏中，选择“ModelArts > Run Training Job”启动此作业。

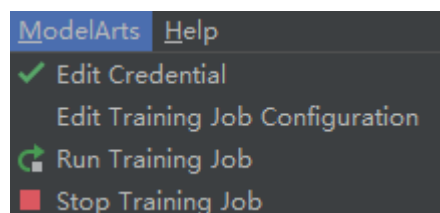
图 4-9 启动作业



停止作业

当训练作业在运行过程中时，您可以在PyCharm菜单栏中，选择“ModelArts > Stop Training Job”停止此作业。

图 4-10 停止作业



4.1.4 查看训练日志

每执行一次训练作业，将产生一个训练作业版本，版本号默认从V0001、V0002递增。您可以通过不同的训练作业版本，查看相应训练作业的日志。

训练作业的输出和日志路径说明

提交训练作业时，系统将自动在您配置的OBS Path中，使用作业名称创建一个新的文件夹，用于存储训练输出的模型、日志和代码。

例如“train-job-01”作业，提交作业时会在“test-modelarts2”桶下创建一个命名为“train-job-01”的文件夹，且此文件夹下分别新建了三个文件夹“output”、“log”、“code”，分别用于存储输出模型、日志和训练代码。“output”文件夹还会根据您的训练作业版本再创建子文件夹，结构示例如下。

```
test-modelarts2
|---train-job-01
|   |---output
|   |---V0001
```

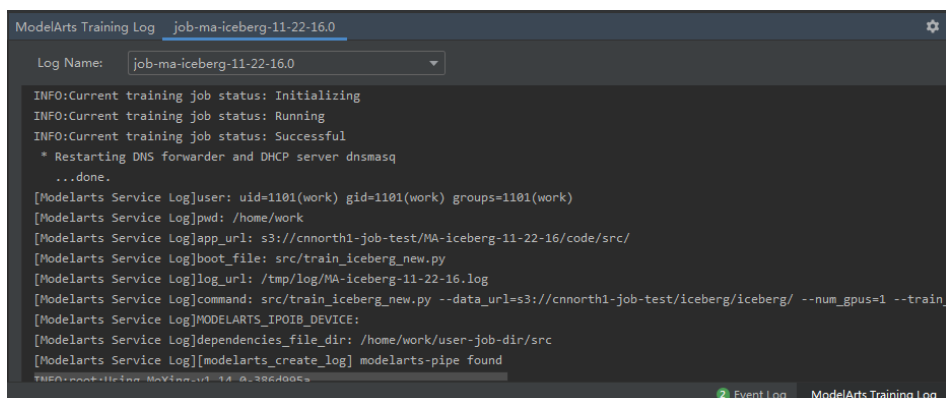


```
    |--V0002  
    |--...  
    |--log  
    |--code
```

在 ToolKit 工具中查看

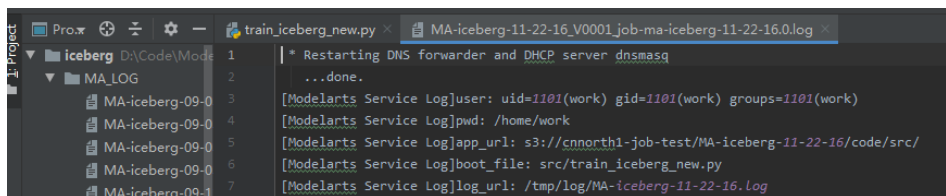
在PyCharm工具中，单击页面右下角的ModelArts Training Log，展示训练日志。

图 4-11 查看训练日志



同时，训练作业运行结束后，系统将自动在工程目录创建一个命名为“MA_LOG”的文件夹，用于保存训练作业日志文件。

图 4-12 工程目录下的日志文件

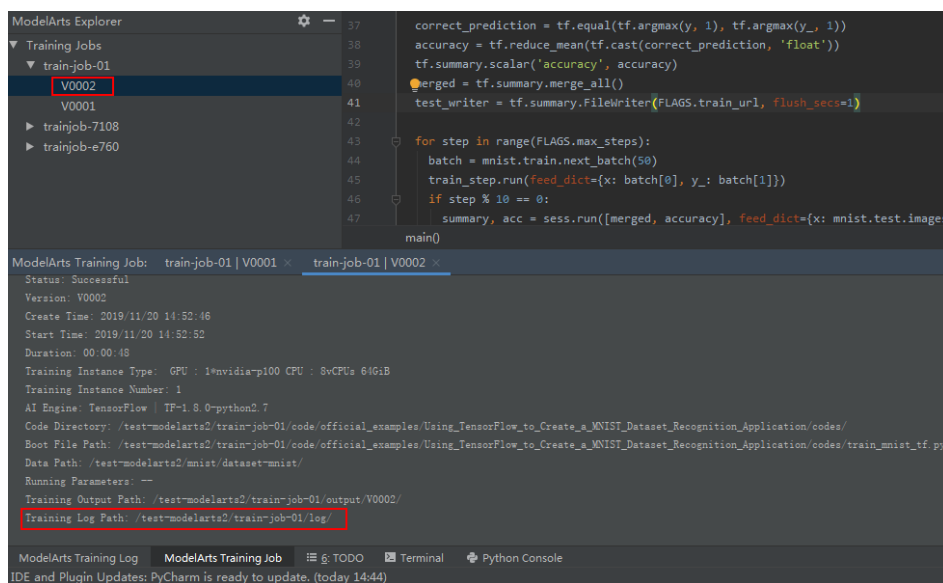


在 OBS 中查看

训练作业运行结束后，系统将日志存储至对应OBS路径下的log目录中。

1. 在PyCharm工具的ModelArts Explorer区域，双击对应训练作业的版本名称，将在ModelArts Training Job区域展示此训练作业的详细配置。
您可以在详细配置中，找到训练日志输出路径，即OBS路径。

