

ModelArts

FAQs

Issue 01
Date 2023-12-06



Copyright © Huawei Technologies Co., Ltd. 2023. All rights reserved.

No part of this document may be reproduced or transmitted in any form or by any means without prior written consent of Huawei Technologies Co., Ltd.

Trademarks and Permissions



HUAWEI and other Huawei trademarks are trademarks of Huawei Technologies Co., Ltd.

All other trademarks and trade names mentioned in this document are the property of their respective holders.

Notice

The purchased products, services and features are stipulated by the contract made between Huawei and the customer. All or part of the products, services and features described in this document may not be within the purchase scope or the usage scope. Unless otherwise specified in the contract, all statements, information, and recommendations in this document are provided "AS IS" without warranties, guarantees or representations of any kind, either express or implied.

The information in this document is subject to change without notice. Every effort has been made in the preparation of this document to ensure accuracy of the contents, but all statements, information, and recommendations in this document do not constitute a warranty of any kind, express or implied.

Huawei Technologies Co., Ltd.

Address: Huawei Industrial Base
Bantian, Longgang
Shenzhen 518129
People's Republic of China

Website: <https://www.huawei.com>

Email: support@huawei.com

Security Declaration

Vulnerability

Huawei's regulations on product vulnerability management are subject to "Vul. Response Process". For details about the policy, see the following website:<https://www.huawei.com/en/psirt/vul-response-process>
For enterprise customers who need to obtain vulnerability information, visit:<https://securitybulletin.huawei.com/enterprise/en/security-advisory>

Contents

1 General Issues.....	1
1.1 What Is ModelArts?.....	1
1.2 What Are the Relationships Between ModelArts and Other Services?.....	1
1.3 What Are the Differences Between ModelArts and DLS?.....	2
1.4 How Do I Obtain an Access Key?.....	2
1.5 How Do I Upload Data to OBS?.....	2
1.6 What Do I Do If the System Displays a Message Indicating that the AK/SK Pair Is Unavailable?.....	3
1.7 How Do I Use ModelArts to Train Models Based on Structured Data?.....	3
1.8 What Are Regions and AZs?.....	3
1.9 How Do I Check Whether ModelArts and an OBS Bucket Are in the Same Region?.....	5
1.10 How Do I View All Files Stored in OBS on ModelArts?.....	5
1.11 Why Error: 403 Forbidden Is Displayed When I Perform Operations on OBS?.....	6
1.12 Where Are Datasets of ModelArts Stored in a Container?.....	7
1.13 What Are the Functions of ModelArts Training and Inference?.....	7
1.14 Can AI-assisted Identification of ModelArts Identify a Specific Label?.....	7
2 Data Management.....	8
2.1 Are There Size Limits for Images to be Uploaded?.....	8
2.2 What Do I Do If Images in a Dataset Cannot Be Displayed?.....	8
2.3 How Do I Integrate Multiple Object Detection Datasets into One Dataset?.....	9
2.4 What Do I Do If Importing a Dataset Failed?.....	9
2.5 Can a Table Dataset Be Labeled?.....	10
2.6 Why Does Data Fail to Be Imported Using the Manifest File?.....	10
2.7 Where Are Labeling Results Stored?.....	11
2.8 How Do I Download Labeling Results to a Local PC?.....	12
2.9 Why Cannot Team Members Receive Emails for a Team Labeling Task?.....	12
2.10 Can I Add Multiple Labeling Boxes to an Object Detection Dataset Image?.....	13
2.11 How Do I Merge Two Datasets?.....	13
2.12 Why Are Images Displayed in Different Angles Under the Same Account?.....	13
2.13 Do I Need to Train Data Again If New Data Is Added After Auto Labeling Is Complete?.....	13
2.14 Why Does the System Display a Message Indicating My Label Fails to Save on ModelArts?.....	14
2.15 Can One Label Be Identified Among Multiple Labels?.....	14
2.16 Why Are Newly Added Images Not Automatically Labeled After Data Amplification Is Enabled?.....	14
2.17 Why Cannot Videos in a Video Dataset Be Displayed or Played?.....	15

2.18 How Do I Add Images to a Validation or Training Dataset?.....	15
2.19 Can I Customize Labels for an Object Detection Dataset?.....	15
2.20 What ModelArts Data Management Can Be Used for?.....	15
2.21 Why Is My Newly Created Bucket Unavailable?.....	17
2.22 Will My Old-Version Datasets Be Cleared After the Old Version Is Discontinued? The existing datasets and the ones newly created in the old version will be retained after the old version is discontinued.....	17
2.23 Why Is My New Dataset Version Unavailable in Versions?.....	17
2.24 How Do I View the Size of a Dataset?.....	18
2.25 How Do I View Labeling Details of a New Dataset?.....	18
2.26 How Do I Export Labeled Data?.....	19
2.27 Why Cannot I Find My Newly Created Dataset?.....	19
2.28 What Do I Do If the Database Quota Is Incorrect?.....	19
2.29 How Do I Split a Dataset?.....	19
2.30 How Do I Delete a Dataset Image?.....	20
2.31 Why Is There No Sample in the ModelArts Dataset Downloaded from AI Gallery and Then an OBS Bucket?.....	20
3 Notebook.....	23
3.1 Constraints.....	23
3.1.1 Is sudo Privilege Escalation Supported?.....	23
3.1.2 Does ModelArts Support apt-get?.....	23
3.1.3 Is the Keras Engine Supported?.....	23
3.1.4 Does ModelArts Support the Caffe Engine?.....	24
3.1.5 Can I Install MoXing in a Local Environment?.....	24
3.1.6 Can Notebook Instances Be Remotely Logged In?.....	24
3.2 Data Upload or Download.....	24
3.2.1 How Do I Upload a File from a Notebook Instance to OBS or Download a File from OBS to a Notebook Instance?.....	24
3.2.2 How Do I Upload Local Files to a Notebook Instance?.....	26
3.2.3 How Do I Import Large Files to a Notebook Instance?.....	26
3.2.4 Where Will the Data Be Uploaded to?.....	26
3.2.5 How Do I Download Files from a Notebook Instance to a Local Computer?.....	27
3.3 Data Storage.....	27
3.3.1 When Creating a Notebook Instance, What Is the Difference If You Select or Do Not Select a Storage Path? Which Directories Can Store Big Data?.....	27
3.3.2 Where Is Data Stored After the Sync OBS Function Is Used?.....	28
3.3.3 How Do I Rename an OBS File?.....	28
3.3.4 Do Files in /cache Still Exist After a Notebook Instance is Stopped or Restarted? How Do I Avoid a Restart?.....	28
3.3.5 How Do I Use the pandas Library to Process Data in OBS Buckets?.....	28
3.4 Environment Configurations.....	28
3.4.1 How Do I Enable the Terminal Function in DevEnviron of ModelArts?.....	28
3.4.2 How Do I Install External Libraries in a Notebook Instance?.....	29
3.4.3 How Do I Create a Soft Link and Switch a CUDA Version?.....	30

3.4.4 Should I Access the Python Environment Same as the Notebook Kernel of the Current Instance in the Terminal?.....	31
3.4.5 How Do I Import a Python Library to a Notebook Instance to Resolve the ModuleNotFoundError Error?.....	31
3.4.6 How Can I Resolve Abnormal Font Display on a ModelArts Notebook Accessed from iOS?.....	32
3.5 Notebook Instances.....	34
3.5.1 What Do I Do If I Cannot Open a Notebook Instance I Created?.....	34
3.5.2 What Should I Do When the System Displays an Error Message Indicating that No Space Left After I Run the pip install Command?.....	34
3.5.3 What Do I Do If "Read timed out" Is Displayed After I Run pip install?.....	34
3.5.4 What Do I Do If the Code Can Be Run But Cannot Be Saved, and the Error Message "save error" Is Displayed?.....	35
3.5.5 What Do I Do If "Error loading notebook" Is Displayed When I Create an IPYNB File?.....	35
3.6 Code Execution.....	36
3.6.1 What Do I Do If a Notebook Instance Won't Run My Code?.....	36
3.6.2 Why Does the Instance Break Down When dead kernel Is Displayed During Training Code Running?	37
3.6.3 What Do I Do If cudaCheckError Occurs During Training?.....	37
3.6.4 What Should I Do If DevEnviron Prompts Insufficient Space?.....	38
3.6.5 Why Does the Notebook Instance Break Down When opencv.imshow Is Used?.....	38
3.6.6 Why Cannot the Path of a Text File Generated in Windows OS Be Found In a Notebook Instance?	38
3.6.7 What Do I Do If Files Fail to Be Saved in JupyterLab?.....	39
3.7 VS Code.....	39
3.7.1 What Do I Do If Installing a Remote Plug-in Failed?.....	39
3.7.2 What Do I Do If a Restarted Notebook Instance Can Be Connected Only After I Locally Delete known_hosts?.....	40
3.7.3 What Do I Do If the Source Code Cannot Be Accessed When I Use VS Code for Debugging?.....	41
3.7.4 What Do I Do If a Message Is Displayed Indicating an Incorrect Username or Email Address When I Use VS Code to Submit Code?.....	42
3.7.5 How Can I View Remote Logs in VS Code?.....	42
3.7.6 How Can I Open the VS Code Configuration File settings.json?.....	42
3.7.7 How Can I Change the VS Code Background Color to Light Green?.....	43
3.7.8 How Can I Set the Default Remote Plug-in in VS Code?.....	43
3.7.9 How Can I Install a Local Plug-in on the Remote End or a Remote Plug-in on the Local End Through VS Code?.....	43
3.8 Failures to Access the Development Environment Through VS Code.....	43
3.8.1 When Do I Do If the VS Code Window Is Not Displayed?.....	43
3.8.2 What Do I Do If a Remote Connection Failed After VS Code Is Opened?.....	44
3.8.3 What Do I Do If Error Message "Could not establish connection to xxx" Is Displayed During a Remote Connection?.....	47
3.8.4 What Do I Do If the Connection to a Remote Development Environment Remains in "Setting up SSH Host xxx: Downloading VS Code Server locally" State for More Than 10 Minutes?.....	48
3.8.5 What Do I Do If the Connection to a Remote Development Environment Remains in the State of "Setting up SSH Host xxx: Downloading VS Code Server locally" for More Than 10 Minutes?.....	50

3.8.6 What Do I Do If the Connection to a Remote Development Environment Remains in the State of "ModelArts Remote Connect: Connecting to instance xxx..." for More Than 10 Minutes?.....	50
3.8.7 What Do I Do If a Remote Connection Is in the Retry State?.....	51
3.8.8 What Do I Do If Error Message "The VS Code Server failed to start" Is Displayed?.....	52
3.8.9 What Do I Do If Error Message "Permissions for 'x:/xxx.pem' are too open" Is Displayed?.....	54
3.8.10 What Do I Do If Error Message "ssh: connect to host xxx.pem port xxx: Connection refused" Is Displayed?.....	55
3.8.11 What Do I Do If Error Message "ssh: connect to host ModelArts-xxx port xxx: Connection timed out" Is Displayed?.....	55
3.8.12 What Do I Do If Error Message "Load key "C:/Users/xx/test1/xxx.pem": invalid format" Is Displayed?.....	56
3.8.13 What Do I Do If Error Message "An SSH installation couldn't be found" or "Could not establish connection to instance xxx: 'ssh' ..." Is Displayed?.....	57
3.8.14 What Do I Do If Error Message "no such identity: C:/Users/xx /test.pem: No such file or directory" Is Displayed?.....	59
3.8.15 What Do I Do If Error Message "Host key verification failed" or "Port forwarding is disabled" Is Displayed?.....	59
3.8.16 What Do I Do If Error Message "Failed to install the VS Code Server" or "tar: Error is not recoverable: exiting now" Is Displayed?.....	61
3.8.17 What Do I Do If Error Message "XHR failed" Is Displayed When a Remote Notebook Instance Is Accessed Through VS Code?.....	62
3.8.18 What Do I Do If It Takes a Long Time to Set Up a Remote Connection After VS Code Is Automatically Upgraded?.....	62
3.9 Others.....	63
3.9.1 How Do I Use Training Code from a Training Jobs After Debugging the Code in a Notebook Instance?.....	64
3.9.2 Why Is the Training Speed Similar When Different Notebook Flavors Are Used?.....	64
3.9.3 How Do I Perform Incremental Training When Using MoXing?.....	64
3.9.4 How Do I View GPU Usage on the Notebook?.....	67
3.9.5 How Can I Obtain GPU Usage Through Code?.....	67
3.9.6 Which Real-Time Performance Indicators of an Ascend Chip Can I View?.....	70
3.9.7 What Are the Relationships Between Files Stored in JupyterLab, Terminal, and OBS?.....	70
3.9.8 How Do I Migrate Data from an Old-Version Notebook Instance to a New-Version One?.....	70
3.9.9 How Do I Use the Datasets Created on ModelArts in a Notebook Instance?.....	73
3.9.10 pip and Common Commands.....	73
3.9.11 What Are Sizes of the /cache Directories for Different Notebook Specifications in DevEnviron?.....	74
4 Training Jobs.....	75
4.1 Functional Consulting.....	75
4.1.1 What Are the Solutions to Underfitting?.....	75
4.1.2 How Do I Obtain a Trained ModelArts Model?.....	76
4.1.3 How Do I Set the Runtime Environment of the AI Engine Scikit_Learn 0.18.1?.....	76
4.2 Reading Data During Training.....	76
4.2.1 How Do I Configure the Input and Output Data for Training Models on ModelArts?.....	76
4.2.2 How Do I Improve Training Efficiency While Reducing Interaction with OBS?.....	77

4.2.3 Why the Data Read Efficiency Is Low When a Large Number of Data Files Are Read During Training?.....	78
4.3 Compiling the Training Code.....	79
4.3.1 How Do I Create a Training Job When a Dependency Package Is Referenced by the Model to Be Trained?.....	79
4.3.2 What Is the Common File Path for Training Jobs?.....	81
4.3.3 How Do I Install a Library That C++ Depends on?.....	81
4.3.4 How Do I Check Whether a Folder Copy Is Complete During Job Training?.....	82
4.3.5 How Do I Load Some Well Trained Parameters During Job Training?.....	82
4.3.6 How Do I Obtain Training Job Parameters from the Boot File of the Training Job?.....	82
4.3.7 Why Can't I Use os.system ('cd xxx') to Access the Corresponding Folder During Job Training?.....	84
4.3.8 How Do I Invoke a Shell Script in a Training Job to Execute the .sh File?.....	84
4.3.9 How Do I Obtain the Path for Storing the Dependency File in Training Code?.....	84
4.4 Creating a Training Job.....	85
4.4.1 What Can I Do If the Message "Object directory size/quantity exceeds the limit" Is Displayed When I Create a Training Job?.....	85
4.4.2 What Are Precautions for Setting Training Parameters?.....	85
4.4.3 What Are Sizes of the /cache Directories for Different Resource Specifications in the Training Environment?.....	86
4.4.4 Is the /cache Directory of a Training Job Secure?.....	87
4.4.5 Why Is a Training Job Always Queuing?.....	87
4.5 Managing Training Job Versions.....	87
4.5.1 Does a Training Job Support Scheduled or Periodic Calling?.....	87
4.6 Viewing Job Details.....	87
4.6.1 How Do I Check Resource Usage of a Training Job?.....	87
4.6.2 How Do I Access the Background of a Training Job?.....	88
4.6.3 Is There Any Conflict When Models of Two Training Jobs Are Saved in the Same Directory of a Container?.....	88
4.6.4 Only Three Valid Digits Are Retained in a Training Output Log. Can the Value of loss Be Changed?.....	88
4.6.5 Can a Trained Model Be Downloaded or Migrated to Another Account? How Do I Obtain the Download Path?.....	88
5 Model Management.....	90
5.1 Importing Models.....	90
5.1.1 How Do I Import the .h5 Model of Keras to ModelArts?.....	90
6 Service Deployment.....	91
6.1 Functional Consulting.....	91
6.1.1 What Types of Services Can Models Be Deployed as on ModelArts?.....	91
6.1.2 What Are the Differences Between Real-Time Services and Batch Services?.....	91
6.1.3 How Do I Select Compute Node Specifications for Deploying a Service?.....	91
6.1.4 What Is the CUDA Version for Deploying a Service on GPUs?.....	92
6.2 Real-Time Services.....	92
6.2.1 What Do I Do If a Conflict Occurs in the Python Dependency Package of a Custom Prediction Script When I Deploy a Real-Time Service?.....	93

6.2.2 What Is the Format of a Real-Time Service API?.....	93
6.2.3 What Do I Do If an Image Fails to Be Pulled When a Real-Time Service Is Deployed, Started, Upgraded, or Modified?.....	93
6.2.4 What Do I Do If an Image Restarts Repeatedly When a Real-Time Service Is Deployed, Started, Upgraded, or Modified?.....	94
6.2.5 What Do I Do If a Container Health Check Failed When a Real-Time Service Is Deployed, Started, Upgraded, or Modified?.....	94
6.2.6 What Do I Do If Resources Are Insufficient When a Real-Time Service Is Deployed, Started, Upgraded, or Modified?.....	94
7 API/SDK.....	95
7.1 Can ModelArts APIs or SDKs Be Used to Download Models to a Local PC?.....	95
7.2 What Installation Environments Do ModelArts SDKs Support?.....	95
7.3 Does ModelArts Use the OBS API to Access OBS Files over an Intranet or the Internet?.....	95

1 General Issues

1.1 What Is ModelArts?

ModelArts is a one-stop AI development platform geared toward developers and data scientists of all skill levels. It enables you to rapidly build, train, and deploy models anywhere (from the cloud to the edge), and manage full-lifecycle AI workflows. ModelArts accelerates AI development and fosters AI innovation with key capabilities, including data preprocessing and auto labeling, distributed training, automated model building, and one-click workflow executing.

The one-stop ModelArts platform covers all stages of AI development, including data processing, AI application creation, and model training and deployment. The underlying layer of ModelArts supports various heterogeneous computing resources. You can flexibly select and use the resources without having to consider the underlying technologies. In addition, ModelArts supports popular open-source AI development frameworks such as TensorFlow. Developers can also use self-developed algorithm frameworks to match their usage habits.

ModelArts aims to achieve simple, convenient AI development.

1.2 What Are the Relationships Between ModelArts and Other Services?

OBS

ModelArts uses Object Storage Service (OBS) to securely and reliably store data and models at low costs. For more details, see *Object Storage Service Console Operation Guide*.

CCE

ModelArts uses Cloud Container Engine (CCE) to deploy models as real-time services. CCE enables high concurrency and provides elastic scaling. For more information about CCE, see *Cloud Container Engine User Guide*.

SWR

To use an AI framework that is not supported by ModelArts, use SoftWare Repository for Container (SWR) to customize an image and import the image to ModelArts for training or inference. For details about SWR, see .

Cloud Eye

ModelArts uses Cloud Eye to monitor online services and model loads in real time and send alarms and notifications automatically. For details about Cloud Eye, see *Cloud Eye User Guide*.

CTS

ModelArts uses Cloud Trace Service (CTS) to record operations for later query, audit, and backtrack operations. For details about CTS, see *Cloud Trace Service User Guide*.

1.3 What Are the Differences Between ModelArts and DLS?

Deep Learning Service (DLS) is a one-stop deep learning platform based on the high-performance computing capabilities . With various optimized neural network models, DLS allows you to easily implement model training and evaluation with the flexibility of on-demand scheduling.

However, DLS supports only the deep learning technologies, while ModelArts integrates both the deep learning and machine learning technologies. In addition, ModelArts is a one-stop AI development platform, which manages the AI development lifecycle from data labeling, algorithm development, to model training and deployment. To be specific, ModelArts contains and supports the functions and features of DLS. The functions related to deep learning can be directly used in ModelArts. If you are a DLS user, you can also migrate the data in DLS to ModelArts.

1.4 How Do I Obtain an Access Key?

Obtaining an Access Key

1. Log in to the console, enter the **My Credentials** page, and choose **Access Keys > Create Access Key**.
2. In the **Create Access Key** dialog box that is displayed, use the login password for verification.
3. Click **OK**, open the **credentials.csv** file, and save the key file as prompted. The access key file is saved in the default downloads folder of the browser. Then, the access key (**Access Key Id** and **Secret Access Key**) is obtained.

1.5 How Do I Upload Data to OBS?

Before using ModelArts to develop AI models, data needs to be uploaded to an OBS bucket. You can log in to the OBS console to create an OBS bucket, create a

folder in it, and upload data. For details about how to upload data, see *Object Storage Service User Guide*.

1.6 What Do I Do If the System Displays a Message Indicating that the AK/SK Pair Is Unavailable?

Issue Analysis

An AK and SK form a key pair required for accessing OBS. Each SK corresponds to a specific AK, and each AK corresponds to a specific user. If the system displays a message indicating that the AK/SK pair is unavailable, it is possible that the account is in arrears or the AK/SK pair is incorrect.

Solution

1. Use the current account to log in to the OBS console and check whether the current account can access OBS.
 - If the account can access OBS, rectify the fault by referring to [2](#).
 - If the account cannot access OBS, contact technical support.
2.
 - If yes, .
 - If not, replace the AK/SK with those created using the current account. For details, see [Access Keys](#)"Obtaining an Access Key".

1.7 How Do I Use ModelArts to Train Models Based on Structured Data?

For more advanced users, ModelArts provides the notebook creation function of DevEnviron for code development. It allows the users to create training tasks with large volumes of data in training jobs and use the engines such as Scikit_Learn, XGBoost, or Spark_MLlib in the development and training processes.

1.8 What Are Regions and AZs?

Concept

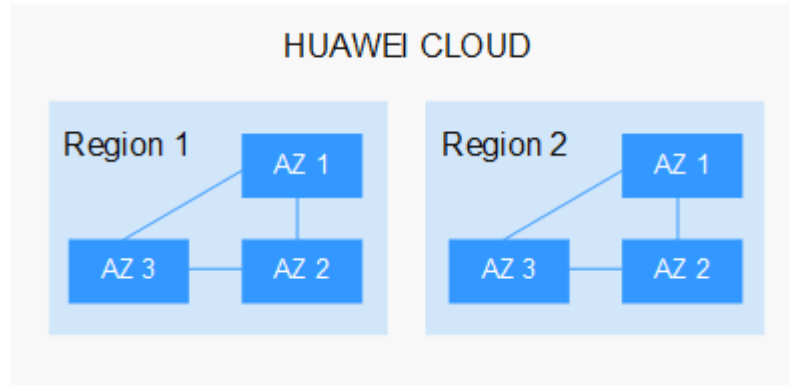
A region and availability zone (AZ) identify the location of a data center. You can create resources in a specific region and AZ.