图 4-13 查找日志存储路径



2. 使用当前帐号登录OBS管理控制台，根据上一个步骤获取的OBS路径，找到对应文件夹，查看训练作业的详细日志。

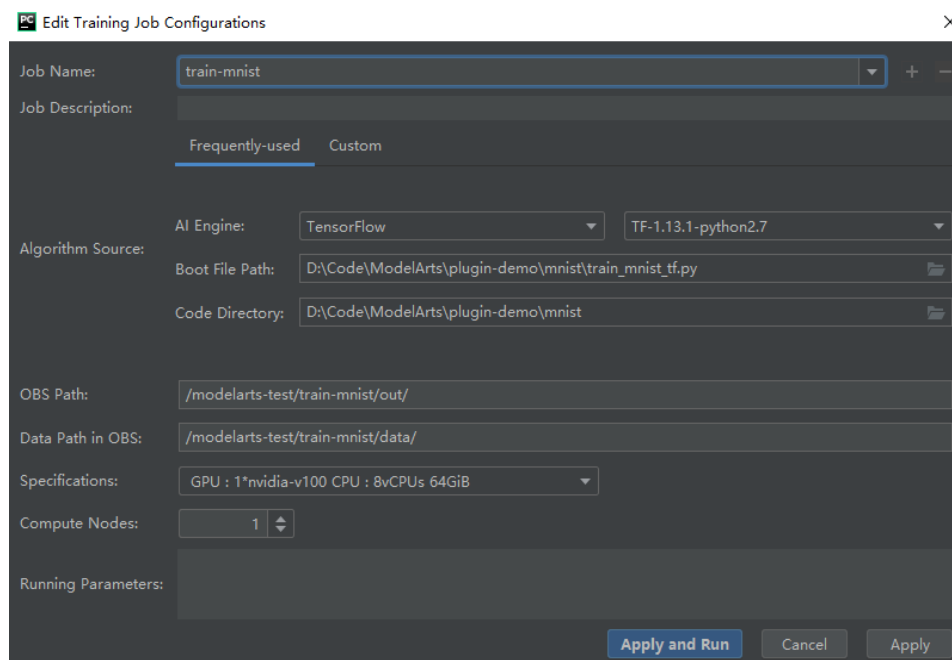
4.1.5 提交不同名称的训练作业

在ModelArts中，训练作业一旦提交后，作业名称不允许修改，每提交一次训练作业，实质上是创建了一个训练作业的新版本。如需提交一个不同名称的训练作业，请参见如下操作指导。

提交一个不同名称的训练作业

例如，当前已有一个手写数字识别的训练算法，提交了“Job Name”为“train-mnist”的训练作业，首次提交后，作业版本为V0001，运行完成后，用户对算法代码进行了优化，再次提交作业，此时系统会在作业“train-mnist”下创建V0002版本。