- Regions are divided based on geographical location and network latency. Public services, such as Elastic Cloud Server (ECS), Elastic Volume Service (EVS), Object Storage Service (OBS), Virtual Private Cloud (VPC), Elastic IP (EIP), and Image Management Service (IMS), are shared within the same region. Regions are classified into universal regions and dedicated regions. A universal region provides universal cloud services for common tenants. A dedicated region provides specific services for specific tenants.
- An AZ contains one or more physical data centers. Each AZ has independent cooling, fire extinguishing, moisture-proof, and electricity facilities. Within an

AZ, computing, network, storage, and other resources are logically divided into multiple clusters. AZs within a region are interconnected using high-speed optical fibers to support cross-AZ high-availability systems.

Figure 1-1 shows the relationship between regions and AZs.

Figure 1-1 Regions and AZs



HUAWEI CLOUD provides services in many regions around the world. Select a region and AZ based on requirements. For more information, see .

Selecting a Region

When selecting a region, consider the following factors:

- Location
 - It is recommended that you select the closest region for low network latency and quick access. Regions within the Chinese mainland provide the same infrastructure, BGP network quality, as well as resource operations and configurations. Therefore, if your target users are on the Chinese mainland, you do not need to consider the network latency differences when selecting a region.
 - If your target users are in Asia Pacific (excluding the Chinese mainland), select the **CN-Hong Kong**, **AP-Bangkok**, or **AP-Singapore** region.
 - If your target users are in Africa, select the **AF-Johannesburg** region.
 - If your target users are in Europe, select the **EU-Paris** region.
 - If your target users are in Latin America, select the **LA-Santiago** region.

NOTE

The **LA-Santiago** region is located in Chile.

- Resource price
 - Resource prices may vary in different regions. For details, see .

Selecting an AZ

When deploying resources, consider your applications' requirements on disaster recovery (DR) and network latency.

- For high DR capability, deploy resources in different AZs within the same region.

- For lower network latency, deploy resources in the same AZ.

Regions and Endpoints

Before you use an API to call resources, specify its region and endpoint. For more details, see .

1.9 How Do I Check Whether ModelArts and an OBS Bucket Are in the Same Region?

If an OBS directory needs to be specified for using ModelArts functions, such as creating training jobs and datasets, ensure that the OBS bucket and ModelArts are in the same region.

Checking Whether the OBS Bucket and ModelArts Are in the Same Region

1. Check the region where the created OBS bucket resides.
 - a. Log in to OBS Console.
 - b. On the **Object Storage Service** page, to search for a bucket, enter a keyword in **Bucket Name**.
In the **Region** column, view the region where the created OBS bucket is located.
2. Check the region where ModelArts is deployed.
Log in to the ModelArts console and view the region where ModelArts resides in the upper left corner.
3. Check whether the region of the created OBS bucket is the same as that of ModelArts. Ensure that the OBS bucket is in the same region as ModelArts.

1.10 How Do I View All Files Stored in OBS on ModelArts?

To view all files stored in OBS when using notebook instances or training jobs, use either of the following methods:

- OBS console
Log in to OBS console using the current account, and search for the OBS buckets, folders, and files.
- You can use an API to check whether a given directory exists. In an existing notebook instance or after creating a new notebook instance, run the following command to check whether the directory exists:

```
import moxing as mox
mox.file.list_directory('obs://bucket_name', recursive=True)
```

If there are a large number of files, wait until the final file path is displayed.

1.11 Why Error: 403 Forbidden Is Displayed When I Perform Operations on OBS?

Symptom

Message "Error: stat:403" is displayed when I use `mox.file.copy_parallel` in ModelArts to perform operations on OBS.

Figure 1-2 Error message

```
ERROR:root:
  stat:403
  errorCode:None
  errorMessage:None
  reason:Forbidden
  request-id:000001752610DE67600F295F15304A6C
  retry:0
```

Possible Causes

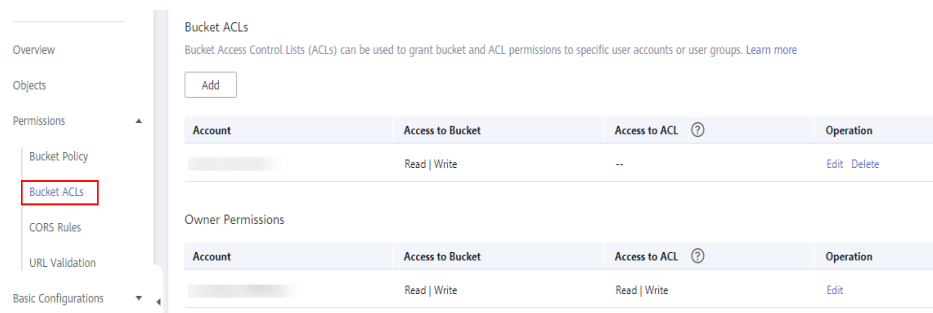
- ModelArts uses an AK/SK for authentication globally, and the AK/SK has been deleted and recreated.
- You do not have the permission to access OBS buckets.
- The OBS bucket policy is incorrect.

Solution

If you access the OBS bucket using an IAM user account, contact the tenant account to perform the following operations:

- For cause 1, go to the global configuration page and reconfigure the authorization.
- For cause 2, log in to the OBS console, search for the target OBS bucket, and click the bucket name to go to the **Overview** page. In the left navigation pane, choose **Permissions > Bucket ACLs**. On the **Bucket ACLs** page that is displayed, check whether the current account has the read and write permissions. If it does not, contact the bucket owner to grant the permissions.

Figure 1-3 Bucket ACLs



- For cause 2, log in to the OBS console, search for the target OBS bucket, and click the bucket name to go to the **Overview** page. In the navigation pane,

choose **Permissions > Bucket Policy** and verify that the current OBS bucket can be accessed by IAM users.

If the fault persists, see for further troubleshooting.

1.12 Where Are Datasets of ModelArts Stored in a Container?

Datasets of ModelArts and data in specific data storage locations are stored in OBS.

1.13 What Are the Functions of ModelArts Training and Inference?

ModelArts training includes ExeML, training management, and dedicated resource pools (for development/training).

ModelArts inference includes AI application management and deployment.

1.14 Can AI-assisted Identification of ModelArts Identify a Specific Label?

After a model with multiple labels is trained and deployed as a real-time service, all the labels are identified. If only one type of label needs to be identified, train a model dedicated for identifying the label. To speed up the label identification, select a high flavor for deploying the model.

2 Data Management

2.1 Are There Size Limits for Images to be Uploaded?

For data management, there are limits on the image size when you upload images to the datasets whose labeling type is object detection or image classification. The size of an image cannot exceed 8 MB, and only JPG, JPEG, PNG, and BMP formats are supported.

Note that for ExeML, the size of an image to be uploaded cannot exceed 5 MB.

Solutions:

- Import the images from OBS. Upload images to any OBS directory and import the images from the OBS directory to an existing dataset.
- Use data source synchronization. Upload images to the input directory or its subdirectory of a dataset, and click **Synchronize Data Source** on the dataset details page to add new images. Note that synchronizing a data source will delete the files deleted from OBS from the dataset. Exercise caution when performing this operation.
- Create a dataset. Upload images to any OBS directory. You can directly use the image directory as the input directory to create the dataset.

2.2 What Do I Do If Images in a Dataset Cannot Be Displayed?

Symptom

Images in a created dataset cannot be displayed during labeling, and they cannot be viewed by clicking them. Alternatively, the system displays a message indicating that an error occurred in image loading.

Possible Cause

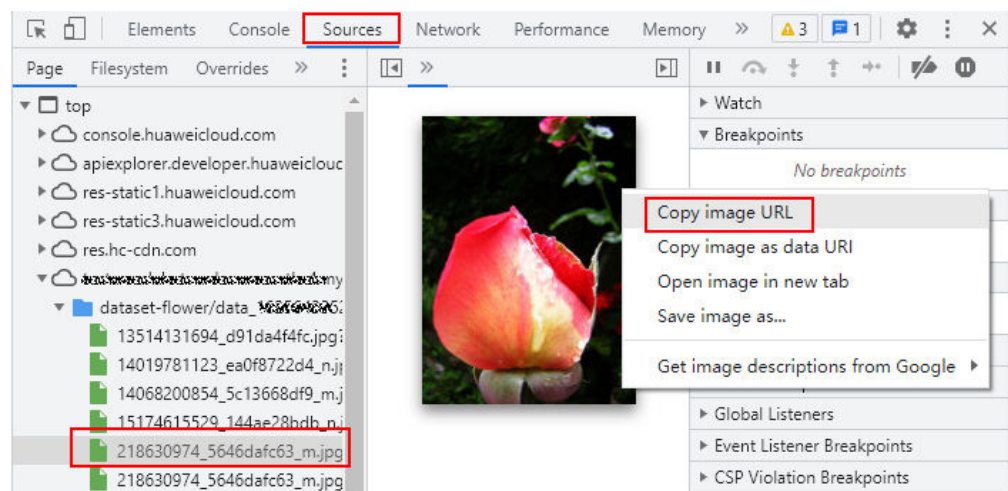
- The local network may be faulty. As a result, OBS cannot be accessed and images cannot be loaded.

- You are not allowed to access the target OBS bucket.
- The OBS bucket or file may be encrypted.
- The OBS storage class does not allow the parallel file system to process images. Therefore, the thumbnails cannot be displayed.

Solution

1. The following uses Google Chrome as an example. Press **F12** to open the browser console, locate the image, and copy the image URL.

Figure 2-1 Obtaining the image URL



2. Enter the URL in a new browser. The "Your Connection Is Not Private" message is displayed. Click **Advanced** on the page and choose **Proceed to <link> (unsafe)** to go to the target website.
3. After the image is successfully accessed, return to the ModelArts console to access the dataset. The image is displayed.

2.3 How Do I Integrate Multiple Object Detection Datasets into One Dataset?

Create a parent directory in an OBS bucket, in the directory add the same number of folders as that of datasets, export one dataset to one folder, and use the parent directory to create a dataset.

Log in to the ModelArts management console and choose **Data Management > Datasets**. Click the target dataset to switch to its **Dashboard** page. Then, click **Export** in the upper right corner of the page to export the dataset to a folder in the OBS parent directory.

2.4 What Do I Do If Importing a Dataset Failed?

The possible cause is that the storage class of the target OBS bucket is incorrect. In this case, select a bucket of the standard storage class to import data.

Region

Regions are geographic areas isolated from each other. Resources are region-specific and cannot be used across regions through internal network connections. For low network latency and quick resource access, select the nearest region. Once a bucket is created, the region cannot be changed.

Data Redundancy Policy Multi-AZ storage Single-AZ storage

Multi-AZ storage improves data availability. Multi-AZ storage has higher price than single-AZ storage. [Pricing details](#)

Once the multi-AZ storage is enabled, it cannot be disabled.

Bucket Name

Naming conventions

- The name must be globally unique in OBS.
- The name must contain 3 to 63 characters. Only lowercase letters, digits, hyphens (-), and periods (.) are allowed.
- The name cannot start or end with a period (.) or hyphen (-), and cannot contain two consecutive periods (..) or contain a period (.) and a hyphen (-) adjacent to each other.
- The name cannot be an IP address.
- If the name contains any periods, a security certificate verification message may appear when you access the bucket or its objects by entering a domain name.
- The name of a bucket or parallel file system can be reused 30 minutes after the bucket or parallel file system is deleted.

Storage Class Standard Infrequent Access Archive

Optimized for frequently accessed (multiple times per month) data such as small and essential files that require low latency.

The storage class of a bucket is inherited by objects uploaded to the bucket by default. You can also change the storage class of an object when uploading it to the bucket. [Learn more](#)

2.5 Can a Table Dataset Be Labeled?

Table datasets cannot be labeled. They are suitable for processing structured data such as tables. Table files are in CSV format. You can preview up to 100 data records in a table.

2.6 Why Does Data Fail to Be Imported Using the Manifest File?

Symptom

Failed to use the manifest file of the published dataset to import data again.

Possible Cause

Data has been changed in the OBS directory of the published dataset, for example, images have been deleted. Therefore, the manifest file is inconsistent with data in the OBS directory. As a result, an error occurs when the manifest file is used to import data again.

Solution

- Method 1 (recommended): Publish a new version of the dataset again and use the new manifest file to import data.
- Method 2: Modify the manifest file on your local PC, search for data changes in the OBS directory, and modify the manifest file accordingly. Ensure that the manifest file is consistent with data in the OBS directory, and then import data using the new manifest file.

2.7 Where Are Labeling Results Stored?

The ModelArts console provides data visualization capabilities, which allows you to view detailed data and labeling information on the console. To learn more about the path for storing labeling results, see the following description.

Background

When creating a dataset in ModelArts, set both **Input Dataset Path** and **Output Dataset Path** to OBS.

- **Input Dataset Path:** OBS path where the raw data is stored.
- **Output Dataset Path:** Under this path, directories are generated based on the dataset version after data is labeled in ModelArts and datasets are published. The manifest files (containing data and labeling information) used in ModelArts are also stored in this path. For details about the files, see .

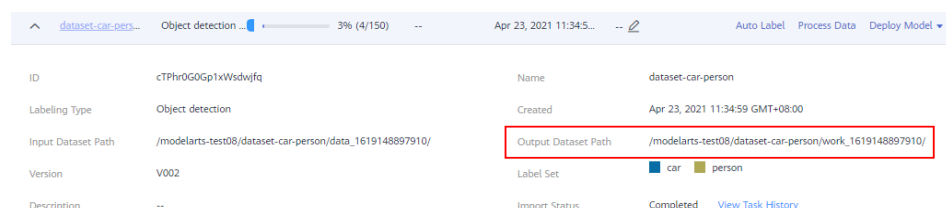
Procedure

1. Log in to the ModelArts console and choose **Data Management > Datasets**.
2. Select your desired dataset and click the triangle icon on the left of the dataset name to expand the dataset details. You can obtain the OBS path set for **Output Dataset Path**.

NOTE

Before obtaining labeling results, ensure that at least one dataset version is available.

Figure 2-2 Dataset details

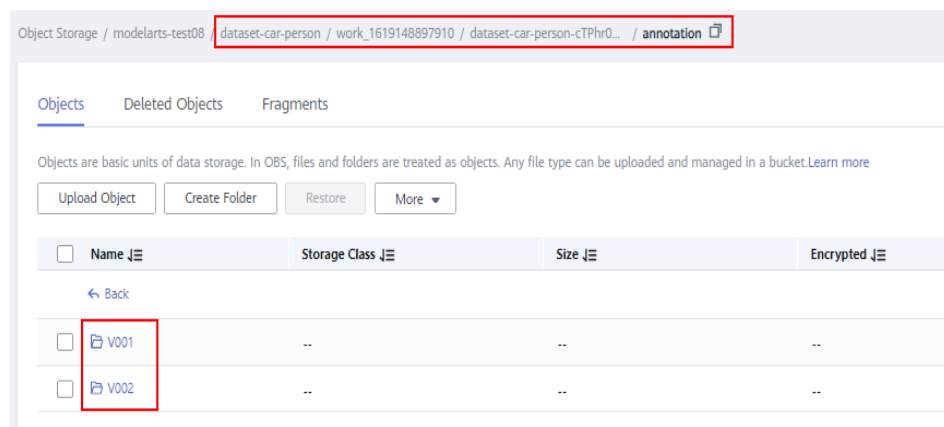


The screenshot shows the details of a dataset named 'dataset-car-person'. The 'Output Dataset Path' is highlighted with a red box. The path is '/modelarts-test08/dataset-car-person/work_1619148897910/'. The 'Label Set' is 'car' and 'person'. The 'Import Status' is 'Completed'.

ID	cTPhr0G0Gp1xWsdwJfQ	Name	dataset-car-person
Labeling Type	Object detection	Created	Apr 23, 2021 11:34:59 GMT+08:00
Input Dataset Path	/modelarts-test08/dataset-car-person/data_1619148897910/	Output Dataset Path	/modelarts-test08/dataset-car-person/work_1619148897910/
Version	V002	Label Set	car person
Description	--	Import Status	Completed View Task History

3. Log in to the OBS console and locate the directory of the corresponding dataset version from the OBS path obtained in 2 to obtain the labeling result of the dataset.

Figure 2-3 Obtaining the labeling result



2.8 How Do I Download Labeling Results to a Local PC?

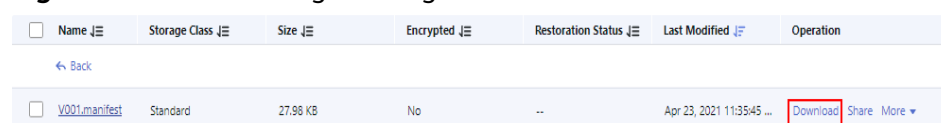
After being published, the labeling information and data in ModelArts datasets are stored as manifest files in the OBS path set for **Output Dataset Path**.

To obtain the OBS path, do as follows:

1. Log in to the ModelArts management console and choose **Data Management > Datasets**.
2. Locate the target dataset and click the triangle icon on the left of the dataset name to expand the dataset details. You can obtain the OBS path set for **Output Dataset Path**.
3. Log in to the OBS management console and locate the version directory from the obtained OBS path to obtain the labeling result of the dataset.

To download the labeling results to a local PC, go to the OBS path where the manifest files are stored and click **Download**.

Figure 2-4 Downloading labeling results



2.9 Why Cannot Team Members Receive Emails for a Team Labeling Task?

The possible causes are as follows:

- All dataset data has been labeled. An email can be sent to team members only if there is unlabeled data in the dataset when the team labeling task is created.
- Team members receive emails for team labeling tasks. No email will be sent when you create a labeling team or add members to a labeling team.
- Your email address has not been configured or has been incorrectly configured. For details about how to configure an email address, see .

- Team members' email addresses are blocked.

2.10 Can I Add Multiple Labeling Boxes to an Object Detection Dataset Image?

Yes.

For an object detection dataset, you can add multiple labeling boxes and labels to an image during labeling. Note that the labeling boxes cannot extend beyond the image boundary.

2.11 How Do I Merge Two Datasets?

Datasets cannot be merged.

However, you can perform the following operations to merge the data of two datasets into one dataset.

For example, to merge datasets A and B, do the following:

1. Publish datasets A and B.
2. Obtain the manifest files of the two datasets from the OBS path set for **Output Dataset Path**.
3. Create empty dataset C and select an empty OBS folder for **Input Dataset Path**.
4. Import the manifest files of datasets A and B to dataset C.

After the import is complete, data in datasets A and B is merged into dataset C. To use the merged dataset, publish dataset C.

2.12 Why Are Images Displayed in Different Angles Under the Same Account?

There are rotation angles of certain images, and the rules of processing such images vary depending on browsers. The following figures show compatibility with browsers.

- L indicates the latest version. L3 indicates the latest three stable browser versions when the product is released.
- If your browser is of an earlier version, the page display will be adversely affected, and the system will prompt you to upgrade your browser.
- If your browser is not compatible with the management console, the system will advise you to upgrade your browser or install a desired browser.

2.13 Do I Need to Train Data Again If New Data Is Added After Auto Labeling Is Complete?

After auto labeling is complete, confirm the labeled data. If you add new data before confirming the labeled data, all unlabeled data will be automatically

labeled again. If you add new data after confirming the labeled data, the data must be trained again.

2.14 Why Does the System Display a Message Indicating My Label Fails to Save on ModelArts?

Symptom

Take the Google Chrome browser as an example. When an image is labeled for the first time, the system displays a message in the upper right corner, indicating that the label fails to save. But when the same image is labeled the second time, a message is displayed, indicating that the label is saved. This issue occurs occasionally. When this issue occurs, the request status is **(failed)net::ERR_ADDRESS_IN_USE**, which is obtained by pressing **F12** on the Google Chrome browser and clicking **Network**.

Name	Status	Type	Initiator	Size	Time	Waterfall
samples	200	xhr	juetu.us.8622	762 B	599 ms	
004009abe08be8af16c818c125e207vorker_ida2a43868ea34e5b6de7443e87d966194	200	xhr	juetu.us.8622	5.1 kB	405 ms	
2021-12-20-11-05-12-676.jpg?AccessKeyId=S78PMLDlW4...34D44%3D&Signature=TWG045puvm9Lco...	200	jpeg	data.label.notificationCtrl.js:1	76.6 kB	312 ms	
001a56ad5410ecb06e5e4859a03c07vorker_ida2a43868ea34e5b6de7443e87d966194	200	xhr	juetu.us.8622	5.3 kB	158 ms	
2021-12-20-11-54-52-9130.jpg?AccessKeyId=MMW58C83...303d&Signature=%2Bc38P%2FQrpnGnDhw...	200	jpeg	data	444 kB	232 ms	
samples	(failed)	net::ERR_ADDRESS_IN...	xhr	0 B	165 ms	
004009abe08be8af16c818c125e207vorker_ida2a43868ea34e5b6de7443e87d966194	200	xhr	juetu.us.8622	5.1 kB	357 ms	
2021-12-20-11-05-12-676.jpg?AccessKeyId=83H2F3R27...pmvR0%3D&Signature=Fxk3hVAe4d05g6...	200	jpeg	data.label.notificationCtrl.js:1	76.6 kB	97 ms	
me	200	xhr	juetu.us.8622	914 B	270 ms	

Possible Cause

The local network is faulty, for example, the network is unstable, or the network configuration is incorrect.

Solution

- Switch to a stable network and try again.
- Initialize the network configuration. To do so, start **cmd** as the administrator, run the **netsh winsock reset** command, and restart the computer. Then, log in to the data labeling platform again.

2.15 Can One Label Be Identified Among Multiple Labels?

After a model is trained with multiple labels and deployed as a real-time service, all the labels are identified. If only one type of label needs to be identified, train a model dedicated for identifying the label. To speed up the label identification, select a high flavor for deploying the model.

2.16 Why Are Newly Added Images Not Automatically Labeled After Data Amplification Is Enabled?

After data amplification is enabled, images newly added in image classification datasets cannot be automatically labeled, but those added in object detection datasets can be.

2.17 Why Cannot Videos in a Video Dataset Be Displayed or Played?

If the issue occurs, check the video format. Only MP4 videos can be displayed and played.

2.18 How Do I Add Images to a Validation or Training Dataset?

You are not allowed to manually add images to a training or validation dataset, but can only set a training and validation ratio. Then, the system randomly allocates the images to the training and validation datasets based on the ratio.

2.19 Can I Customize Labels for an Object Detection Dataset?

Yes. You can add custom labels to the label set of an object detection dataset by modifying the dataset.

Figure 2-5 Modify Dataset

Modify Dataset ×

Name

Description
0/256

Label Set

<input type="text" value="blue"/>	▼	+	✖
<input type="text" value="none"/>	▼	+	✖

+ Add Label

OK Cancel

2.20 What ModelArts Data Management Can Be Used for?

The functions provided ModelArts data management vary depending on the type of the dataset.

Data set Type	Labeling Type	Creating a Dataset	Importing Data	Exporting Data	Publishing a Dataset	Modifying a Dataset	Managing Dataset Versions	Auto Grouping	Data Feature Engineering
Files	Image classification	Supported	Supported	Supported	Supported	Supported	Supported	Supported	Supported
	Object detection	Supported	Supported	Supported	Supported	Supported	Supported	Supported	Supported
	Image segmentation	Supported	Supported	Supported	Supported	Supported	Supported	Supported	N/A
	Sound classification	Supported	Supported	N/A	Supported	Supported	Supported	N/A	N/A
	Speech labeling	Supported	Supported	N/A	Supported	Supported	Supported	N/A	N/A
	Speech paragraph labeling	Supported	Supported	N/A	Supported	Supported	Supported	N/A	N/A
	Text classification	Supported	Supported	N/A	Supported	Supported	Supported	N/A	N/A
	Name entity recognition	Supported	Supported	N/A	Supported	Supported	Supported	N/A	N/A
	Text triplet	Supported	Supported	N/A	Supported	Supported	Supported	N/A	N/A

Data set Type	Labeling Type	Creating a Dataset	Importing Data	Exporting Data	Publishing a Dataset	Modifying a Dataset	Managing Dataset Versions	Auto Grouping	Data Feature Engineering
	Videos	Supported	Supported	N/A	Supported	Supported	Supported	N/A	N/A
	Free format	Supported	N/A	Supported	Supported	Supported	Supported	N/A	N/A
Tables	Tables	Supported	Supported	N/A	Supported	Supported	Supported	N/A	N/A

2.21 Why Is My Newly Created Bucket Unavailable?

Ensure that the created bucket and ModelArts are in the same region. Additionally, the bucket is not encrypted. ModelArts does not support encrypted OBS buckets.

2.22 Will My Old-Version Datasets Be Cleared After the Old Version Is Discontinued? The existing datasets and the ones newly created in the old version will be retained after the old version is discontinued.

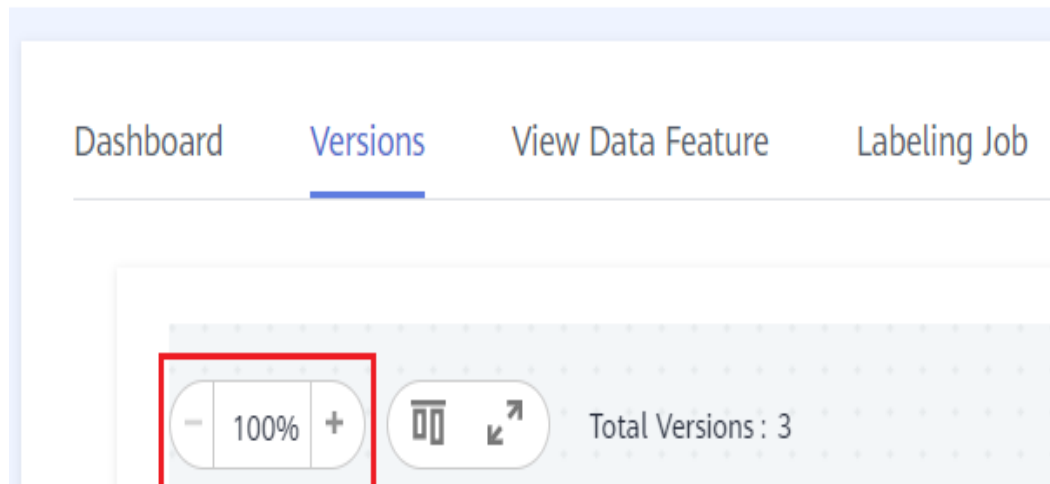
The new version is compatible with the old version. However, the datasets created in the new version cannot be displayed in the dataset list of the old version.

2.23 Why Is My New Dataset Version Unavailable in Versions?

The version list can be zoomed in or out. Zoom out the page before searching.

Click the name of the target dataset to go to the dataset overview page. Then, zoom out the **Versions** page.

Datasets / Dashboard(dataset-animal-swt)

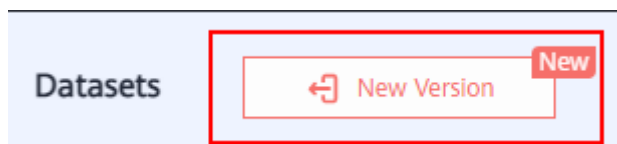


2.24 How Do I View the Size of a Dataset?

Only the number of samples in a dataset is collected in data management. There is no entrance to view the dataset size.

2.25 How Do I View Labeling Details of a New Dataset?

1. Log in to the ModelArts management console and choose **Data Management > Datasets** from the navigation pane on the left.
2. Go to the new dataset version as prompted.



3. Locate the target dataset by name and click its name. The **Dashboard** tab page is displayed.
4. On the **Dashboard** tab page, click **View Details** in the **Labeling Information** area.

Labeling Information

● Image classification [View Details](#)

Name	Labels
test	3

2.26 How Do I Export Labeled Data?

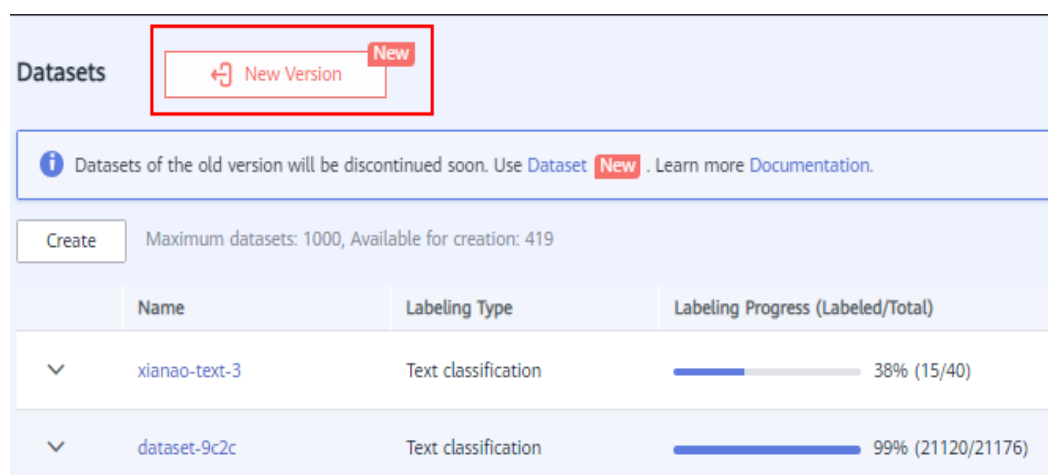
Only datasets of image classification, object detection, and image segmentation types can be exported.

- For image classification datasets, only the label files in TXT format can be exported.
- For object detection datasets, only XML label files in Pascal VOC format can be exported.
- For image segmentation datasets, only XML label files in Pascal VOC format and mask images can be exported.

For other types of datasets, use to publish the datasets.

2.27 Why Cannot I Find My Newly Created Dataset?

The datasets of the new version are not displayed on the dataset page of the old version. To view the datasets of the new version, switch to the dataset page of the new version.



The screenshot shows the 'Datasets' page in ModelArts. At the top, there is a 'New Version' button with a red 'New' badge. Below it, an information message states: 'Datasets of the old version will be discontinued soon. Use Dataset **New**. Learn more [Documentation](#).' A 'Create' button is visible, along with the text 'Maximum datasets: 1000, Available for creation: 419'. Below this is a table of datasets:

	Name	Labeling Type	Labeling Progress (Labeled/Total)
▼	xianao-text-3	Text classification	38% (15/40)
▼	dataset-9c2c	Text classification	99% (21120/21176)

2.28 What Do I Do If the Database Quota Is Incorrect?

The quota for the datasets of both the old and new versions is 100. On the dataset page of the new version, all created datasets are displayed. However, the dataset page of the old version does not display the new-version datasets. Go to the dataset page of the new version to view the datasets.

2.29 How Do I Split a Dataset?

When you publish a dataset, only the dataset of the image classification, object detection, text classification, or sound classification type supports data splitting.

By default, data splitting is disabled. After this function is enabled, set the training and validation ratios.

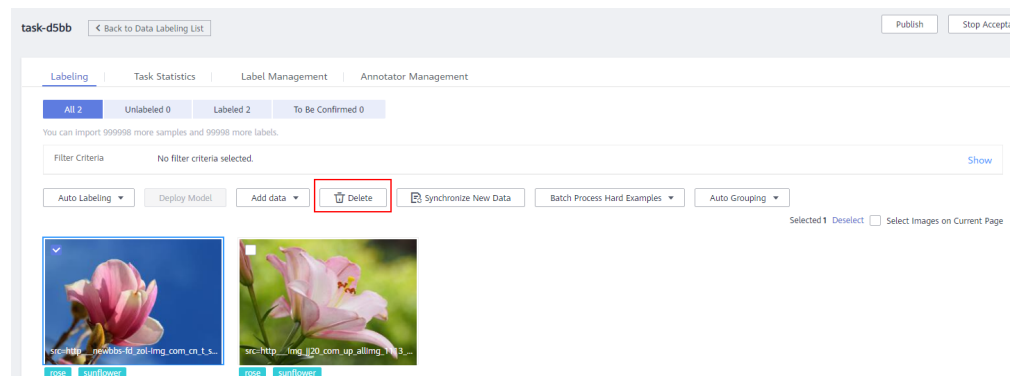
Enter a value ranging from 0 to 1 for the training set ratio. After the training set ratio is set, the validation set ratio is determined. The sum of the training set ratio and the validation set ratio is 1.

The training set ratio is the ratio of sample data used for model training. The validation set ratio is the ratio of the sample data used for model validation. The training and validation ratios affect the performance of training templates.

2.30 How Do I Delete a Dataset Image?

1. Log in to the ModelArts management console. In the navigation pane, choose **Data Management > Label Data**. The data labeling list is displayed. Click the dataset from which you want to delete images. The labeling details page is displayed.
2. On the **All**, **Unlabeled**, or **Labeled** tab page, select the images to be deleted or click **Select Images on Current Page**, and click **Delete** in the upper left corner to delete them. In the displayed dialog box, select or deselect **Delete the source files from OBS** as required. After confirmation, click **Yes** to delete the images.

If a tick is displayed in the upper left corner of an image, the image is selected. If no image is selected on the page, the **Delete** button is unavailable.



2.31 Why Is There No Sample in the ModelArts Dataset Downloaded from AI Gallery and Then an OBS Bucket?

Check the format of the data downloaded from AI Gallery. For example, compressed packages and Excel files will be ignored. The following table lists the supported formats.

Data set Type	Labeling Type	Creating a Dataset	Importing Data	Exporting Data	Publishing a Dataset	Modifying a Dataset	Managing Dataset Versions	Auto Grouping	Data Feature Engineering
Files	Image classification	Supported	Supported	Supported	Supported	Supported	Supported	Supported	Supported
	Object detection	Supported	Supported	Supported	Supported	Supported	Supported	Supported	Supported
	Image segmentation	Supported	Supported	Supported	Supported	Supported	Supported	Supported	N/A
	Sound classification	Supported	Supported	N/A	Supported	Supported	Supported	N/A	N/A
	Speech labeling	Supported	Supported	N/A	Supported	Supported	Supported	N/A	N/A
	Speech paragraph labeling	Supported	Supported	N/A	Supported	Supported	Supported	N/A	N/A
	Text classification	Supported	Supported	N/A	Supported	Supported	Supported	N/A	N/A
	Name entity recognition	Supported	Supported	N/A	Supported	Supported	Supported	N/A	N/A
	Text triplet	Supported	Supported	N/A	Supported	Supported	Supported	N/A	N/A

Data set Type	Labeling Type	Creating a Dataset	Importing Data	Exporting Data	Publishing a Dataset	Modifying a Dataset	Managing Dataset Versions	Auto Grouping	Data Feature Engineering
	Videos	Supported	Supported	N/A	Supported	Supported	Supported	N/A	N/A
	Free format	Supported	N/A	Supported	Supported	Supported	Supported	N/A	N/A
Tables	Tables	Supported	Supported	N/A	Supported	Supported	Supported	N/A	N/A

3 Notebook

3.1 Constraints

3.1.1 Is sudo Privilege Escalation Supported?

For security purposes, notebook instances do not support sudo privilege escalation.

3.1.2 Does ModelArts Support apt-get?

Currently, **Terminal** in ModelArts DevEnviron does not support **apt-get**. You can use **a** to support it.

3.1.3 Is the Keras Engine Supported?

Notebook instances in **DevEnviron** support the Keras engine. The Keras engine is not supported in job training and model deployment (inference).

Keras is an advanced neural network API written in Python. It is capable of running on top of TensorFlow, CNTK, or Theano. Notebook instances in **DevEnviron** support **tf.keras**.

How Do I View Keras Versions?

1. On the ModelArts management console, create a notebook instance with image **TensorFlow-1.13** or **TensorFlow-1.15**.
2. Access the notebook instance. In JupyterLab, run **!pip list** to view Keras versions.

Figure 3-1 Viewing Keras versions

```

Untitled.ipynb
+ ✂ 📄 ▶ ■ ↻ ⏩ Code ⌵ ⌚ git
idna 3.3
importlib-metadata 4.11.3
importlib-resources 5.7.1
iniconfig 1.1.1
ipykernel 6.7.0
ipython 7.31.1
ipython-genutils 0.2.0
jedi 0.18.1
Jinja2 3.1.2
jinja2-time 0.2.0
jmespath 0.10.0
joblib 1.1.0
jsonschema 4.4.0
jupyter-client 7.1.2
jupyter-core 4.9.1
Keras 2.2.4
Keras-Applications 1.0.8
Keras-Preprocessing 1.1.2
keyboard 0.13.5
lxml 4.8.0.post20220315201753
  
```

3.1.4 Does ModelArts Support the Caffe Engine?

The Python 2 environment of ModelArts supports Caffe, but the Python 3 environment does not support it.

3.1.5 Can I Install MoXing in a Local Environment?

No. MoXing can be used only on ModelArts.

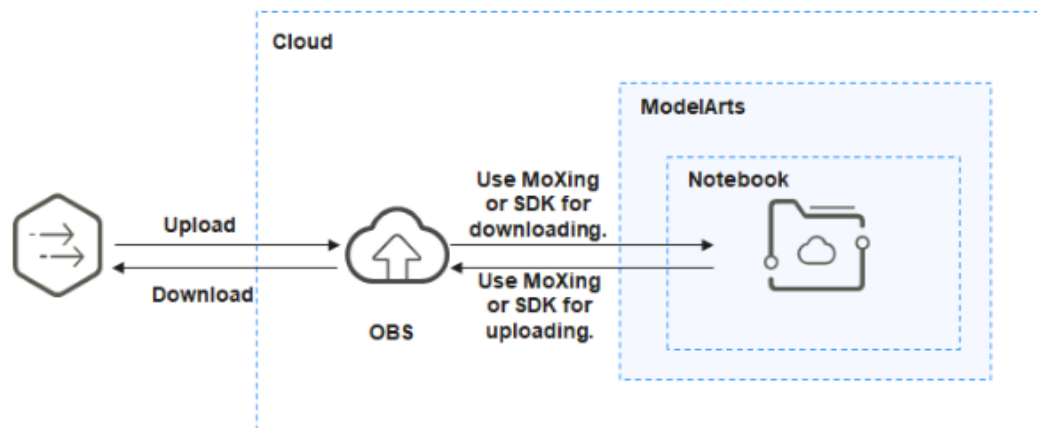
3.1.6 Can Notebook Instances Be Remotely Logged In?

Notebook instances of the old version cannot be remotely logged in.

3.2 Data Upload or Download

3.2.1 How Do I Upload a File from a Notebook Instance to OBS or Download a File from OBS to a Notebook Instance?

In a notebook instance, you can call the ModelArts MoXing API or SDK to exchange data with OBS for uploading a file to OBS or downloading a file from OBS to the notebook instance.

Figure 3-2 Uploading or downloading a file

Method 1: Using MoXing to Upload and Download a File

Developed by the ModelArts team, MoXing is a distributed training acceleration framework built on open-source deep learning engines such as TensorFlow and PyTorch. MoXing makes model coding easier and more efficient.

MoXing provides a set of file object APIs for reading and writing OBS files.

Sample code:

```
import moxing as mox

# Download the OBS folder sub_dir_0 from OBS to a notebook instance.
mox.file.copy_parallel('obs://bucket_name/sub_dir_0', '/home/ma-user/work/sub_dir_0')
# Download the OBS file obs_file.txt from OBS to a notebook instance.
mox.file.copy('obs://bucket_name/obs_file.txt', '/home/ma-user/work/obs_file.txt')

# Upload the OBS folder sub_dir_0 from a notebook instance to OBS.
mox.file.copy_parallel('/home/ma-user/work/sub_dir_0', 'obs://bucket_name/sub_dir_0')
# Upload the OBS file obs_file.txt from a notebook instance to OBS.
mox.file.copy('/home/ma-user/work/obs_file.txt', 'obs://bucket_name/obs_file.txt')
```

Method 2: Using SDK to Upload and Download a File

Call the ModelArts SDK for downloading a file from OBS.

Sample code: Download **file1.txt** from OBS to **/home/ma-user/work/** in the notebook instance. All the bucket name, folder name, and file name are customizable.

```
from modelarts.session import Session
session = Session()
session.obs.download_file(src_obs_file="obs://bucket-name/dir1/file1.txt", dst_local_dir="/home/ma-user/work/")
```

Call the ModelArts SDK for downloading a folder from OBS.

Sample code: Download **dir1** from OBS to **/home/ma-user/work/** in the notebook instance. The bucket name and folder name are customizable.

```
from modelarts.session import Session
session = Session()
session.obs.download_dir(src_obs_dir="obs://bucket-name/dir1/", dst_local_dir="/home/ma-user/work/")
```

Call the ModelArts SDK for uploading a file to OBS.

Sample code: Upload **file1.txt** in the notebook instance to OBS bucket **obs://bucket-name/dir1/**. All the bucket name, folder name, and file name are customizable.

```
from modelarts.session import Session
session = Session()
session.obs.upload_file(src_local_file='/home/ma-user/work/file1.txt', dst_obs_dir='obs://bucket-name/dir1/')
```

Call the ModelArts SDK for uploading a folder to OBS.

Sample code: Upload **/work/** in the notebook instance to **obs://bucket-name/dir1/work/** of **bucket-name**. The bucket name and folder name are customizable.

```
from modelarts.session import Session
session = Session()
session.obs.upload_dir(src_local_dir='/home/ma-user/work/', dst_obs_dir='obs://bucket-name/dir1/')
```

3.2.2 How Do I Upload Local Files to a Notebook Instance?

For details about how to upload files to JupyterLab in notebook of the new version, see [Uploading Files to JupyterLab](#).

3.2.3 How Do I Import Large Files to a Notebook Instance?

- **Large files (files larger than 100 MB)**

Use OBS to upload large files. To do so, use OBS Browser to upload a local file to an OBS bucket and use ModelArts SDK to download the file from OBS to a notebook instance.

For details about how to use ModelArts SDK or MoXing to download files from OBS, see [How Do I Upload a File from a Notebook Instance to OBS or Download a File from OBS to a Notebook Instance?](#)

- **Folders**

Compress a folder into a package and upload the package in the same way as uploading a large file. After the package is uploaded to a notebook instance, decompress it on the **Terminal** page.

```
unzip xxx.zip # Directly decompress the package in the path where the package is stored.
```

For more details, search for the decompression command in mainstream search engines.

3.2.4 Where Will the Data Be Uploaded to?

Data may be stored in OBS or EVS, depending on which kind of storage you have configured for your Notebook instances:

- **OBS**

After you click **upload**, the data is directly uploaded to the target OBS path specified when the notebook instance was created.

- **EVS**

After you click **upload**, the data is uploaded to the instance container, that is, the **~/work** directory on the **Terminal** page.

3.2.5 How Do I Download Files from a Notebook Instance to a Local Computer?

For details about how to download files from JupyterLab in notebook of the new version, see .

3.3 Data Storage

3.3.1 When Creating a Notebook Instance, What Is the Difference If You Select or Do Not Select a Storage Path? Which Directories Can Store Big Data?

Selecting an Instance with a Storage Path (EVS or OBS)

The storage path for a Notebook instance can be OBS or EVS.

- OBS
All files in the notebook list are read and written based on the content in the selected OBS path. Any additions, modifications, or deletions are performed on the content in the corresponding OBS path. The operations are not related to the instance space. To synchronize the content to the instance space, select the content and click **Sync OBS** to synchronize the selected content to the container space.
- EVS
All files in the notebook list are read and written based on the content in the selected EVS path. Any additions, modifications, or deletions are performed on the content in the corresponding EVS path. The operations are not related to the instance space. The default disk space is 5 GB, and there is no charge for this storage. If the disk space exceeds 5 GB, the additional space is charged by GB from the time when a notebook instance is created to the time the notebook instance is deleted. By default, ModelArts uses the ultra-high I/O EVS disks and cannot be modified. If the disk space exceeds 5 GB, the additional space is billed based on ultra-high I/O EVS pricing. If EVS is used, you can mount your data to the `~/work` directory.