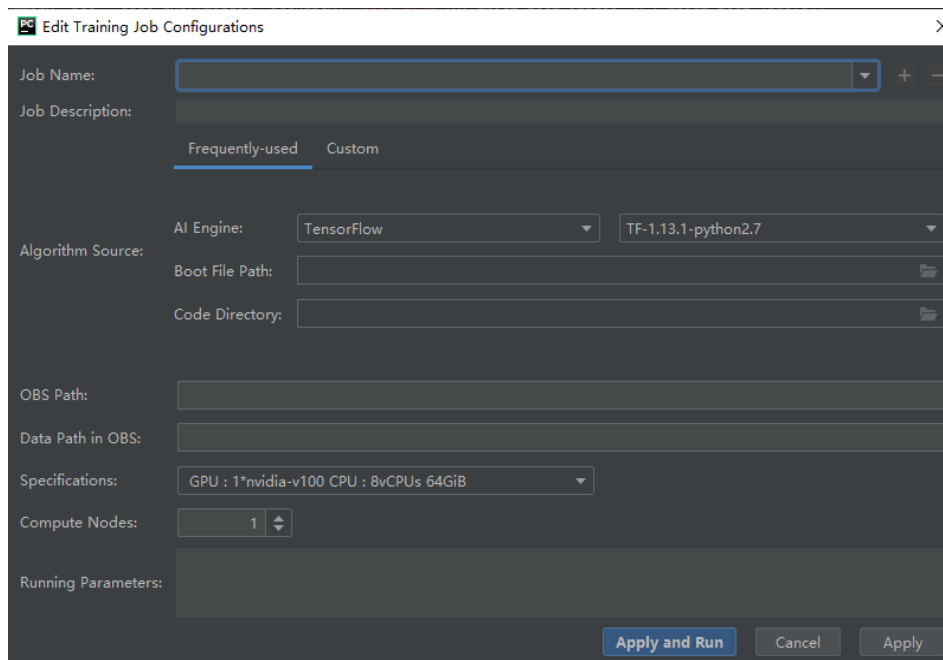
图 4-14 提交训练作业



当用户再有一个新的训练算法，如物体检测，此时想提交一个“Job Name”为“dog-cat-recognize”的训练作业，可执行如下步骤完成。

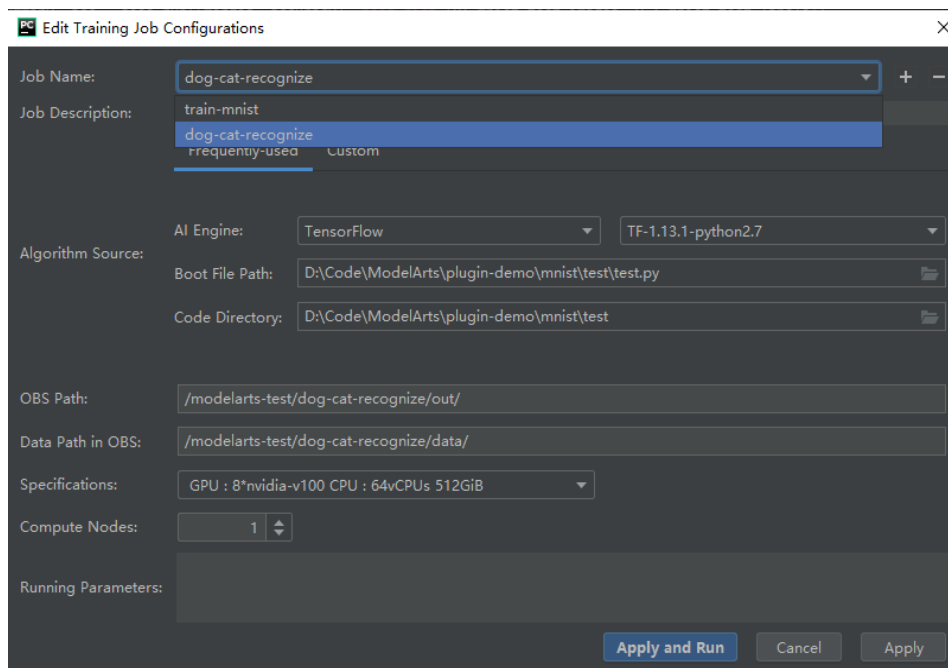
1. 单击“Edit Training Job Configuration”配置框右上角的加号，添加一个新的训练作业配置。填写完毕后单击“Apply and Run”提交训练作业即可。

图 4-15 添加一个新的训练作业配置



2. 当再次打开“Edit Training Job Configuration”配置界面时，可以看到“Job Name”的下拉菜单中显示了两个作业的名称，单击名称后可以查看作业的配置信息。

图 4-16 配置中包含 2 个作业名称



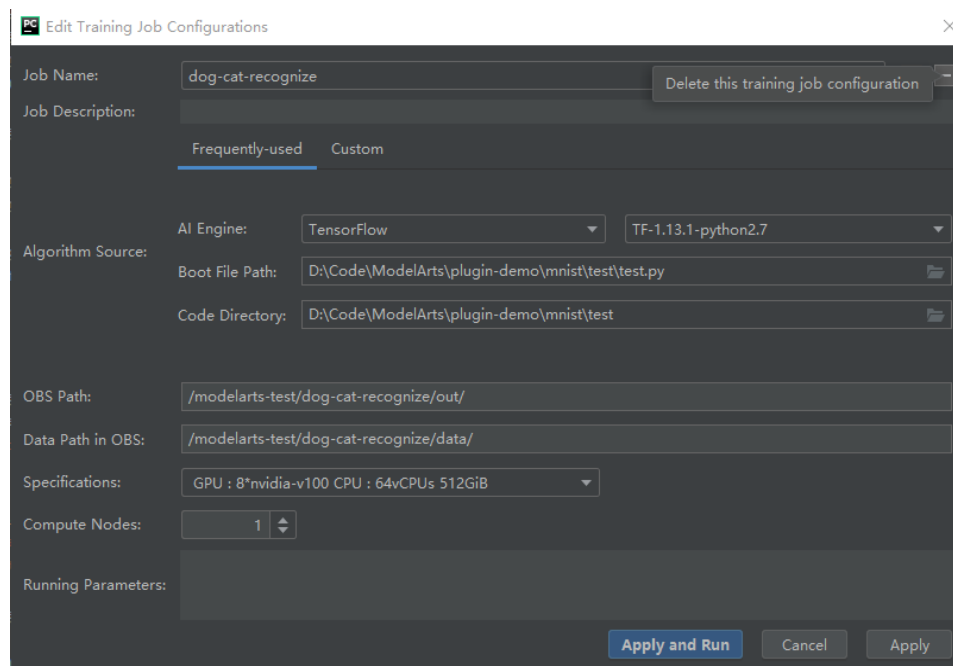
删除其中一个训练作业

当存在多个不同名称的训练作业时，如果不需要保留本地的训练作业配置信息时，可以单击右上角的减号，删除本地作业配置。

说明

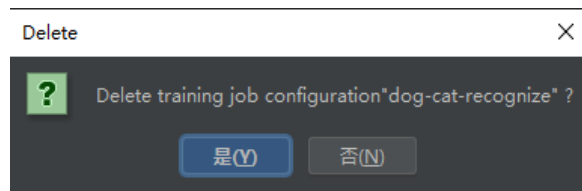
该操作是删除本地的配置信息，不会删除ModelArts管理控制台的作业。

图 4-17 删除训练作业配置信息



单击减号后，系统将弹出确认框，确认信息无误后，单击“是”即可删除本地保存的作业配置。

图 4-18 确认删除



说明

配置信息删除后无法恢复，请谨慎操作。

4.2 部署上线

针对训练好的模型，您也可以在ToolKit工具中一键部署为在线服务。

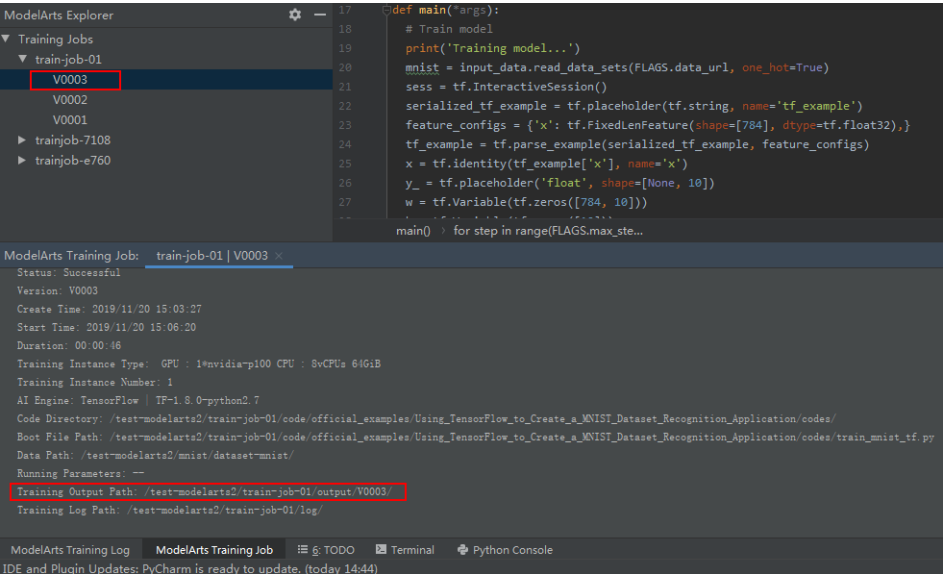
说明

当前仅支持部署为在线服务，不支持部署批量服务和边缘服务。

背景信息

- 已完成模型训练，且训练作业状态为成功。
- 针对ModelArts控制台已有的训练作业，也支持在PyCharm ToolKit工具中一键部署上线。
- 针对训练好的模型，在部署前，需完成推理脚本及配置文件的开发。推理脚本和配置文件的开发规范请参见[模型包规范](#)、[模型配置文件编写说明](#)、[模型推理代码编写说明](#)。
- 推理代码和配置文件需存储至模型输出位置。在PyCharm工具中运行的训练作业，其模型输出位置，可双击作业版本，在“ModelArts Training Job”区域中“Training Output Path”参数值即为此训练作业模型输出位置，此路径为OBS路径。

图 4-19 模型输出位置



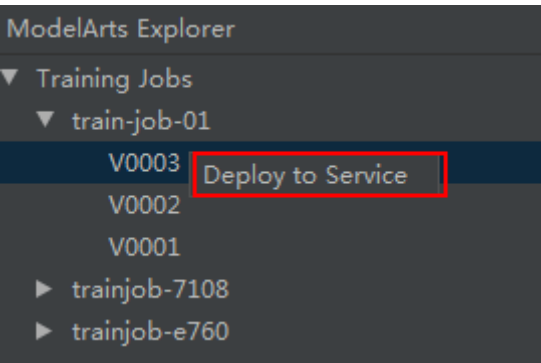
- 部署完成的服务，将呈现在ModelArts管理控制台的“部署上线 > 在线服务”列表中。您也可以在ModelArts管理控制台，管理已部署完成的在线服务，如预测、启停等操作。

部署模型

在完成模型训练，推理代码和配置文件编写，同时，已将推理代码和配置文件上传至对应的OBS路径下，即训练作业输出位置的model目录。您可以执行如下步骤，将模型部署为在线服务。