The notebook instance of the EVS type does not support the **Sync OBS** function.

NOTE

If big data needs to be stored, EVS is recommended.

Instance Without a Specified Storage Path

The read and write operations on all files in the notebook list are performed on the content in the container, and are not related to OBS or EVS. If you store objects in OBS, you can use the MoXing APIs to access the objects. The default storage space is 5 GB.

3.3.2 Where Is Data Stored After the Sync OBS Function Is Used?

1. Log in to the ModelArts management console, and choose **DevEnviron > Notebooks**.
2. In the **Operation** column of the target notebook instance in the notebook list, click **Open** to go to the **Jupyter** page.
3. On the **Files** tab page of the **Jupyter** page, select the target file and click **Sync OBS** in the upper part of the page to synchronize the file. The file is stored in the **~/work** directory of the instance.
4. On the **Files** tab page of the **Jupyter** page, click **New** and select **Terminal**. The **Terminal** page is displayed.
5. Run the following command to go to the **~/work** directory.

```
cd work
```
6. Run the **ls** command in the **~/work** directory to view the files.

3.3.3 How Do I Rename an OBS File?

OBS files cannot be renamed on OBS Console. To rename an OBS file, call a MoXing API. Run the command in an existing notebook instance or after creating a new notebook instance.

The following example renames **obs_file.txt** to **obs_file_2.txt**.

```
import moxing as mox
mox.file.rename('obs://bucket_name/obs_file.txt', 'obs://bucket_name/obs_file_2.txt')
```

3.3.4 Do Files in /cache Still Exist After a Notebook Instance is Stopped or Restarted? How Do I Avoid a Restart?

/cache is a temporary directory and will not be saved. After an instance using OBS storage is stopped, data in the **~/work** directory will be deleted. After a notebook instance is restarted, all cached data except the data in the OBS bucket is lost, and your model or code is unavailable.

To avoid a restart, do not train heavy-load jobs that consume large amounts of CPU, GPU, or memory resources in DevEnviron.

3.3.5 How Do I Use the pandas Library to Process Data in OBS Buckets?

When using libraries, for example, pandas, you can download data from OBS to a notebook instance for local processing. For details, see .

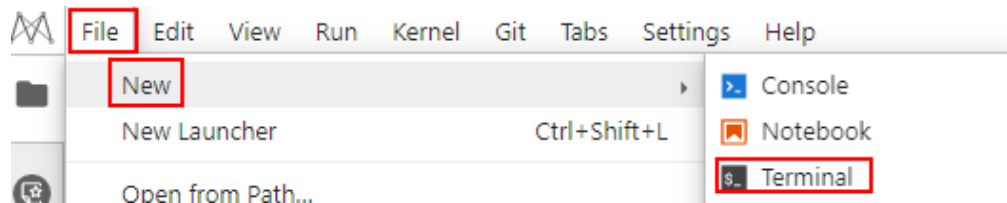
3.4 Environment Configurations

3.4.1 How Do I Enable the Terminal Function in DevEnviron of ModelArts?

1. Log in to the ModelArts management console, and choose **DevEnviron > Notebooks**.

2. Create a notebook instance. When the instance is running, click **Open** in the **Operation** column. The **JupyterLab** page is displayed.
3. Choose **File > New > Terminal**. The **Terminal** page is displayed.

Figure 3-3 Going to the **Terminal** page



3.4.2 How Do I Install External Libraries in a Notebook Instance?

Multiple environments have been integrated into ModelArts Notebook. These environments contain Jupyter Notebook and Python packages, including TensorFlow, MXNet, Caffe, PyTorch, and Spark. You can use **pip install** to install external libraries in Jupyter Notebook or on the **Terminal** page.

Installing External Libraries in Jupyter Notebook

You can use Jupyter Notebook to install external libraries in the **TensorFlow-1.8** environment. For example, to install Shapely:

1. Open a notebook instance.
2. On the Jupyter Notebook dashboard, choose **New > TensorFlow-1.8**.
3. In the new notebook instance, enter the following command in the code input bar:

```
!pip install Shapely
```

Installing External Libraries on the Terminal Page

You can use **pip** to install external libraries in the **TensorFlow-1.8** environment on the **Terminal** page. For example, to install Shapely:

1. Open a notebook instance.
2. On the Jupyter Notebook dashboard, choose **New > Terminal**.
3. Enter the following command in the code input bar to obtain the command for activating **TensorFlow-1.8** and activate it:

```
cat /home/ma-user/README
```

```
source /home/ma-user/anaconda3/bin/activate TensorFlow-1.8
```

NOTE

If you use another engine, replace **TensorFlow-1.8** in the command with the name and version of the engine.

Figure 3-4 Activating the environment

```
sh-4.3$cat /home/ma-user/README
Please use one of following command to start the specified framework environment.

for Conda-python3 ----- source /home/ma-user/anaconda3/bin/activate base
for MXNet-1.2.1 ----- source /home/ma-user/anaconda3/bin/activate MXNet-1.2.1
for PySpark-2.3.2 ----- source /home/ma-user/anaconda3/bin/activate PySpark-2.3.2
for Pytorch-1.0.0 ----- source /home/ma-user/anaconda3/bin/activate Pytorch-1.0.0
for TensorFlow-1.13.1 ----- source /home/ma-user/anaconda3/bin/activate TensorFlow-1.13.1
for TensorFlow-1.8 ----- source /home/ma-user/anaconda3/bin/activate TensorFlow-1.8
for XGBoost-Sklearn ----- source /home/ma-user/anaconda3/bin/activate XGBoost-Sklearn
```

4. Type the following command in the code input bar to install Shapely:

pip install Shapely

NOTE

A new independent running environment is opened when a ModelArts training job is created. The new environment is not associated with the packages installed in the notebook environment. You need to add `os.system('pip install xxx')` to the boot code before importing the installation package.

For example, if you need to use the Shapely dependencies for a training job, add the following code to the boot code after the notebook instance is installed:

```
import os
os.system('pip install Shapely')
import Shapely
```

3.4.3 How Do I Create a Soft Link and Switch a CUDA Version?

NOTE

This section applies only to notebook of the old version. The function is available to only existing users of the old-version notebook.

In the new-version notebook, there is only one CUDA version. Therefore, the CUDA version cannot be switched.

Create a soft link.

```
sudo ln -snf xxx
```

Check the CUDA versions in the environment.

```
ll /usr/local | grep cuda
```

Switch the CUDA version by setting *version* to the CUDA version.

```
sudo ln -snf /usr/local/cuda-{version} /usr/local/cuda
```

View the CUDA version in the new-version notebook. The following shows an example.

Figure 3-5 Checking the CUDA version in the current environment

```
ll /usr/local | grep cuda
lrwxrwxrwx 1 root 9 Feb 9 09:28 cuda -> cuda-10.2/
drwxr-xr-x 12 root 4096 Feb 10 09:28 cuda-10.2/
```

In the preceding example, the CUDA version is 10.2, which cannot be switched. If you run the version switchover command, an error will be reported.

3.4.4 Should I Access the Python Environment Same as the Notebook Kernel of the Current Instance in the Terminal?

The following uses the TensorFlow-1.8 engine as an example. The operations on other engines are similar. You only need to replace the engine name and version number in the commands.

1. Open a notebook instance.
2. On the displayed Jupyter dashboard, click **New** and select **Terminal** from the shortcut menu.
3. The **README** file exists in the user directory on the **Terminal**. The file describes how to switch different Python environments.
 - Enter the following command in the code input bar to search for a virtual environment name:
cat /home/ma-user/README
 - Enter the following command in the code input bar to activate different environments: For example, for the TensorFlow-1.8 engine in the Python2 environment, you can replace ****** with **TensorFlow-1.8** in the command line. You can replace it with the engine of another version. Example:
source /home/ma-user/anaconda2/bin/activate **

3.4.5 How Do I Import a Python Library to a Notebook Instance to Resolve the ModuleNotFoundError Error?

Importing the Python Library to a Notebook Instance with EVS Disks Mounted

1. Obtain the address of the Python library to be imported, and follow the instructions in **Adding Folders to sys.path** of Python 3 to import the Python library. There are two methods to import the Python library. The widely used method is to use the **PYTHONPATH** environment variable.
2. After the Python library is imported, you can view your **PYTHONPATH** in the notebook instance.
 - a. View the **PYTHONPATH** in **IPYNB**. Run the following command in the code input bar to view the **PYTHONPATH**. If the returned address is the same as the address of the Python library that you set, the Python library has been imported.
!echo \$PYTHONPATH
 - b. View the **PYTHONPATH** on the **Terminal** page. Run the following command to view the **PYTHONPATH**. If the returned address is the same as the address of the Python library that you set, the Python library has been imported.
echo \$PYTHONPATH

Importing the Python Library to a Notebook Instance with OBS Storage

The methods of importing the Python library vary according to the size of the python library file.

1. If the size of the python library file is less than 100 MB, use either of the following methods:
 - Upload the Python files to OBS, and synchronize the Python files from OBS to the notebook instance . Follow the instructions in Adding Folders to sys.path of Python 3 to import the Python library. Use the **PYTHONPATH** environment variable to import the library.
 - Upload the Python files to OBS, and synchronize the Python files from OBS to the notebook instance using SDKs. Follow the instructions in Adding Folders to sys.path of Python 3 to import the Python library. You are advised to use the **PYTHONPATH** environment variable to import the library.
2. If the size of the Python library file is greater than 100 MB, use the following method:

Upload the Python files to OBS, and synchronize the Python files from OBS to the notebook instance using MoXing, and then MoXing can use the OBS files. Follow the instructions in Adding Folders to sys.path of Python 3 to import the Python library. Use the **PYTHONPATH** environment variable to import the library.

After the Python library is imported, you can view your **PYTHONPATH** in the notebook instance.

1. View the **PYTHONPATH** in **IPYNB**. Run the following command in the code input bar to view the **PYTHONPATH**. If the returned address is the same as the address of the Python library that you set, the Python library has been imported.
!echo \$PYTHONPATH
2. View the **PYTHONPATH** on the **Terminal** page. Run the following command to view the **PYTHONPATH**. If the returned address is the same as the address of the Python library that you set, the Python library has been imported.
echo \$PYTHONPATH

3.4.6 How Can I Resolve Abnormal Font Display on a ModelArts Notebook Accessed from iOS?

Symptom

When a ModelArts notebook is accessed from iOS, the font is displayed abnormally.

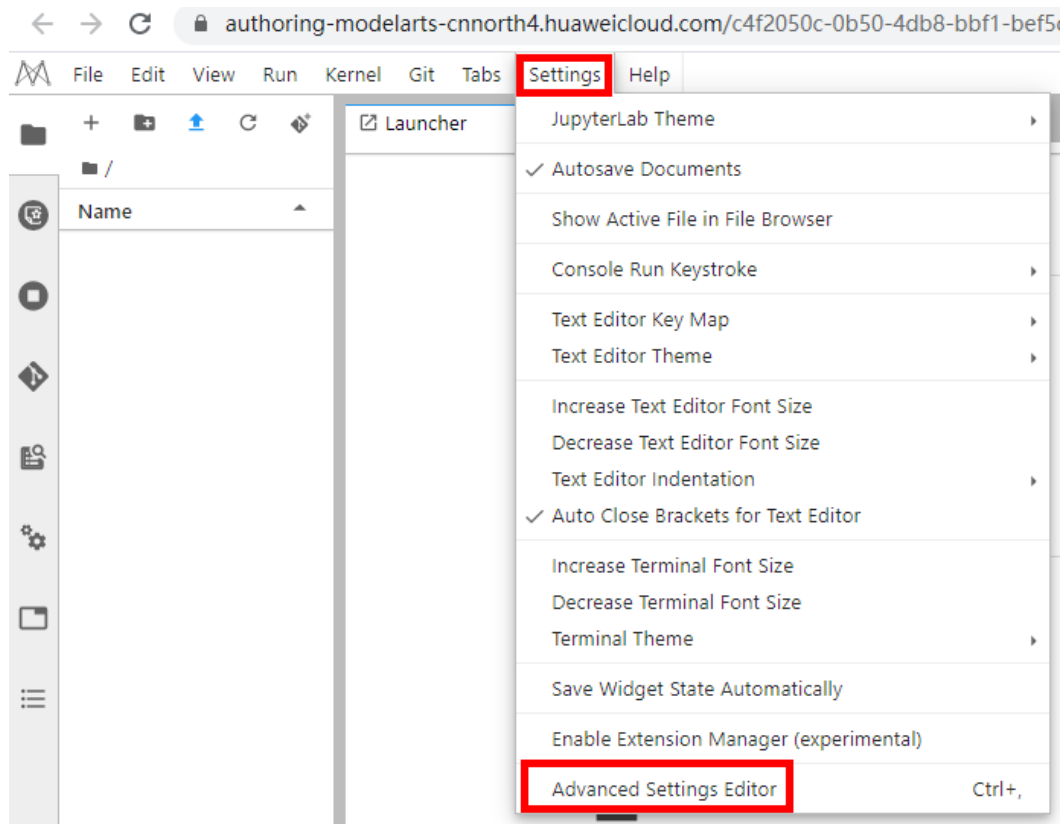
Solution

Set **fontFamily** of **Terminal** to **Menlo**.

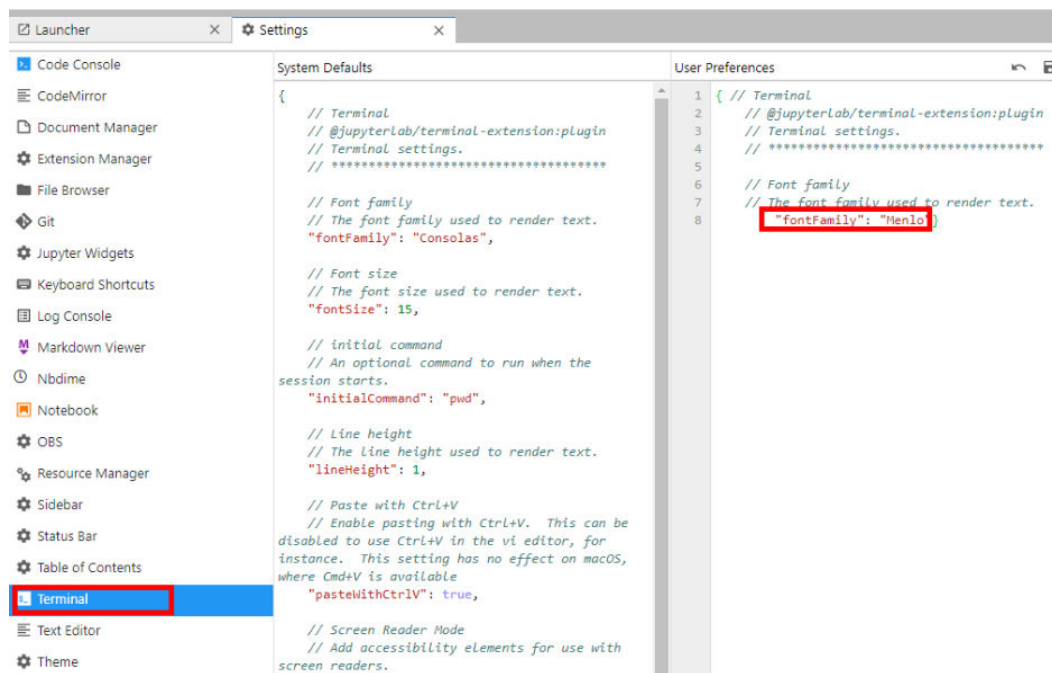
Procedure

- Step 1** Log in to the ModelArts management console and choose **DevEnviron > Notebook**.
- Step 2** Locate the row containing the target notebook instance and click **Open** in the **Operation** column. The **JupyterLab** page is displayed.

Step 3 On the **JupyterLab** page, choose **Settings > Advanced Settings Editor**. The **Settings** tab page is displayed.



Step 4 Choose **Terminal** in the navigation pane on the left and set **fontFamily** to **Menlo**.



----End

3.5 Notebook Instances

3.5.1 What Do I Do If I Cannot Open a Notebook Instance I Created?

If you cannot open your created notebook instance, refer to the following error codes for troubleshooting.

404 Error

If this error is reported when an IAM user creates an instance, the IAM user does not have the permissions to access the corresponding storage location (OBS bucket).

Solution

1. Log in to the OBS console using the primary account and grant access permissions for the OBS bucket to the IAM user. For details about the operation, see .
2. After the IAM user obtains the permissions, log in to the ModelArts console, delete the instance, and use the OBS path to create a notebook instance.

503 Error

If this error is reported, it is possible that the instance is consuming too many resources. If this is the case, stop the instance and restart it.

3.5.2 What Should I Do When the System Displays an Error Message Indicating that No Space Left After I Run the pip install Command?

Symptom

In the notebook instance, error message "No Space left..." is displayed after the **pip install** command is run.

Solution

You are advised to run the **pip install --no-cache **** command instead of the **pip install **** command. Adding the **--no-cache** parameter can solve such problem.

3.5.3 What Do I Do If "Read timed out" Is Displayed After I Run pip install?

Symptom

After I run **pip install** in a notebook instance, the system displays error message "ReadTimeoutError..." or "Read timed out..."

```
sh-4.3$ pip install torch==1.7.0 torchvision==0.8.0 torchaudio==0.7.0 matplotlib pyyaml tqdm sklearn h5py tensorboard pandas
Looking in indexes: http://pip-notebook.modelarts.com:8888/repository/pypi/simple/
Collecting torch==1.7.0
  WARNING: Retrying (Retry(total=4, connect=None, read=None, redirect=None, status=None)) after connection broken by 'ReadTimeou
tError("HTTPConnectionPool(host='pip-notebook.modelarts.com', port=8888): Read timed out. (read timeout=15)")': /repository/pypi
/packages/torch/1.7.0/torch-1.7.0-cp37-cp37m-manylinux1_x86_64.whl
  WARNING: Retrying (Retry(total=3, connect=None, read=None, redirect=None, status=None)) after connection broken by 'ReadTimeou
tError("HTTPConnectionPool(host='pip-notebook.modelarts.com', port=8888): Read timed out. (read timeout=15)")': /repository/pypi
/packages/torch/1.7.0/torch-1.7.0-cp37-cp37m-manylinux1_x86_64.whl
  WARNING: Retrying (Retry(total=2, connect=None, read=None, redirect=None, status=None)) after connection broken by 'ReadTimeou
tError("HTTPConnectionPool(host='pip-notebook.modelarts.com', port=8888): Read timed out. (read timeout=15)")': /repository/pypi
/packages/torch/1.7.0/torch-1.7.0-cp37-cp37m-manylinux1_x86_64.whl
  WARNING: Retrying (Retry(total=1, connect=None, read=None, redirect=None, status=None)) after connection broken by 'ReadTimeou
tError("HTTPConnectionPool(host='pip-notebook.modelarts.com', port=8888): Read timed out. (read timeout=15)")': /repository/pypi
/packages/torch/1.7.0/torch-1.7.0-cp37-cp37m-manylinux1_x86_64.whl
^CERROR: operation cancelled by user
WARNING: You are using pip version 21.0.1; however, version 21.1.2 is available.
You should consider upgrading via the '/opt/conda/bin/python -m pip install --upgrade pip' command.
```

Solution

Run `pip install --upgrade pip` and then `pip install`.

3.5.4 What Do I Do If the Code Can Be Run But Cannot Be Saved, and the Error Message "save error" Is Displayed?

If the notebook instance can run the code but cannot save it, the error message "save error" is displayed when you save the file. In most cases, this error is caused by a security policy of Web Application Firewall (WAF).

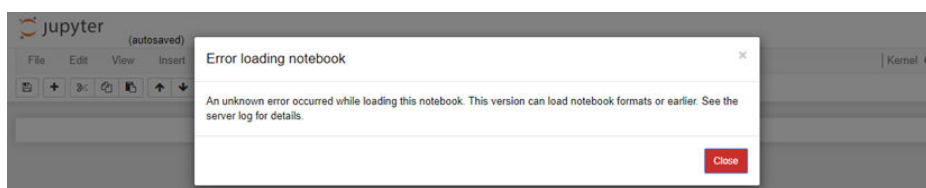
On the current page, some characters in your input or output of the code are intercepted, because they are considered to be a security risk. Submit a service ticket and contact system administrator to check and handle the problem.

3.5.5 What Do I Do If "Error loading notebook" Is Displayed When I Create an IPYNB File?

Symptom

On the Jupyter page, "Error loading notebook" is displayed when an IPYNB file is created.

Figure 3-6 Error message



Possible Cause

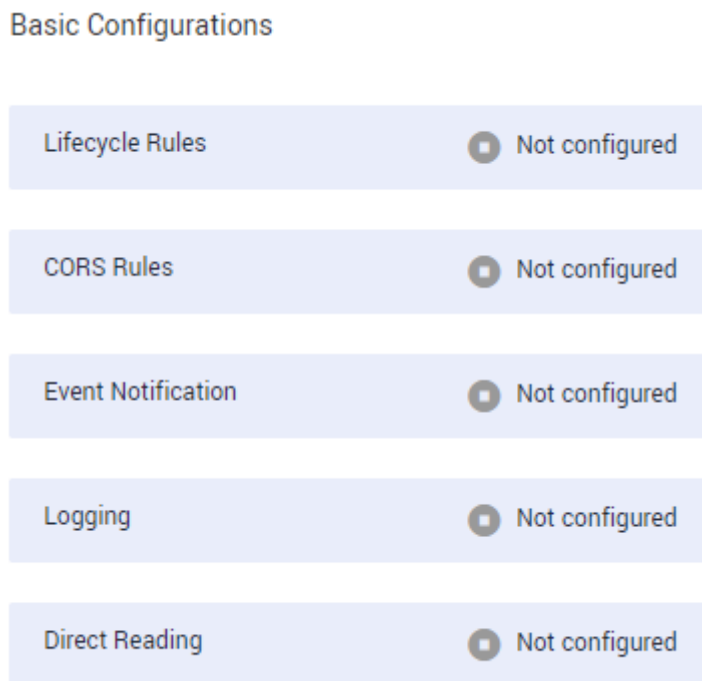
The root cause of this error is the attributes of the OBS bucket that you use to create a notebook instance. If the storage class of the OBS bucket is **Archive** and the **Direct Reading** function is disabled, the IPYNB file fails to be created in ModelArts Notebook.