- 在ModelArts Explorer区域对应的训练作业版本中，右键选择“Deploy to Service”。

图 4-20 Deploy to Service



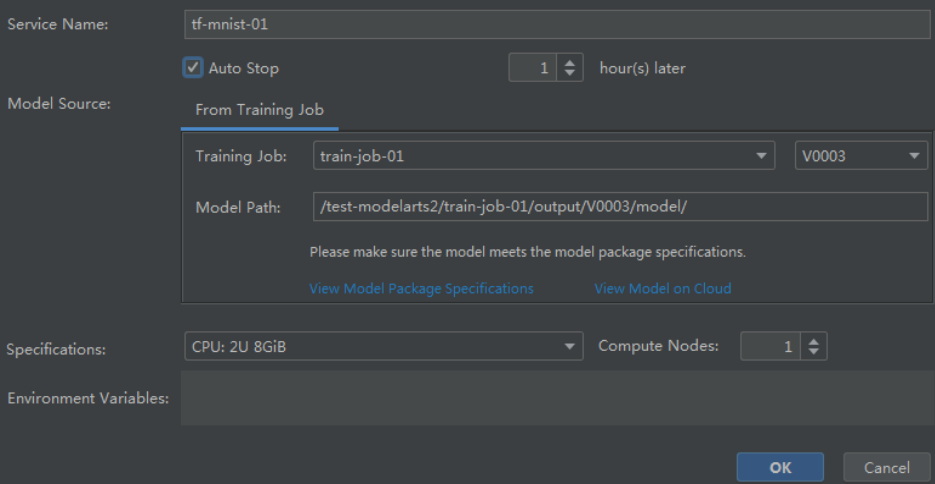
- 在弹出的对话框中，填写部署模型所需信息，详细参数说明请参见表4-2。

表 4-2 部署服务的参数说明

参数	说明
Service Name	在线服务的名称，请按照界面提示规则填写。

参数	说明
Auto Stop	启用该参数并设置时间后，服务将在指定时间后自动停止。 启用该功能后，默认设置为1小时后自动停止，您也可以根据实际情况设置自动停止的时间，如5小时之后。
Model Source	如果您是通过具体训练作业版本启动的部署任务，“Training Job”和“Model Path”将自动填写完成。您也可以根据实际情况，去更改配置。 <ul style="list-style-type: none">“Training Job”：表示训练作业的名称，以及其对应的版本。“Model Path”：表示训练输出模型的存储路径，为OBS路径。
Specifications	选择部署在线服务时使用的资源，目前支持“CPU 2核 8GiB”和“CPU：2核 8GiB GPU：1*P4”两种规格。
Compute Nodes	设置资源的实例个数。如果节点个数设置为1，表示后台的计算模式是单机模式；如果节点个数设置大于1，表示后台的计算模式为分布式的。请根据实际编码情况选择计算模式。
Environment Variables	设置环境变量，注入环境变量到容器实例。多个环境变量使用英文分号隔开。

图 4-21 设置部署服务的相关参数



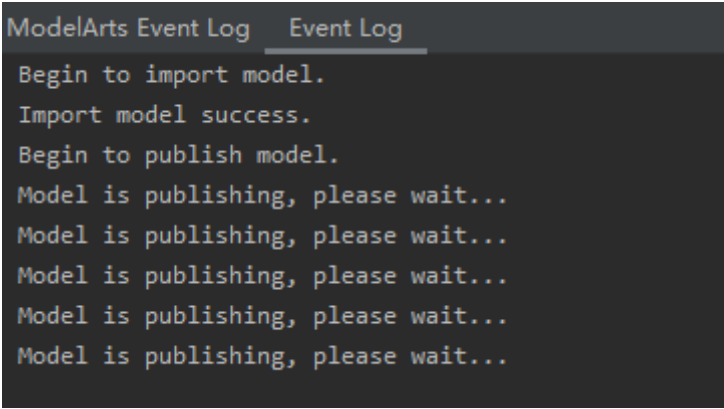
- 参数填写完成后，单击“OK”启动服务部署。部署任务启动后，工具左下角的“Event Log”实时展示部署状态。

部署为在线服务耗费时间较长，请耐心等待。部署完成后，将在界面中提供在线服务的链接，您可以单击链接快速跳转至ModelArts管理控制台的在线服务。注意，首次访问时需输入帐号密码登录。

说明

针对部署完成的服务，当前只支持在ModelArts管理控制台做预测操作。

图 4-22 部署状态



The screenshot shows a dark-themed window titled "ModelArts Event Log" with a sub-tab "Event Log". The log contains the following text:

```
Begin to import model.  
Import model success.  
Begin to publish model.  
Model is publishing, please wait...  
Model is publishing, please wait...  
Model is publishing, please wait...  
Model is publishing, please wait...  
Model is publishing, please wait...
```


5 OBS 上传下载

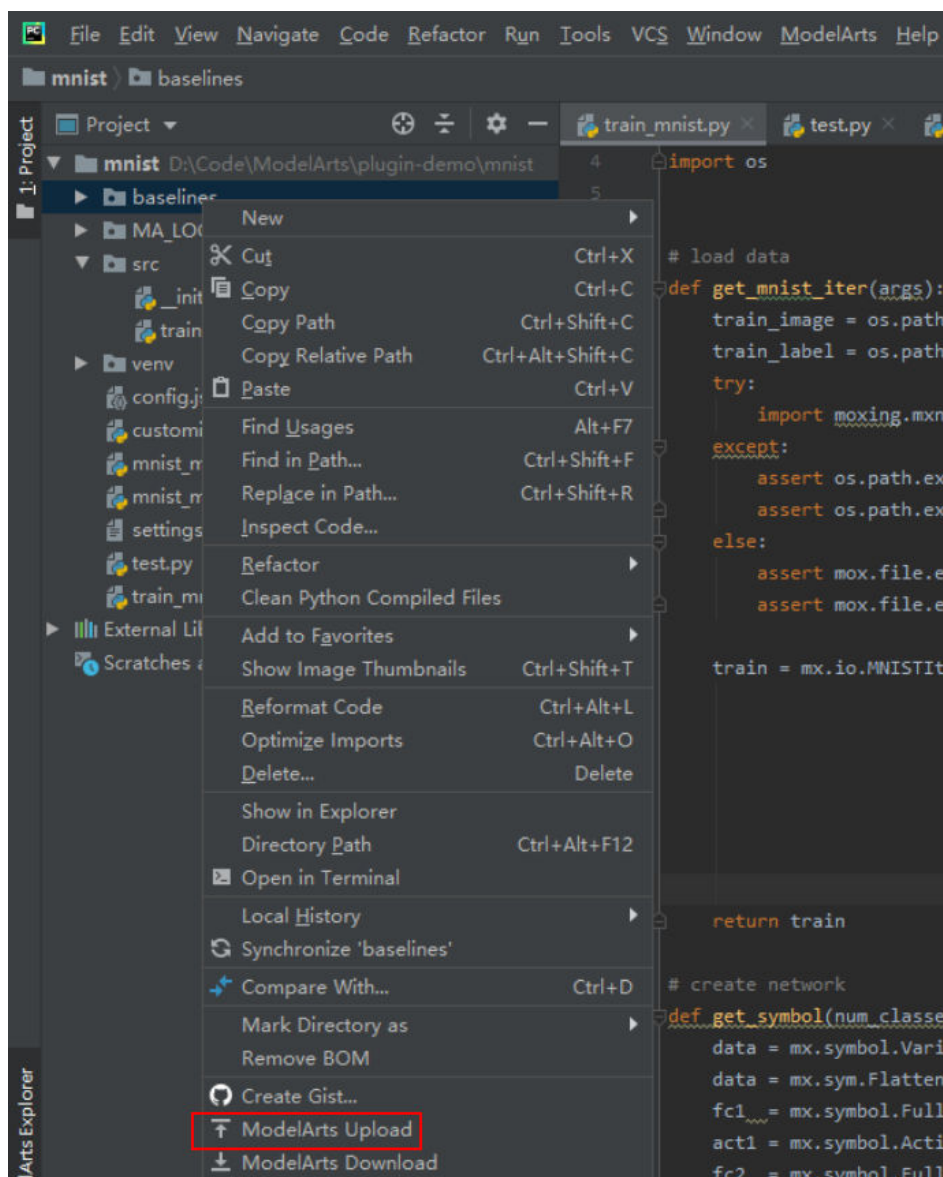
当用于训练的数据需要存储至OBS时，您可以使用PyCharm TookKit提供的OBS上传下载功能，在PyCharm界面将本地文件和文件夹上传到OBS，或者将OBS上的文件或文件夹下载到本地。

上传至OBS后，您的训练脚本中，可直接定义为对应的OBS路径。

将本地文件或文件夹上传至 OBS

1. 在Project窗口单击鼠标右键，在弹出的选项中选择“ModelArts Upload”。

图 5-1 ModelArts Upload

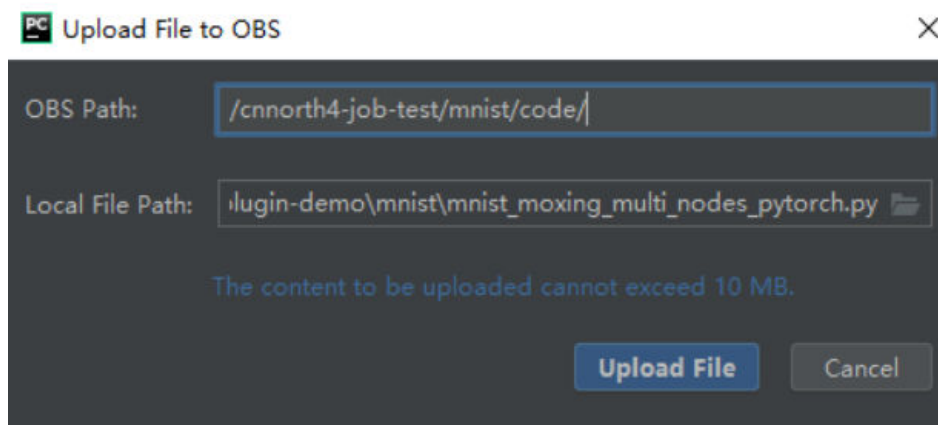


2. 在弹出的配置框中填写“OBS Path”和“Local File Path”，上传的文件或文件夹大小建议不超过10MB。配置完成后单击“Upload File”。

“OBS Path”：表示本地文件需要上传的OBS路径。

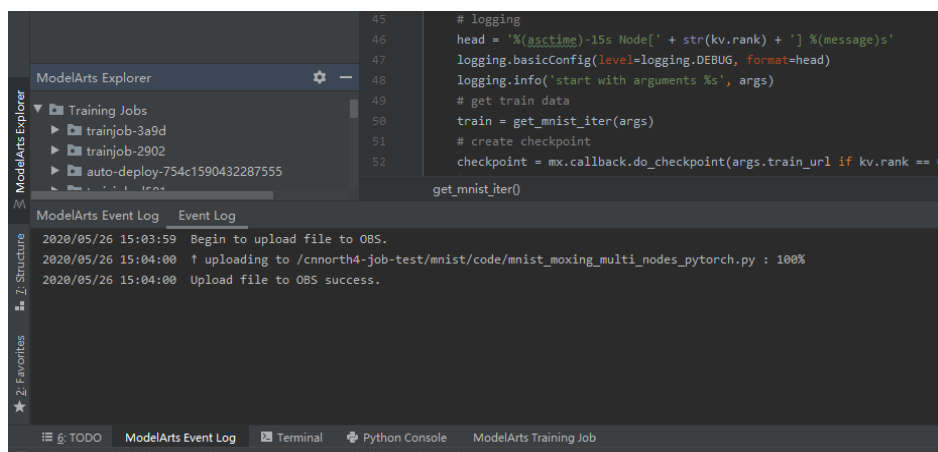
“Local File Path”：表示需上传的文件或文件夹所在的本地目录。如果上传文件，此参数请指定对应的具体文件。