Solution

Log in to OBS Console, select the bucket corresponding to the notebook instance, and click the bucket name to go to the bucket details page. In the **Basic Configurations** area, locate the row of **Direct Reading**, click it, and select **Enable**.

in the dialog box that is displayed. After the configuration is completed, the **IPYNB** file can be created for the notebook instance.

Figure 3-7 Enabling Direct Reading



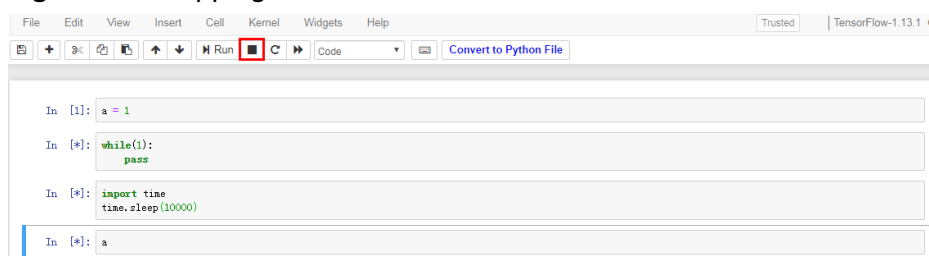
3.6 Code Execution

3.6.1 What Do I Do If a Notebook Instance Won't Run My Code?

If a notebook instance fails to execute code, you can locate and rectify the fault as follows:

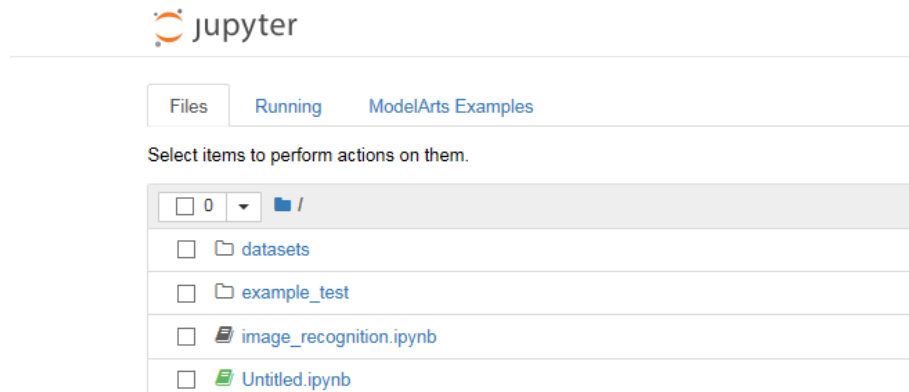
1. If the execution of a cell is suspended or lasts for a long time (for example, the execution of the second and third cells in [Figure 3-8](#) is suspended or lasts for a long time, causing execution failure of the fourth cell) but the notebook page still responds and other cells can be selected, click **interrupt the kernel** highlighted in a red box in the following figure to stop the execution of all cells. The notebook instance retains all variable spaces.

Figure 3-8 Stopping all cells



2. If the notebook page does not respond, close the notebook page and the ModelArts console. Then, open the ModelArts console and access the notebook instance again. The notebook instance retains all the variable spaces that exist when the notebook instance is unavailable.

Figure 3-9 Accessing the notebook instance again



3. If the notebook instance still cannot be used, access the **Notebook** page on the ModelArts console and stop the notebook instance. After the notebook instance is stopped, click **Start** to restart the notebook instance and open it. The instance will have preserved all the spaces for the variables that were unable to run.

3.6.2 Why Does the Instance Break Down When dead kernel Is Displayed During Training Code Running?

The notebook instance breaks down during training code running due to insufficient memory caused by large data volume or excessive training layers.

After this error occurs, the system automatically restarts the notebook instance to fix the instance breakdown. In this case, only the breakdown is fixed. If you run the training code again, the failure will still occur. To solve the problem of insufficient memory, you are advised to create a new notebook instance and use a resource pool of higher specifications, such as a GPU or dedicated resource pool, to run the training code. An existing notebook instance that has been successfully created cannot be scaled up using resources with higher specifications.

3.6.3 What Do I Do If cudaCheckError Occurs During Training?

Symptom

The following error occurs when the training code is executed in a notebook:

```
cudaCheckError() failed : no kernel image is available for execution on the device
```

Possible Cause

Parameters **arch** and **code** in **setup.py** have not been set to match the GPU compute power.

Solution

For Tesla V100 GPUs, the GPU compute power is **-gencode arch=compute_70,code=[sm_70,compute_70]**. Set the compilation parameters in **setup.py** accordingly.

3.6.4 What Should I Do If DevEnviron Prompts Insufficient Space?

If space is insufficient, you are advised to use notebook instances of the EVS type.

Upload code and data to an OBS bucket for the original notebook instance by referring to [How Do I Upload a File from a Notebook Instance to OBS or Download a File from OBS to a Notebook Instance?](#). Then, create a notebook instance of the EVS type, and download files from OBS to the new notebook instance.

3.6.5 Why Does the Notebook Instance Break Down When `opencv.imshow` Is Used?

Symptom

When `opencv.imshow` is used in a notebook instance, the notebook instance breaks down.

Possible Causes

The `cv2.imshow` function in OpenCV malfunctions in a client/server environment such as Jupyter. However, Matplotlib does not have this problem.

Solution

Display images by referring to the following example. Note that OpenCV displays BGR images while Matplotlib displays RGB images.

Python:

```
from matplotlib import pyplot as plt
import cv2
img = cv2.imread('Image path')
plt.imshow(cv2.cvtColor(img, cv2.COLOR_BGR2RGB))
plt.title('my picture')
plt.show()
```

3.6.6 Why Cannot the Path of a Text File Generated in Windows OS Be Found In a Notebook Instance?

Symptom

When a text file generated in Windows is used in a notebook instance, the text content cannot be read and an error message may be displayed indicating that the path cannot be found.

Possible Causes

The notebook instance runs Linux and its line feed format (CRLF) differs from that (LF) in Windows.

Solution

Convert the file format to Linux in your notebook instance.

Shell:

```
dos2unix File name
```

3.6.7 What Do I Do If Files Fail to Be Saved in JupyterLab?

Symptom

When a file is saved in JupyterLab, an error message is displayed.

File Save Error for rebar_count.ipynb

Failed to fetch

Dismiss

Possible Cause

A third-party plug-in has been installed on the browser, and the proxy intercepts the request. As a result, the file cannot be saved.

Solution

Disable the plug-in and save file again.

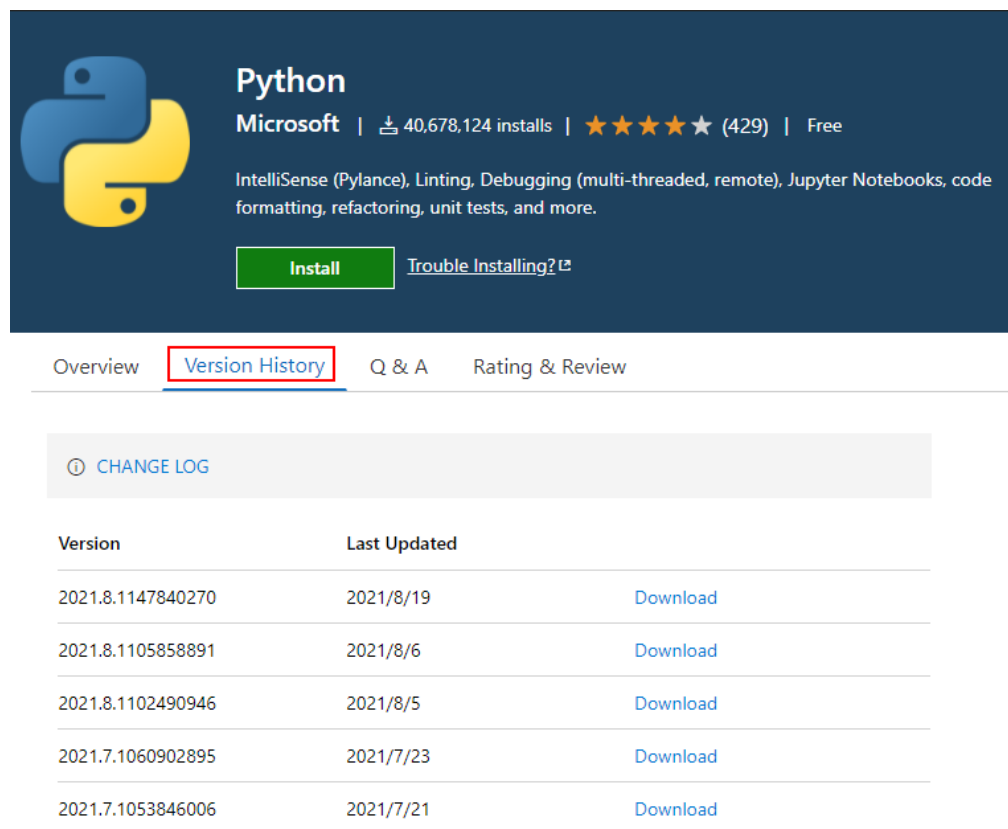
3.7 VS Code

3.7.1 What Do I Do If Installing a Remote Plug-in Failed?

Method 1 (recommended): Using an offline package

1. Log in at the [official VS Code website](#) and search for the target Python plug-in.
2. Click the **Version History** tab of the plug-in and download the offline installation package.

Figure 3-10 Offline installation package of the Python plug-in



3. In local VS Code, drag the downloaded VSIX file to the remote notebook.
4. Right-click the file and choose **Install Extension VSIX** from the shortcut menu.

Method 2: Setting the default remote plug-in

Set the default remote plug-in in VS Code by following the instructions provided in [How Can I Set the Default Remote Plug-in in VS Code?](#) This enables automatic plug-in installation when the plug-in is connected.

Method 3: Taking measures provided at [official VS Code website](#)

Tips (adjust parameter settings as needed):

```
"remote.SSH.connectTimeout": 10,
"remote.SSH.maxReconnectionAttempts": null,
"remote.downloadExtensionsLocally": true,
"remote.SSH.useLocalServer": false,
"remote.SSH.localServerDownload": "always",
```

3.7.2 What Do I Do If a Restarted Notebook Instance Can Be Connected Only After I Locally Delete known_hosts?

To resolve this issue, set notebook parameters **StrictHostKeyChecking no** and **UserKnownHostsFile=/dev/null** in the local **ssh config** file.

```
Host roma-local-cpu
  HostName x.x.x.x # IP address
  Port 22522
  User ma-user
```

```
IdentityFile C:/Users/my.pem
StrictHostKeyChecking no
ForwardAgent yes
```

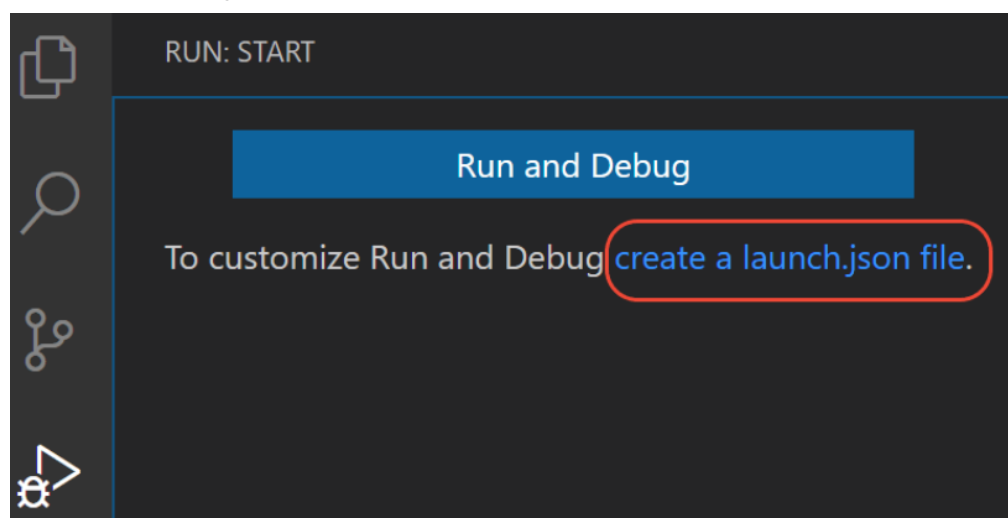
Note: SSH logins are insecure because the `known_hosts` file will be ignored during the logins.

3.7.3 What Do I Do If the Source Code Cannot Be Accessed When I Use VS Code for Debugging?

If the `launch.json` file already exists, go to step 3.

Step 1: Open `launch.json`.

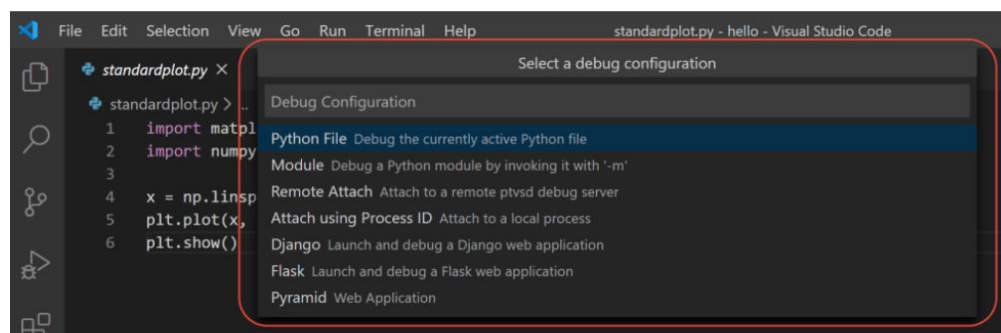
- Method 1: Click **Run (Ctrl+Shift+D)** in the menu bar on the left and click **create a launch.json file**.



- Method 2: In the menu bar, choose **Run > Open configurations**.

Step 2: Select a language.

To set a Python language, select **Python File** in **Select a debug configuration**. The operations for setting other languages are similar.



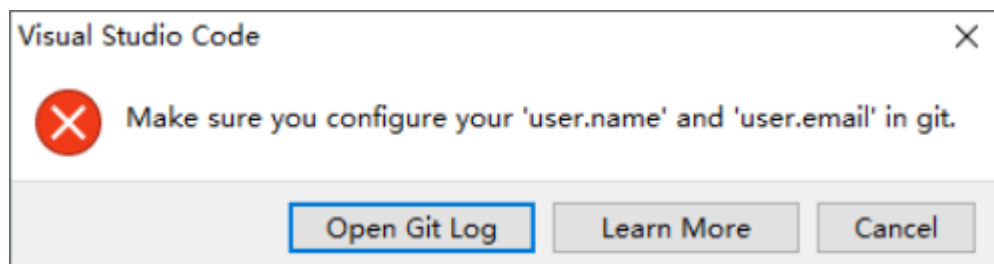
Step 3: Set `justMyCode` to `False` in `launch.json`.

```
{
  // Use IntelliSense to learn about possible attributes.
  // Hover to view descriptions of existing attributes.

  "version": "0.2.0",
  "configurations": [
```

```
{
  "name": "Python: Current file",
  "type": "python",
  "request": "launch",
  "program": "${file}",
  "console": "integratedTerminal",
  "justMyCode": false
}
```

3.7.4 What Do I Do If a Message Is Displayed Indicating an Incorrect Username or Email Address When I Use VS Code to Submit Code?



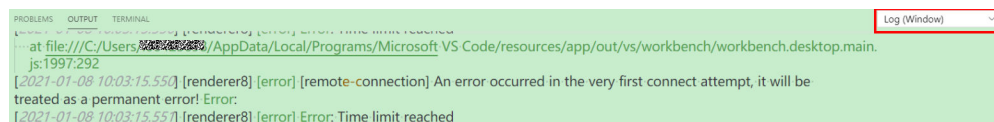
1. In VS Code, press **Ctrl+Shift+P**.
2. Search for **Python: Select Interpreter** and select the target Python.
3. Choose **Terminal > New Terminal**. The CLI of the remote container is displayed.
4. On the VS Code terminal, run the following commands and submit the code again:


```
git config --global user.email xxx@xxx.com # Change the email address to yours.
git config --global user.name xxx # Change the username to yours.
```

3.7.5 How Can I View Remote Logs in VS Code?

1. In VS Code, press **Ctrl+Shift+P**.
2. Search for **show logs**.
3. Choose **Remote Server**.

Alternatively, switch logs in the red box shown in the following figure.



3.7.6 How Can I Open the VS Code Configuration File settings.json?

1. In VS Code, press **Ctrl+Shift+P**.
2. Search for **Open Settings (JSON)**.

3.7.7 How Can I Change the VS Code Background Color to Light Green?

Add the following settings to the VS Code configuration file **settings.json**:

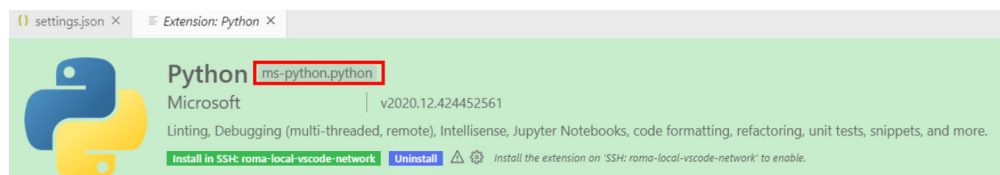
```
"workbench.colorTheme": "Atom One Light",
"workbench.colorCustomizations": {
  "[Atom One Light]": {
    "editor.background": "#C7EDCC",
    "sideBar.background": "#e7f0e7",
    "activityBar.background": "#C7EDCC",
  },
},
```

3.7.8 How Can I Set the Default Remote Plug-in in VS Code?

Add **remote.SSH.defaultExtensions**, for example, for automatically installing Python and Maven plug-ins, to the VS Code configuration file **settings.json**.

```
"remote.SSH.defaultExtensions": [
  "ms-python.python",
  "vscjava.vscode-maven"
],
```

To obtain a plug-in name, click the plug-in in VS Code.



3.7.9 How Can I Install a Local Plug-in on the Remote End or a Remote Plug-in on the Local End Through VS Code?

1. In VS Code, press **Ctrl+Shift+P**.
2. Search for **install local** and select the plug-in as required.

3.8 Failures to Access the Development Environment Through VS Code

3.8.1 When Do I Do If the VS Code Window Is Not Displayed?

Possible cause: The VS Code is not installed or the installed version is outdated.

Solution: Download and install the VS Code.



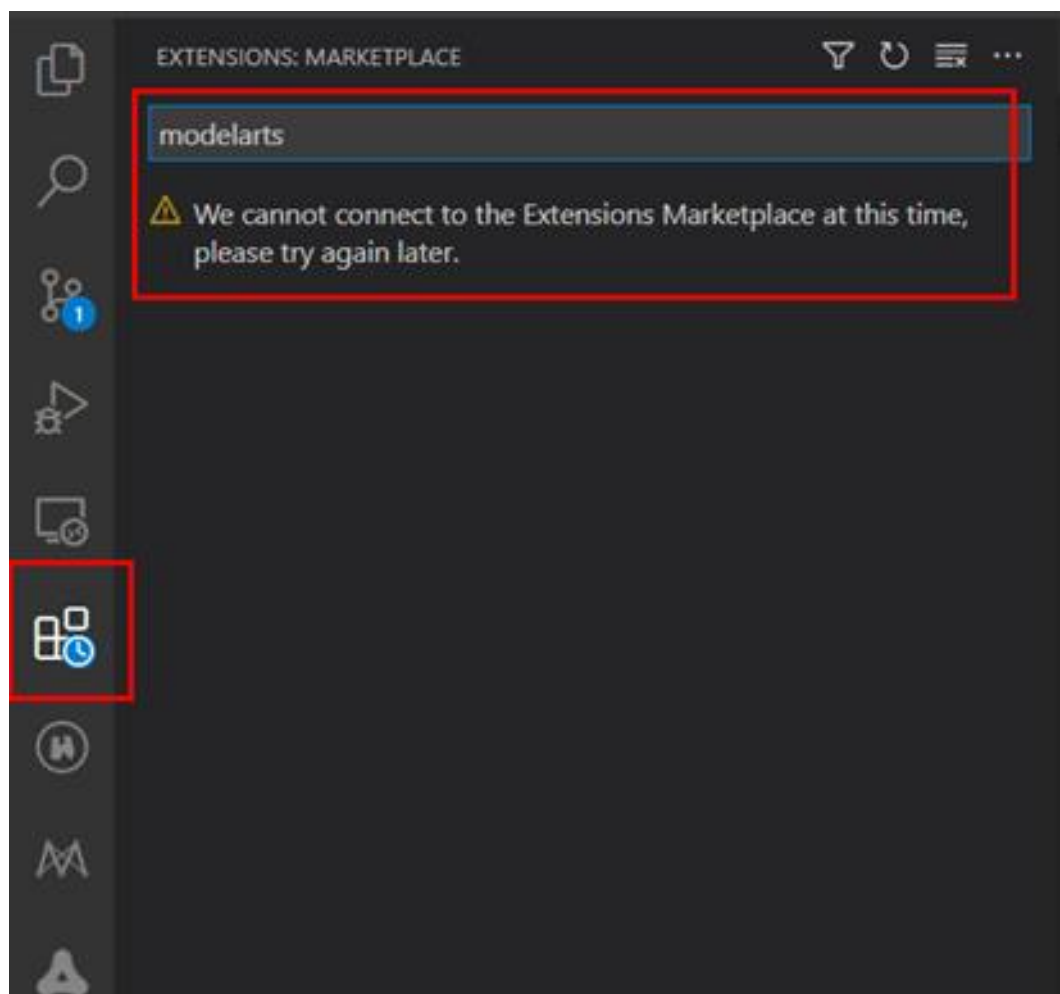
3.8.2 What Do I Do If a Remote Connection Failed After VS Code Is Opened?

NOTICE

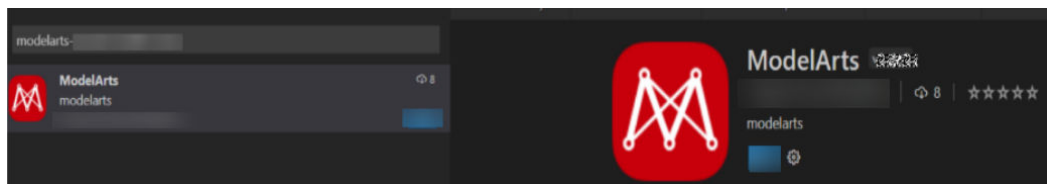
If your local PC runs Linux, see possible cause 2.

Possible cause 1: Automatically installing the VS Code plug-in ModelArts failed.

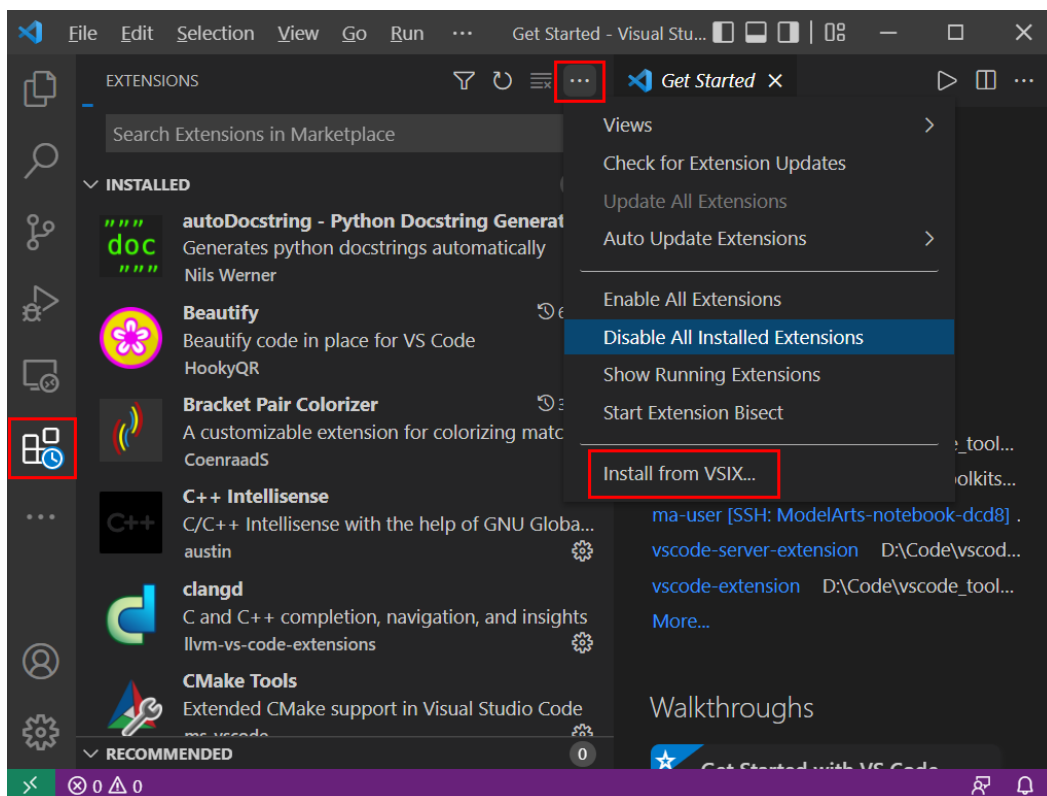
Solution 1: Verify that the VS Code network is accessible. Search for **ModelArts** in the VS Code marketplace. If the following information is displayed, a network error occurred. In this case, switch to another proxy or use another network.



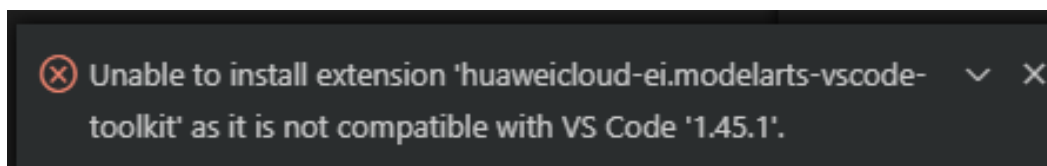
Search for **ModelArts** again. If the following information is displayed, the network is normal. Then, switch back to the ModelArts management console and try to access VS Code again.



Solution 2: If the VS Code marketplace cannot be accessed, manually download and install ModelArts and [Remote-SSH](#).

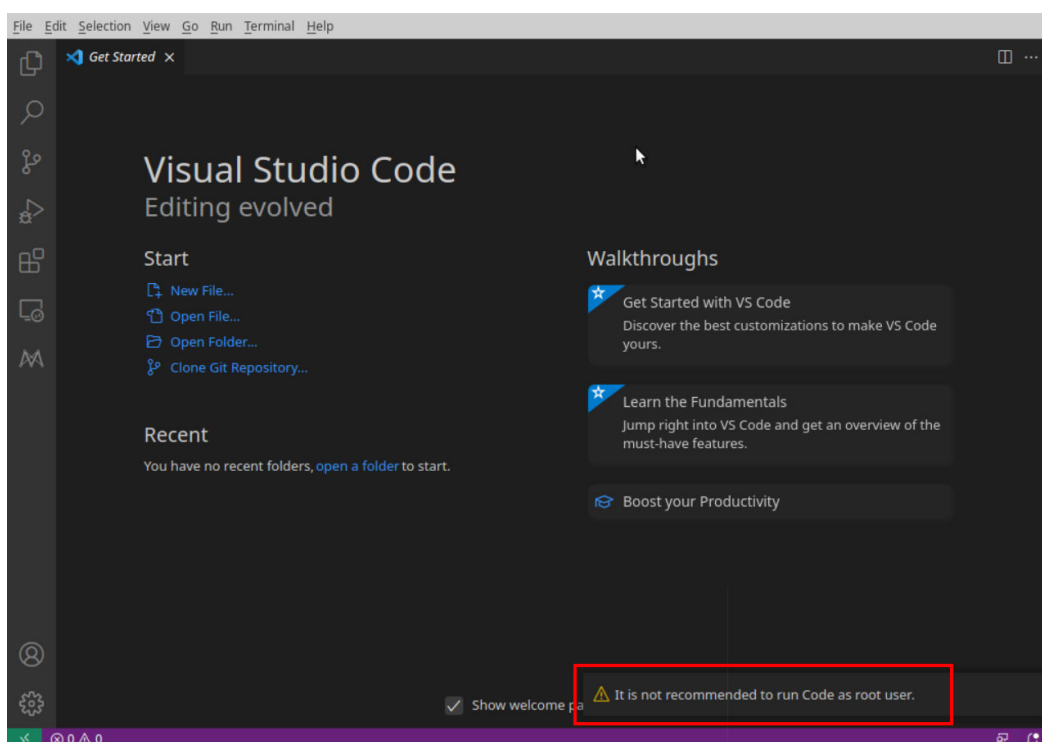


Solution 3: If the error message shown in the following figure is displayed, the VS Code version is outdated. Upgrade the VS Code to 1.57.1 or the latest version.



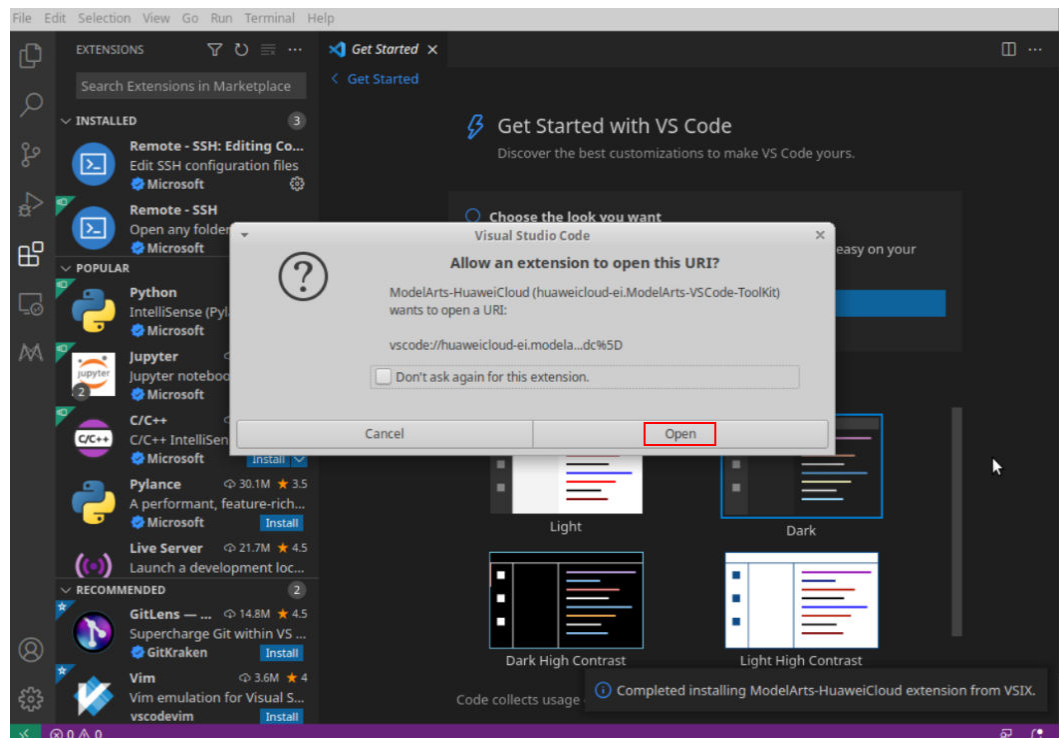
Possible cause 2: The local PC runs Linux, and VS Code is installed as user **root**. When you access VS Code, the information "It is not recommended to run Code as root user" is displayed.

```
root@ecs-.../VSCode# sudo dpkg -i code_1.67.2-1652812855_amd64.deb
Selecting previously unselected package code.
(Reading database ... 199224 files and directories currently installed.)
Preparing to unpack code_1.67.2-1652812855_amd64.deb ...
Unpacking code (1.67.2-1652812855) ...
Setting up code (1.67.2-1652812855) ...
Processing triggers for gnome-menus (3.13.3-11ubuntu1.1) ...
Processing triggers for desktop-file-utils (0.23-1ubuntu3.18.04.2) ...
Processing triggers for mime-support (3.60ubuntu1) ...
Processing triggers for shared-mime-info (1.9-2) ...
root@ecs-.../VSCode# code
You are trying to start Visual Studio Code as a super user which isn't recommended. If this was intended, please add the argument '--no-sandbox' and specify an alternate user data directory using the '--user-data-dir' argument.
root@ecs-dctest:/dongcong/VSCode# code
You are trying to start Visual Studio Code as a super user which isn't recommended. If this was intended, please add the argument '--no-sandbox' and specify an
```



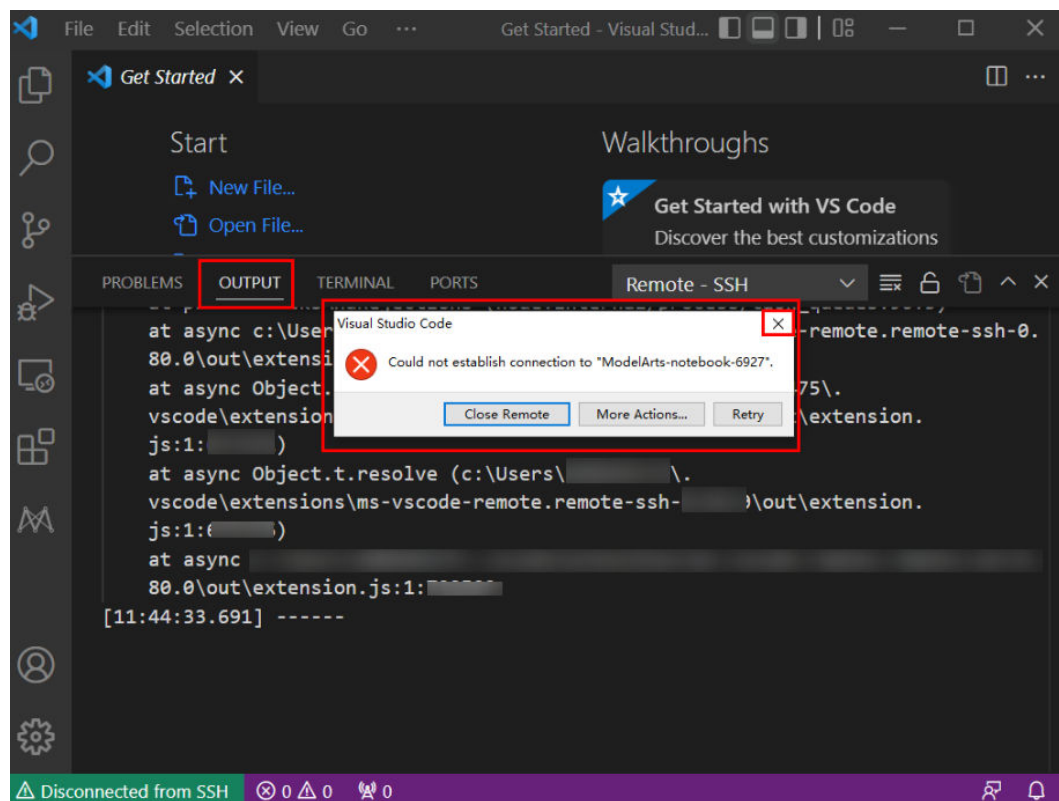
Solution: Install VS Code as a non-root user, return to the ModelArts management console, and click **Access VS Code**.

```
~/VSCode$ sudo dpkg -i code_1.67.2-1652812855_amd64.deb
[sudo] password for dc:
(Reading database ... 200705 files and directories currently installed.)
Preparing to unpack code_1.67.2-1652812855_amd64.deb ...
Unpacking code (1.67.2-1652812855) over (1.67.2-1652812855) ...
Setting up code (1.67.2-1652812855) ...
Processing triggers for gnome-menus (3.13.3-11ubuntu1.1) ...
Processing triggers for desktop-file-utils (0.23-1ubuntu3.18.04.2) ...
Processing triggers for mime-support (3.60ubuntu1) ...
Processing triggers for shared-mime-info (1.9-2) ...
~/VSCode$ code
```

3.8.3 What Do I Do If Error Message "Could not establish connection to xxx" Is Displayed During a Remote Connection?

Symptom

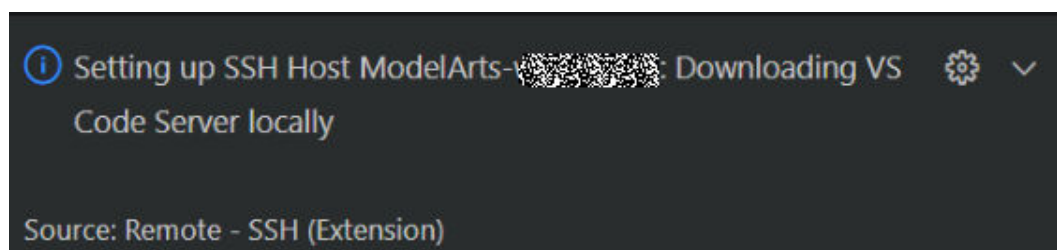


Possible cause: Establishing the remote SSH connection to the instance through VS Code failed.

Solution: Close the displayed dialog box, view the error information in **OUTPUT**, and rectify the fault by referring to the troubleshooting methods provided in the following sections.

3.8.4 What Do I Do If the Connection to a Remote Development Environment Remains in "Setting up SSH Host xxx: Downloading VS Code Server locally" State for More Than 10 Minutes?

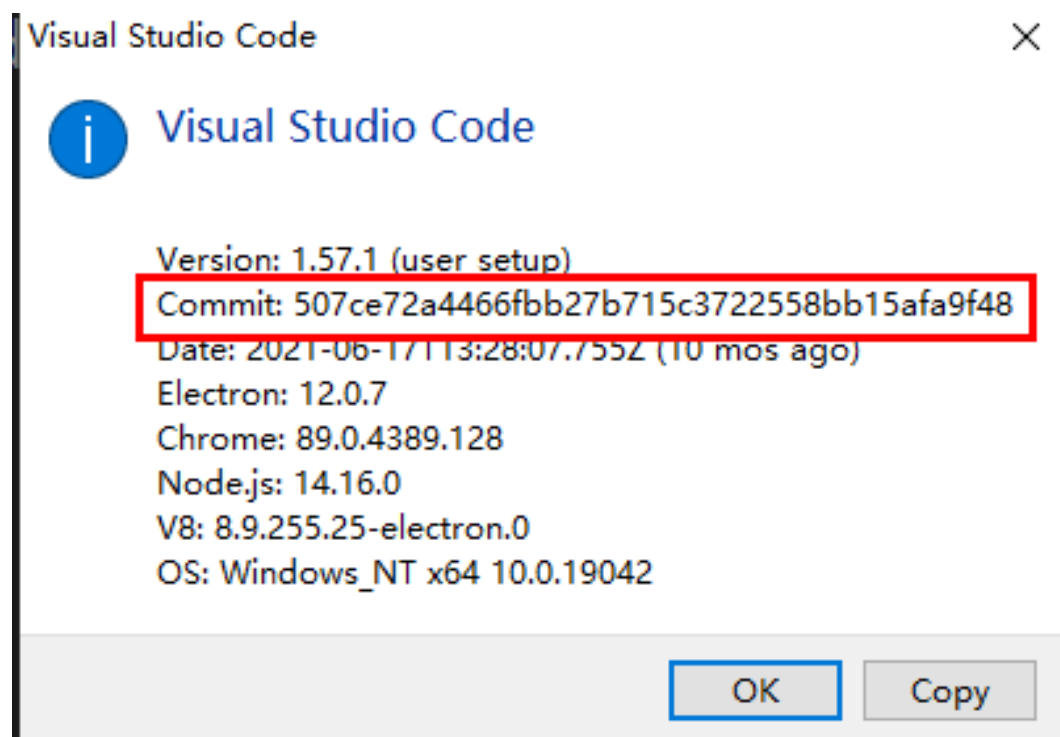
Symptom



Possible cause: The local network is faulty. As a result, it takes a long time to automatically install the VS Code server remotely.

Solution: Manually install the VS Code server.

1. Obtain the VS Code commit ID.

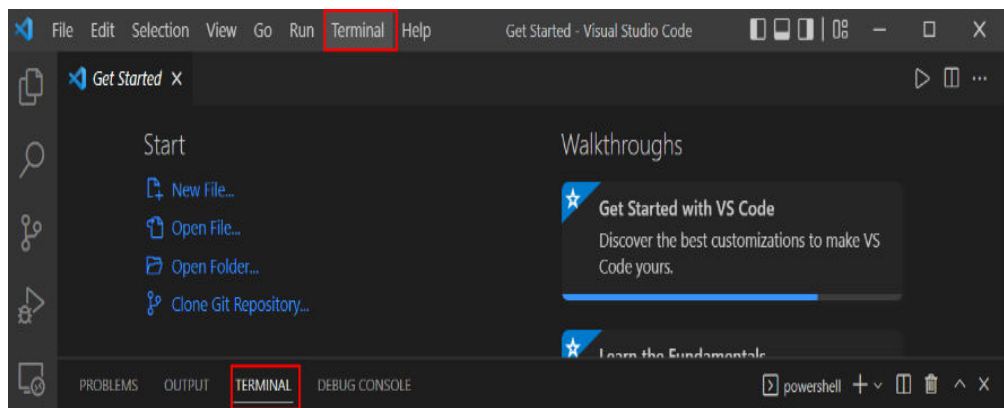


2. Download the VS Code server package of the required version. Select Arm or x86 based on the CPU architecture of the development environment.

 **NOTE**

Replace $\${commitID}$ in the following link with the commit ID obtained in **1**:

- For Arm, download **vscode-server-linux-arm64.tar.gz**.
[https://update.code.visualstudio.com/commit:\\${commitID}/server-linux-arm64/stable](https://update.code.visualstudio.com/commit:${commitID}/server-linux-arm64/stable)
 - For x86, download **vscode-server-linux-x64.tar.gz**.
[https://update.code.visualstudio.com/commit:\\${commitID}/server-linux-x64/stable](https://update.code.visualstudio.com/commit:${commitID}/server-linux-x64/stable)
3. Access the remote environment.
 Switch to **Terminal** in VS Code.