图 5-2 填写文件路径



3. 在“ModelArts Event Log”窗口中可以看到上传的日志。

图 5-3 查看上传日志



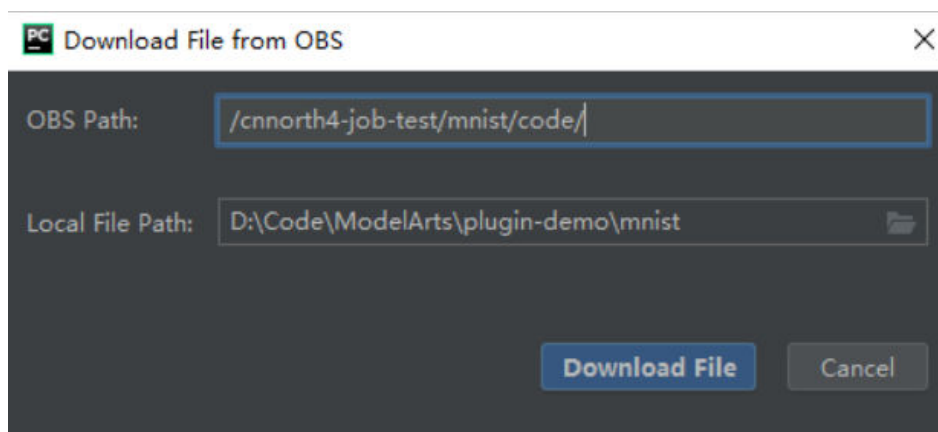
从 OBS 下载文件或文件夹

1. 在Project窗口单击鼠标右键，在弹出的选项中选择“ModelArts Download”。
2. 在弹出的配置框中填写“OBS Path”和“Local File Path”，配置完成后，单击“Download File”。

“OBS Path”：表示下载文件的OBS路径。当下载的是具体文件时，此处需指定为此文件的文件名。

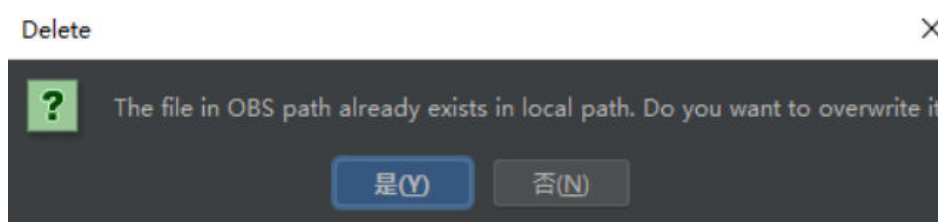
“Local File Path”：表示下载文件需要存储的本地路径。

图 5-4 从 OBS 下载文件



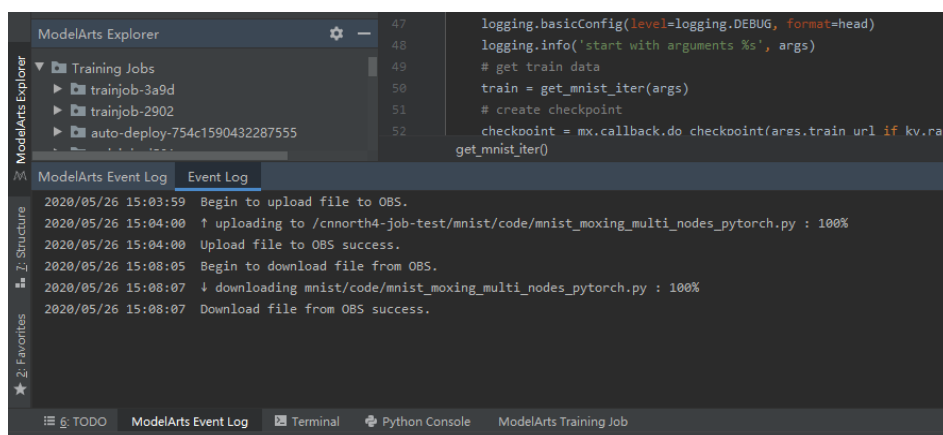
3. 当本地的目录下已经存在同名文件时，将提示是否选择覆盖，选择“是”表示覆盖，选择“否”表示有冲突的文件就跳过了不下载了，其他没有冲突的文件继续下载。请根据实际情况进行选择。

图 5-5 选择是否覆盖



4. 在“ModelArts Event Log”窗口中可以看到下载的日志。

图 5-6 查看下载日志



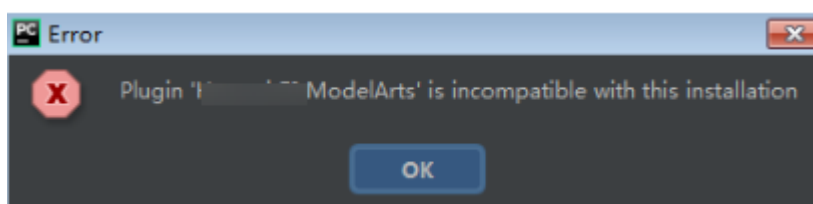
6 常见问题

6.1 安装 ToolKit 工具时出现错误，如何处理？

问题

在安装ToolKit工具过程中，出现如下错误。

图 6-1 错误提示



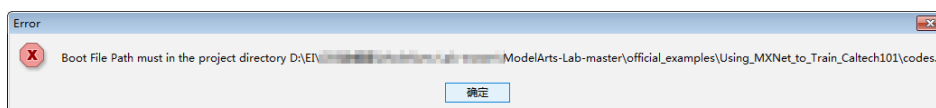
回答

此问题是因为插件版本和PyCharm版本不一致导致的，需要获取和PyCharm同一版本的插件安装，即2019.2或以上版本。

6.2 为什么无法启动训练？

如果启动脚本选择了不属于本工程的代码，则无法启动训练，错误信息如下图所示。建议将启动脚本添加至本工程，或者是打开启动脚本所在工程后，再启动训练作业。

图 6-2 错误信息



6.3 提交训练作业时，出现 xxx isn't existed in train version 错误

问题现象

提交训练作业时，出现xxx isn't existed in train version错误，如下所示。

图 6-3 xxx isn't existed in train_version 错误

```
Begin to check training configuration.
Begin to upload training code.
No modified code to upload.
Begin to get training job pre version.
Http Response Code : 400
{"is_success":false,"error_code":"ModelArts.0102","error_message":"193526 isn't existed in train_version."}
ModelArts Training is Finished.
```

原因分析

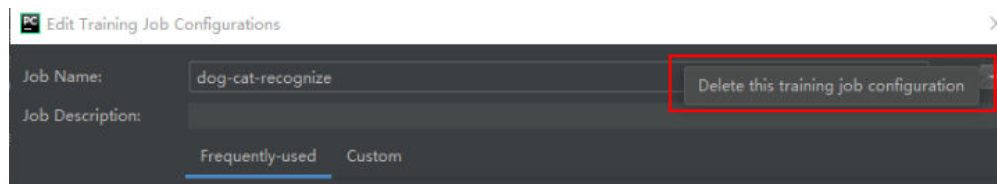
出现上述错误，是因为用户通过PyCharm ToolKit提交了训练作业后，又登录了ModelArts管理控制台删除了该训练作业导致的。

由于PyCharm ToolKit本地记录了云端ModelArts的训练作业ID，如果手动在ModelArts管理控制台删除，本地提交作业时候会报找不到该ID的作业。

解决措施

如果已经在ModelArts管理控制台删除了作业，需要在ToolKit也同步删除掉本地的配置信息，单击“Edit Training Configuration”，找到作业名称，单击右上角的减号并确认删除。

图 6-4 删除配置信息



在弹出的确认对话框中，确认信息无误后，单击“是”删除对应配置信息。删除后您可以创建新的训练作业配置并提交训练作业。

6.4 提交训练作业报错

当运行训练作业时，出现“Invalid OBS path”错误。

图 6-5 训练作业报错

```
No modified code to upload.
Begin to get training job pre version.
Begin to create training job.
Http Response Code : 400
{"is_success":false,"error_code":"ModelArts.0404","error_message":"Invalid OBS path '/{bucket}/{prefix}/code/{code}/'."}
Create training job failed.
ModelArts Training is Finished.
```

该问题可以通过以下方法排查：

- 如果是首次使用ModelArts的新用户，需要登录ModelArts管理控制台，在全局配置页面完成访问授权配置，推荐使用委托授权的方式。完成全局配置后，再次提交作业即可。
- 请排查所填写的Data Path in OBS是否存在，文件夹下是否有数据文件，如果没有，需要在OBS创建目录并上传训练数据到该目录。

6.5 部署上线时，出现错误

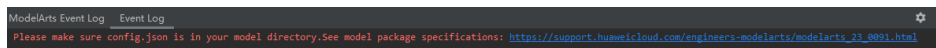
在部署上线前，您需要基于训练后的模型编写配置文件和推理代码。

如果您的模型存储路径下，缺少配置文件“`confi.json`”，或者缺少推理代码“`customize_service.py`”时，将出现错误，错误信息如下图所示。

解决方案：

请参考[模型包规范](#)写配置文件和推理代码，并存储至需部署的模型所在OBS目录下。

图 6-6 错误信息



6.6 如何查看 PyCharm ToolKit 的错误日志

PyCharm ToolKit的错误日志记录在PyCharm的“`idea.log`”中，以Windows为例，该文件的路径在“`C:\Users\xxx\IdeaIC2019.2\system\log\idea.log`”。

在日志中搜索“`modelarts`”，可以查看所有和PyCharm ToolKit相关的日志。

A 修订记录

发布日期	修订记录
2022-03-31	增加PyCharm ToolKit (latest版本) 工具包, 该版本支持对接新版训练作业。
2021-11-05	更新PyCharm ToolKit工具包版本为PyCharm-ToolKit-2.2.1。
2021-08-30	增加PyCharm版本说明。 前提条件 。
2021-07-08	更新PyCharm ToolKit工具下载地址。
2020-08-31	更新PyCharm ToolKit版本, 新增支持使用自定义镜像创建训练作业的方式。 提交训练作业 (旧版训练)
2020-06-05	<ul style="list-style-type: none">更新PyCharm ToolKit版本, 新增OBS上传下载、提交不同名称作业等功能。 提交不同名称的训练作业 OBS上传下载根据新版工具, 更新如下操作指导。 提交训练作业 (旧版训练) 查看训练作业详情 部署上线新增如下常见问题。 提交训练作业时, 出现xxx isn't existed in train_version错误 提交训练作业报错 如何查看PyCharm ToolKit的错误日志
2020-05-06	更新使用限制的描述。 PyCharm ToolKit介绍
2019-12-05	第一次正式发布。