Run the following command in VS Code Terminal to access the remote development environment:

```
ssh -tt -o StrictHostKeyChecking=no -i ${IdentityFile} ${User}@${HostName} -p ${Port}
```

Parameters:

- **IdentityFile**: Path to the local key
- **User**: Username, for example, **ma-user**
- **HostName**: IP address
- **Port**: Port number

Name	notebook-4002	Flavor	modelarts.vm.cpu.2u
Status	Stopped	Image	pytorch1.4-cuda10.1-cudnn7-ubuntu18.04
ID	dfc45125-4258-4564-b178-4865343815a7	Created At	May 18, 2022 16:19:08 GMT+08:00
Storage Path	/home/ma-user/work/	Updated At	May 18, 2022 18:33:53 GMT+08:00
Storage Capacity	50 GB (Default)	Address	ssh://mauser@xxx.modelarts.com:31205
Whitelist	--	Authentication	KeyPair-9559

4. Manually install the VS Code server.

Run the following commands on the VS Code Terminal to clear the residual data (replace $\${commitID}$ in the commands with the commit ID obtained in **1**):

```
rm -rf /home/ma-user/.vscode-server/bin/${commitID}/*
mkdir -p /home/ma-user/.vscode-server/bin/${commitID}
```

Upload the VS Code server package to the development environment.

```
exit
scp -i xxx.pem -P 31205 Local path to the VS Code server package ma-user@xxx:/home/ma-user/.vscode-server/bin
ssh -tt -o StrictHostKeyChecking=no -i ${IdentityFile} ${User}@${HostName} -p ${Port}
```

Parameters:

- **IdentityFile:** Path to the local key
- **User:** Username, for example, **ma-user**
- **HostName:** IP address
- **Port:** Port number

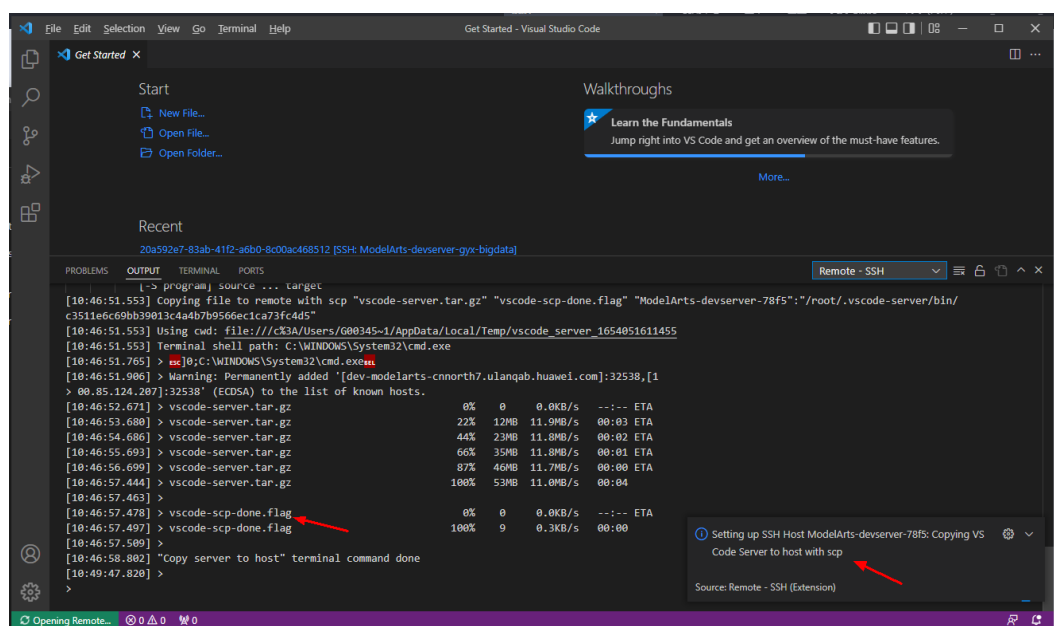
Take Arm as an example. Decompress the VS Code server package to **\$HOME/.vscode-server/bin**. Replace *`\${commitID}* in the command with the commit ID obtained in 1.

```
cd /home/ma-user/.vscode-server/bin
tar -zxf vscode-server-linux-arm64.tar.gz
mv vscode-server-linux-arm64/* ${commitID}
```

5. Establish the remote connection again.

3.8.5 What Do I Do If the Connection to a Remote Development Environment Remains in the State of "Setting up SSH Host xxx: Downloading VS Code Server locally" for More Than 10 Minutes?

Symptom

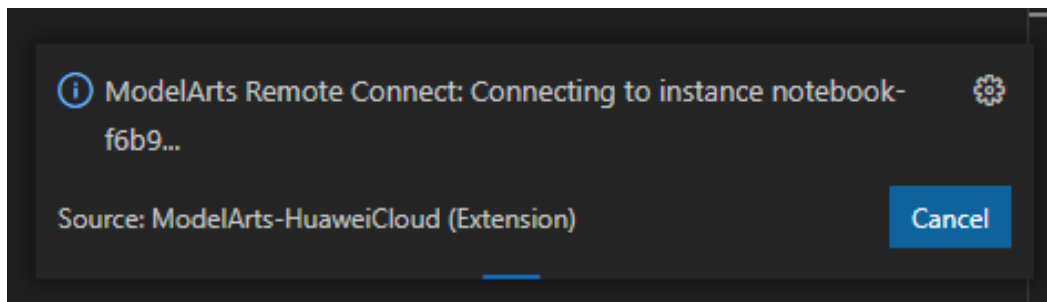


Possible cause: The log shows that **vscode-scp-done.flag** has been locally uploaded, but it is not received on the remote end.

Solution: Close all VS Code windows, return to the ModelArts management console, and click **Access VS Code**.

3.8.6 What Do I Do If the Connection to a Remote Development Environment Remains in the State of "ModelArts Remote Connect: Connecting to instance xxx..." for More Than 10 Minutes?

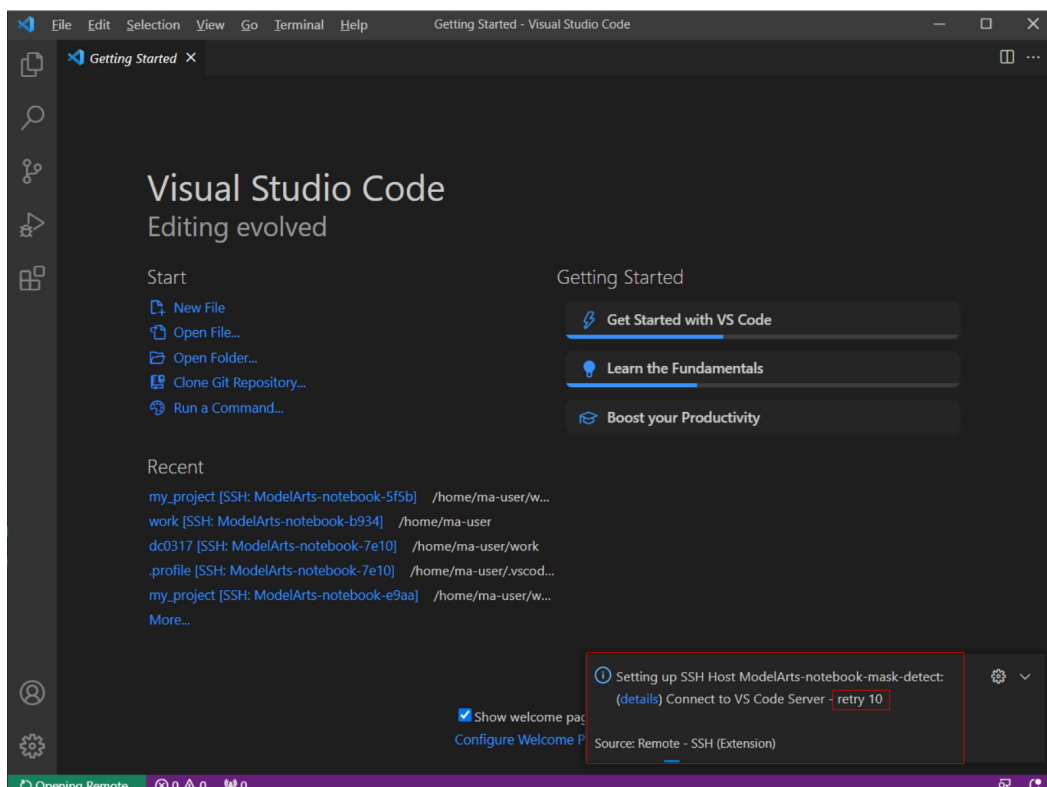
Symptom



Solution: Click **Cancel**, return to the ModelArts management console, and click **Access VS Code**.

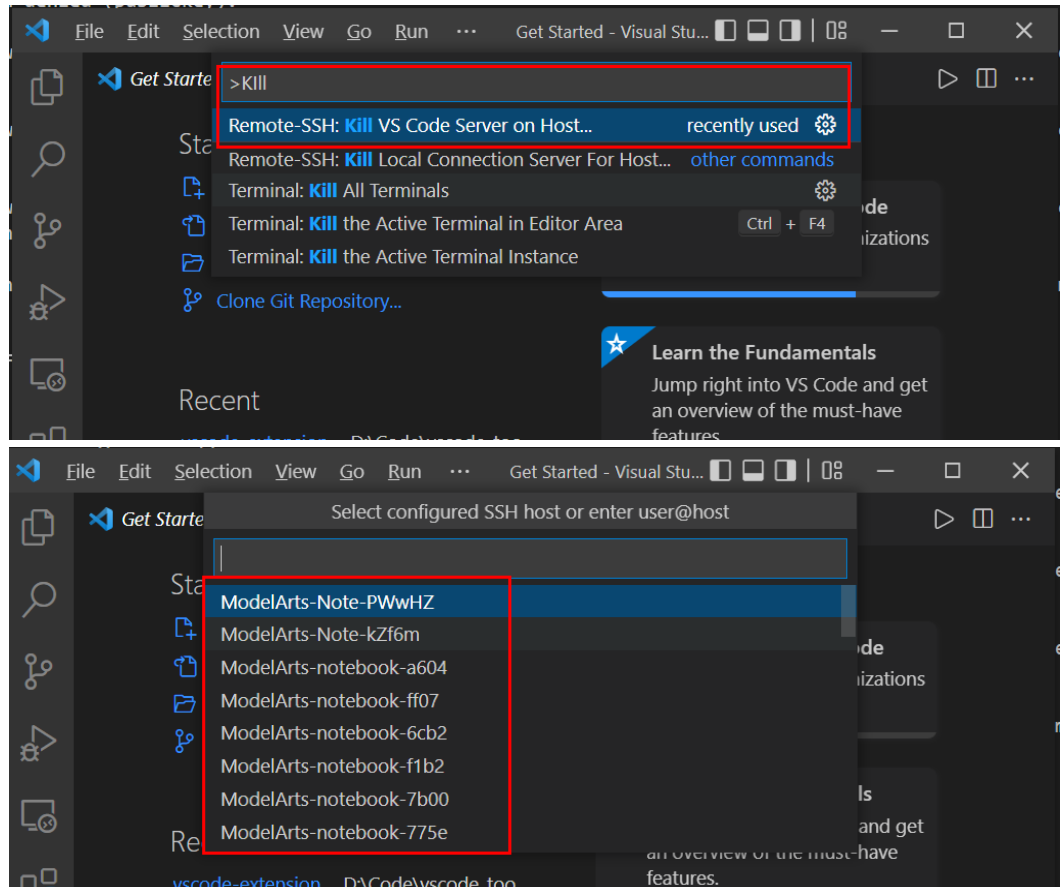
3.8.7 What Do I Do If a Remote Connection Is in the Retry State?

Symptom



Possible cause: Downloading the VS Code server failed before, leading to residual data. As a result, new download cannot be performed.

Solution 1 (performed locally): Open the command panel (**Ctrl+Shift+P** for Windows and **Cmd+Shift+P** for Mac), search for **Kill VS Code Server on Host**, and locate the affected instance, which will be automatically cleared. Then, establish the connection again.



Solution 2 (performed remotely): Delete the files that are being used in /home/ma-user/.vscode-server/bin/ on the VS Code Terminal. Then, establish the connection again.

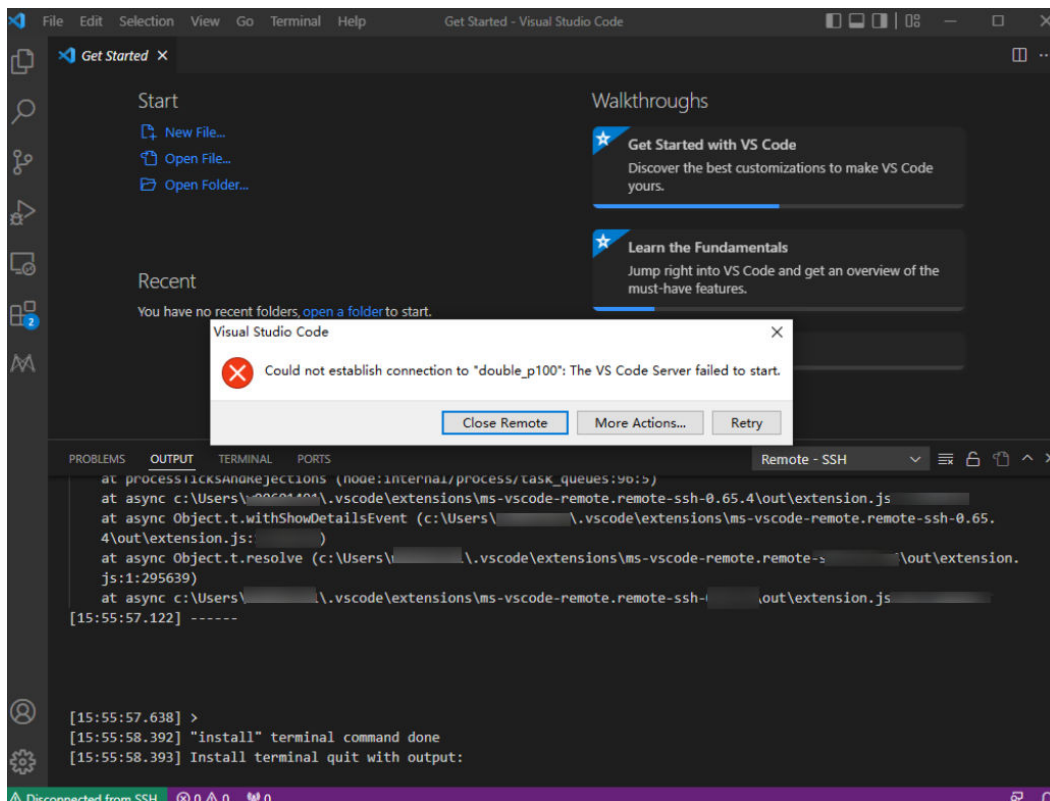
```
ssh -tt -o StrictHostKeyChecking=no -i ${IdentityFile} ${User}@${HostName} -p ${Port}
rm -rf /home/ma-user/.vscode-server/bin/
```

Parameters:

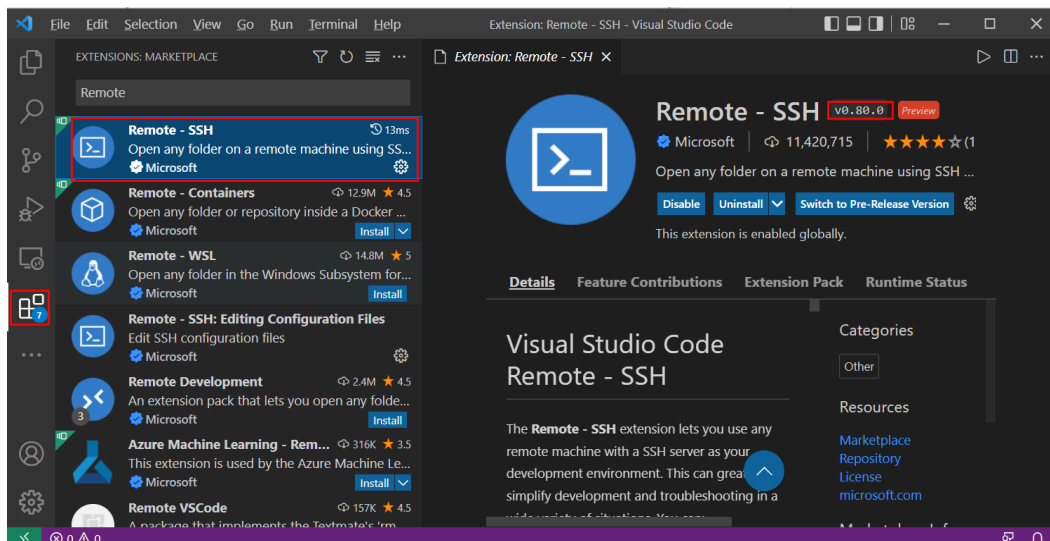
- **IdentityFile**: Path to the local key
- **User**: Username, for example, **ma-user**
- **HostName**: IP address
- **Port**: Port number

3.8.8 What Do I Do If Error Message "The VS Code Server failed to start" Is Displayed?

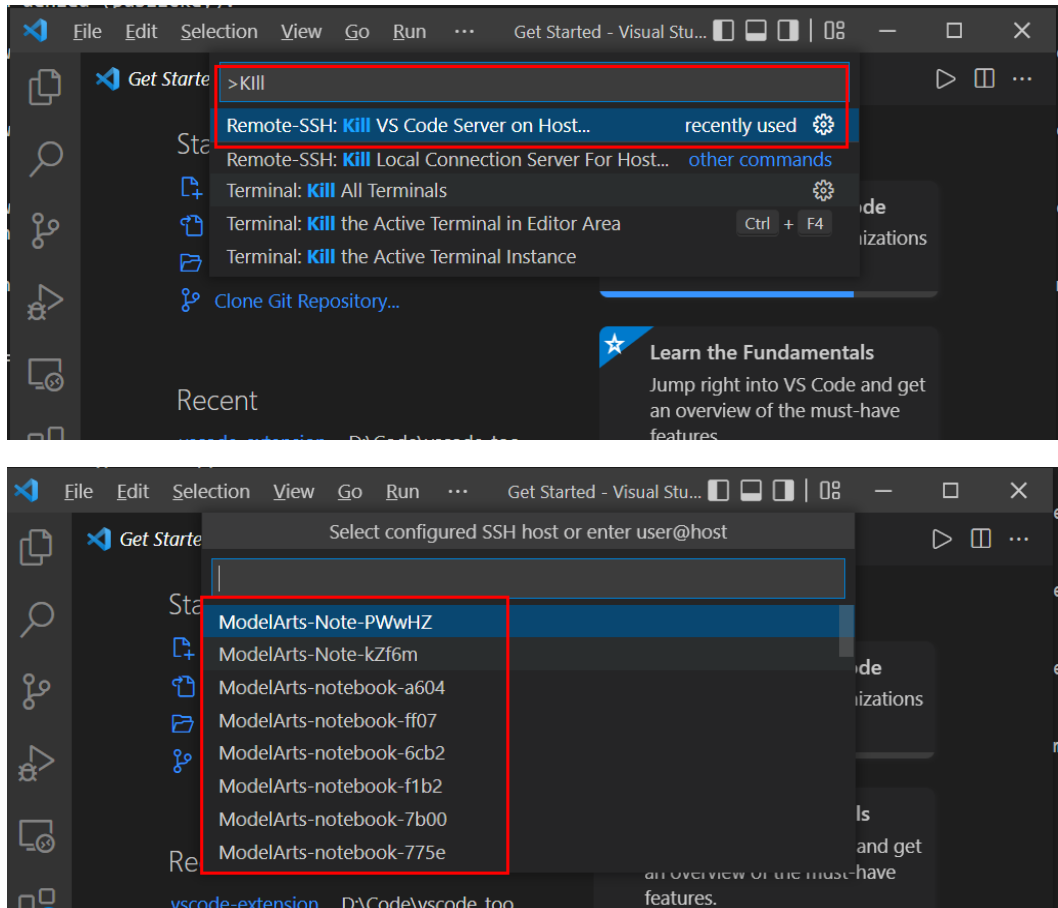
Symptom



Solution: 1. Check whether the VS Code version is 1.65.0 or later. If so, check the Remote-SSH version. If the Remote-SSH version is earlier than 0.76.1, upgrade Remote-SSH.

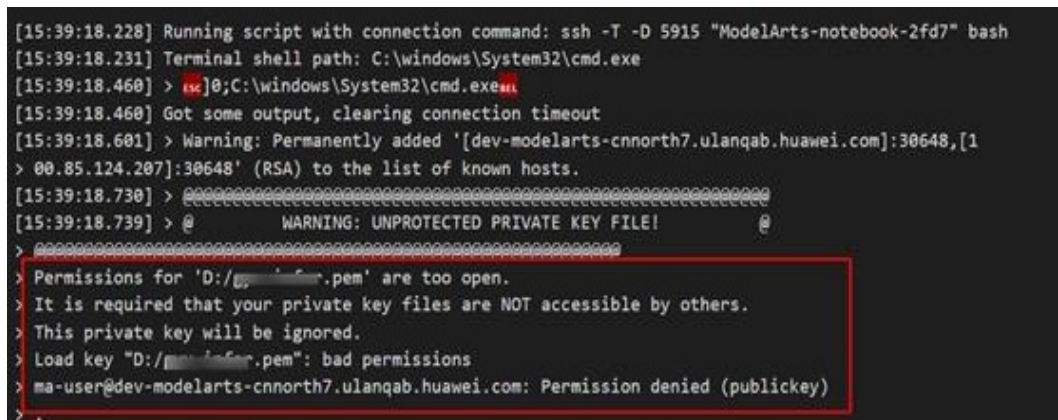


2. Open the command panel (**Ctrl+Shift+P** for Windows and **Cmd+Shift+P** for Mac), search for **Kill VS Code Server on Host**, and locate the affected instance, which will be automatically cleared. Then, establish the connection again.



3.8.9 What Do I Do If Error Message "Permissions for 'x:/xxx.pem' are too open" Is Displayed?

Symptom



Possible cause 1: The key file is not stored in the specified path. For details, see the [security restrictions](#) or [VS Code document](#).

Solution: Place the key file in a specified path or its sub-path:

Windows: **C:\Users\{{user}}**

Mac or Linux: **Users/{{user}}**

Possible cause 2: For Mac or Linux, the permission on the key file or the folder where the key is stored may be incorrect.

Solution: [Check the file and folder permissions](#).

Local SSH file and folder permissions

macOS / Linux:

On your local machine, make sure the following permissions are set:

Folder / File	Permissions
<code>.ssh</code> in your user folder	<code>chmod 700 ~/.ssh</code>
<code>.ssh/config</code> in your user folder	<code>chmod 600 ~/.ssh/config</code>
<code>.ssh/id_rsa.pub</code> in your user folder	<code>chmod 600 ~/.ssh/id_rsa.pub</code>
Any other key file	<code>chmod 600 /path/to/key/file</code>

Windows:

The specific expected permissions can vary depending on the exact SSH implementation you are using. We recommend using the out of box [Windows 10 OpenSSH Client](#).

In this case, make sure that all of the files in the `.ssh` folder for your remote user on the SSH host is owned by you and no other user has permissions to access it. See the [Windows OpenSSH wiki](#) for details.

For all other clients, consult your client's documentation for what the implementation expects.

3.8.10 What Do I Do If Error Message "ssh: connect to host xxx.pem port xxx: Connection refused" Is Displayed?

Symptom

```
[15:00:31.447] Running script with connection command: ssh -T -D 11839
"ModelArts-xxxx" bash
[15:00:31.449] Terminal shell path: C:\windows\System32\cmd.exe
[15:00:31.681] > rsc]0;C:\windows\System32\cmd.exe*ri
[15:00:31.681] Got some output, clearing connection timeout
[15:00:52.731] > ssh: connect to host ModelArts-xxxx port xxx
Connection timed out
[15:00:52.742] >
[15:00:54.019] "install" terminal command done
[15:00:54.020] Install terminal quit with output: i
[15:00:54.020] Received install output:
[15:00:54.020] Failed to parse remote port from server output
[15:00:54.022] Resolver error: Error:
```

Possible cause: The instance is not running.

Solution: Log in to the ModelArts management console and check the status of the instance. If the instance is stopped, start it. If the instance is in other states, such as **Error**, stop and then start it. After the instance status changes to **Running**, perform the remote connection again.

3.8.11 What Do I Do If Error Message "ssh: connect to host ModelArts-xxx port xxx: Connection timed out" Is Displayed?

Symptom


```
[15:00:31.447] Running script with connection command: ssh -T -D 11839
"ModelArts-..." bash
[15:00:31.449] Terminal shell path: C:\windows\System32\cmd.exe
[15:00:31.681] > [ESC]0;C:\windows\System32\cmd.exe[ESC]
[15:00:31.681] Got some output, clearing connection timeout
[15:00:52.731] > ssh: connect to host ModelArts-... port ...
Connection timed out
[15:00:52.742] > ...
[15:00:54.019] "install" terminal command done
[15:00:54.020] Install terminal quit with output: i...
[15:00:54.020] Received install output: ...
[15:00:54.020] Failed to parse remote port from server output
[15:00:54.022] Resolver error: Error:
```

Possible cause 1: The whitelisted IP addresses configured for the instance are different from the ones used in the local network.

Solution: so that the whitelisted IP addresses are the same as those used in the local network or disable the whitelist.

Possible cause 2: The local network is inaccessible.

Solution: Check the local network and network restrictions.

3.8.12 What Do I Do If Error Message "Load key "C:/Users/xx/test1/xxx.pem": invalid format" Is Displayed?

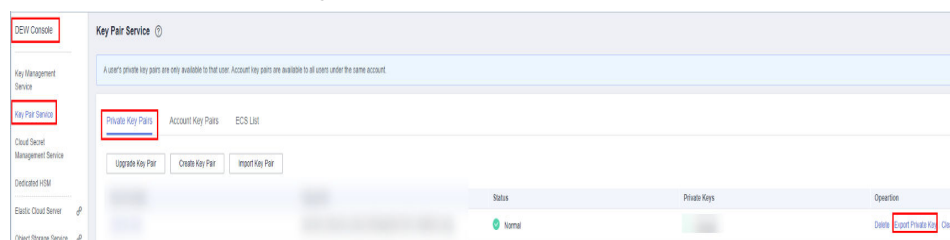
Symptom

```
[17:20:18.402] Running script with connection command: ssh -T -D 8578 "ModelArts-notebook-2fd7" bash
[17:20:18.404] Terminal shell path: C:\windows\System32\cmd.exe
[17:20:18.630] > [ESC]0;C:\windows\System32\cmd.exe[ESC]
[17:20:18.630] Got some output, clearing connection timeout
[17:20:18.777] > Warning: Permanently added '[dev-modelarts-cnnorth7.ulanqab.huawei.com]:30648,[1
> 00.85.124.207]:30648' (RSA) to the list of known hosts.
[17:20:18.904] > Load key "C:/Users/.../test1/...r.pem": invalid format
[17:20:18.922] > ma-user@dev-modelarts-cnnorth7.ulanqab.huawei.com: Permission denied (publickey)
```

Possible cause: The content or format of the key file is incorrect.

Solution: Use the correct key file for remote access. If there is no correct key file locally or the file is damaged, perform the following operations:

1. Log in to the console, search for **DEW**. On the DEW management console, choose **Key Pair Service** and click **Private Key Pairs**. Then, view and download the correct key file.

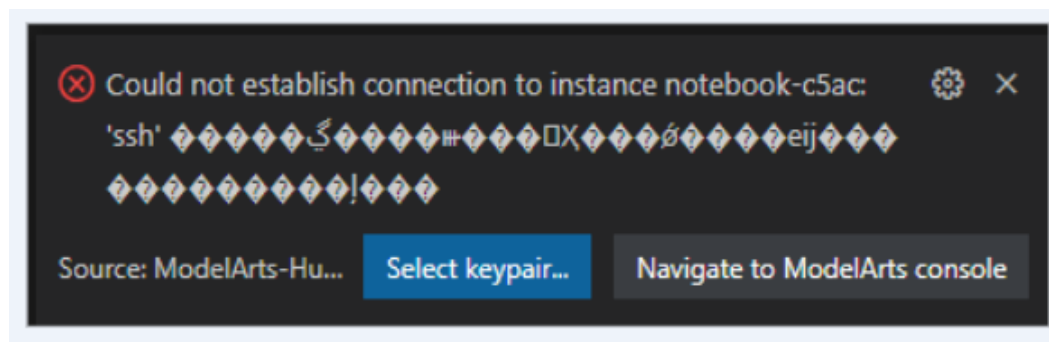
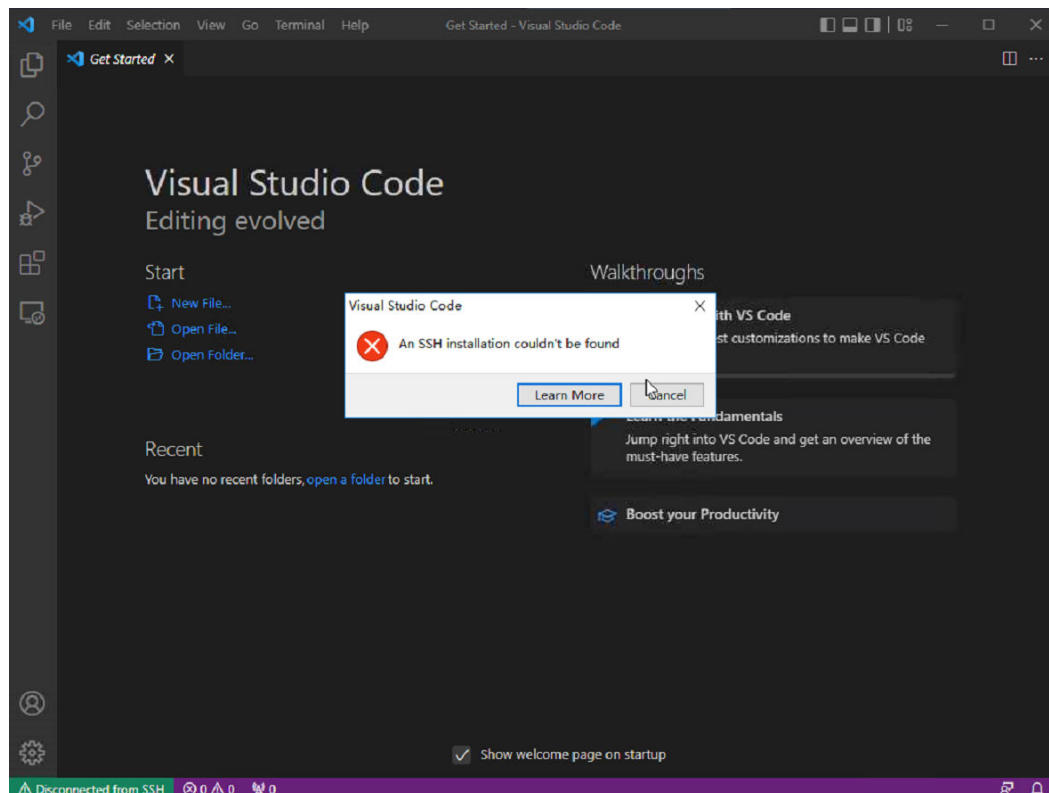


2. If the key cannot be downloaded and the originally downloaded key was lost, create a new development environment instance and a new key file.

Replacing a key file in a running development environment will be supported later.

3.8.13 What Do I Do If Error Message "An SSH installation couldn't be found" or "Could not establish connection to instance xxx: 'ssh' ..." Is Displayed?

Symptom



Possible cause: OpenSSH is not installed in the current environment or is not installed in the default path. For details, see the [VS Code document](#).

Solution

1. If OpenSSH is not installed in the current environment, [download and install it](#).

Installing a supported SSH client

OS	Instructions
Windows 10 1803+ / Server 2016/2019 1803+	Install the Windows OpenSSH Client .
Earlier Windows	Install Git for Windows .
macOS	Comes pre-installed.
Debian/Ubuntu	Run <code>sudo apt-get install openssh-client</code>
RHEL / Fedora / CentOS	Run <code>sudo yum install openssh-clients</code>

VS Code will look for the `ssh` command in the PATH. Failing that, on Windows it will attempt to find `ssh.exe` in the default Git for Windows install path. You can also specifically tell VS Code where to find the SSH client by adding the `remote.SSH.path` property to `settings.json`.

If OpenSSH fails to be installed, manually [download the OpenSSH installation package](#) and perform the following operations:

- Download the .zip package and decompress it into **C:\Windows\System32**.
- In **C:\Windows\System32**, open CMD as the administrator and run the following command:

```
powershell.exe -ExecutionPolicy Bypass -File install-sshd.ps1
```
- Add **C:\Program Files\OpenSSH-xx** (in which the SSH executable .exe file is stored) to environment system variables.
- Open CMD again and run `ssh`. If the following information is displayed, the installation is successful. Otherwise, go to **1.e** and **1.f**.

```
C:\windows\system32>ssh
usage: ssh [-46AaCfGgKkMnNqsTtVvXxYy] [-B bind_interface]
          [-b bind_address] [-c cipher_spec] [-D [bind_address:]port]
          [-E log_file] [-e escape_char] [-F configfile] [-I pkcs11]
          [-i identity_file] [-J [user@]host[:port]] [-L address]
          [-l login_name] [-m mac_spec] [-O ctl_cmd] [-o option] [-p port]
          [-Q query_option] [-R address] [-S ctl_path] [-W host:port]
          [-w local_tun[:remote_tun]] destination [command]
```

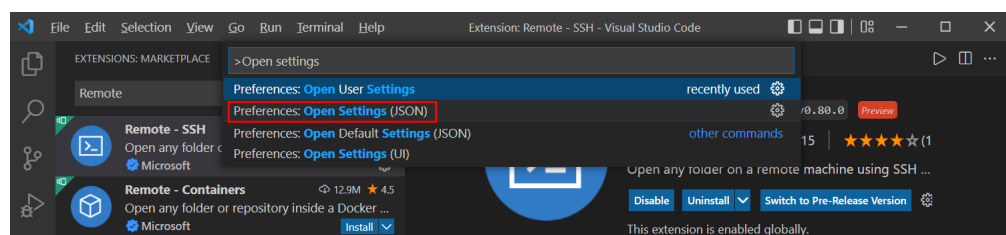
- Enable port 22 (default OpenSSH port) on the firewall and run the following command in Command Prompt:

```
netsh advfirewall firewall add rule name=sshd dir=in action=allow protocol=TCP localport=22
```
- Run the following command to start OpenSSH:

```
Start-Service sshd
```

- If OpenSSH is not installed in the default path, open the command panel (**Ctrl+Shift+P** for Windows and **Cmd+Shift+P** for Mac).

Search for **Open settings**.



Add the remote.SSH.path to **settings.json**, for example, "remote.SSH.path": "installation path of the local OpenSSH".

```
{
  "extensions.autoCheckUpdates": false,
  "extensions.autoUpdate": false,
  "remote.SSH.remotePlatform": {
    "ModelArts-notebook-": "linux"
  },
  "remote.SSH.path": "D:/OpenSSH-Win64/ssh.exe"
}
```

3.8.14 What Do I Do If Error Message "no such identity: C:/Users/xx /test.pem: No such file or directory" Is Displayed?

Symptom

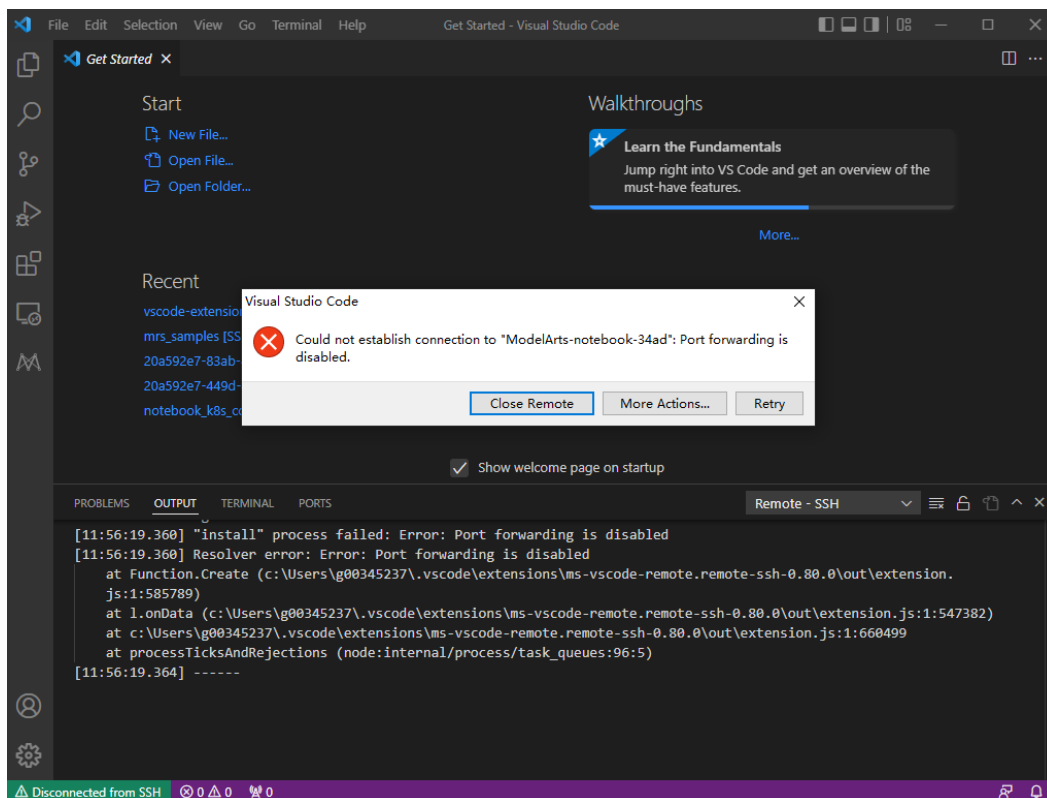
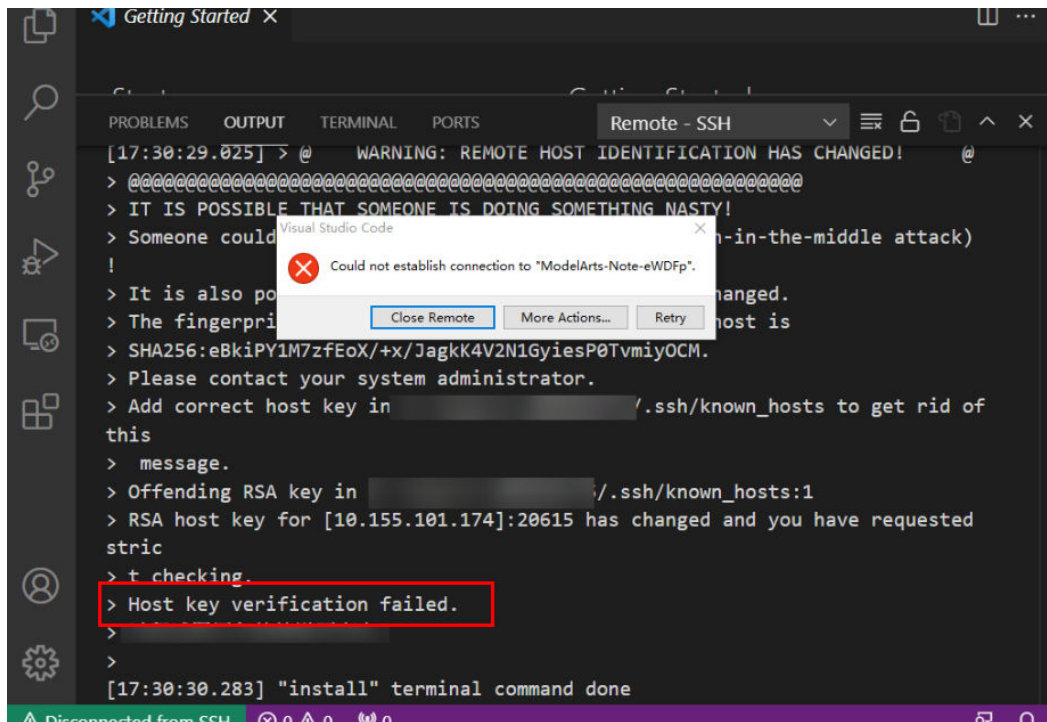
```
[17:27:44.947] Running script with connection command: ssh -T -D 8866 "ModelArts-notebook-2fd7" bash
[17:27:44.948] Terminal shell path: C:\windows\System32\cmd.exe
[17:27:45.179] > [exe]0;C:\windows\System32\cmd.exe[exe]
[17:27:45.179] Got some output, clearing connection timeout
[17:27:45.318] > Warning: Permanently added '[dev-modelarts-cnnorth7.ulanqab.huawei.com]:30648,[1
> 00.85.124.207]:30648' (RSA) to the list of known hosts.
[17:27:45.438] > no such identity: C:/Users/[redacted]/test.pem: No such file or directory
[17:27:45.455] > ma-user@dev-modelarts-cnnorth7.ulanqab.huawei.com: Permission denied (publickey)
> .
```

Possible cause: The key file is not in the path, or the name of the key file in the path has been changed.

Solution: Correct the path to the key file.

3.8.15 What Do I Do If Error Message "Host key verification failed" or "Port forwarding is disabled" Is Displayed?

Symptom



Possible cause: After the notebook instance is restarted, the public key changes. The alarm is generated when OpenSSH detected the key change.

Solution

- Add **-o StrictHostKeyChecking=no** for remote access through the CLI in VS Code.

```
ssh -tt -o StrictHostKeyChecking=no -i ${IdentityFile} ${User}@${HostName} -p ${Port}
```

Parameters:

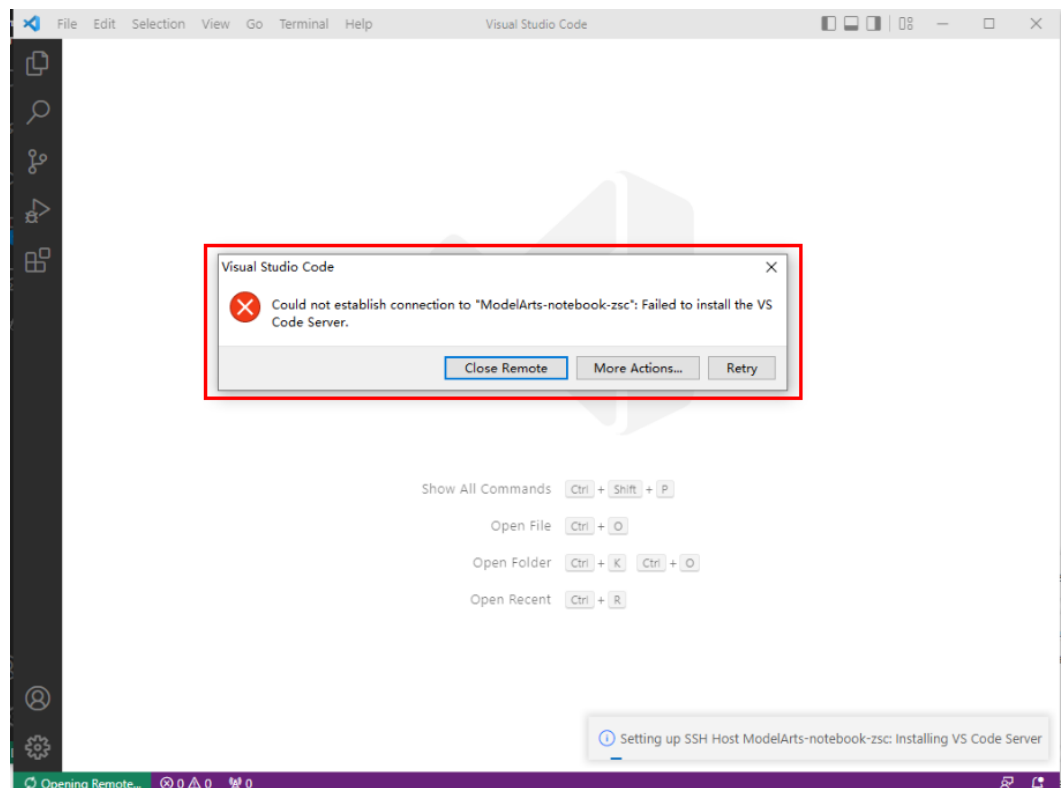
- **IdentityFile**: Path to the local key
- **User**: Username, for example, **ma-user**
- **HostName**: IP address
- **Port**: Port number
- Add **StrictHostKeyChecking no** and **UserKnownHostsFile=/dev/null** to the local **ssh config** file for manual configuration of remote access in VS Code.

```
Host xxx  
HostName x.x.x.x # IP address  
Port 22522  
User ma-user  
IdentityFile C:/Users/my.pem  
StrictHostKeyChecking no  
UserKnownHostsFile=/dev/null  
ForwardAgent yes
```

Note that SSH logins will be insecure after the preceding parameters are added because the **known_hosts** file will be ignored during the logins.

3.8.16 What Do I Do If Error Message "Failed to install the VS Code Server" or "tar: Error is not recoverable: exiting now" Is Displayed?

Symptom




```
[17:53:24.382] > vscode-scp-done.flag 100% 9 0.2KB/s 00:00
[17:53:24.756] > Found flag and server on host
[17:53:24.765] > d3aeabcaa9c5%2%
> tar --version:
[17:53:24.789] > tar (GNU tar) 1.30
> Copyright (C) 2017 Free Software Foundation, Inc.
> License GPLv3+: GNU GPL version 3 or later <https://gnu.org/licenses/gpl.html>.
> This is free software: you are free to change and redistribute it.
> There is NO WARRANTY, to the extent permitted by law.
>
> Written by John Gilmore and Jay Fenlason.
[17:53:24.796] > tar: This does not look like a tar archive
>
> gzip: stdin: unexpected end of file
> tar: Child returned status 1
> tar: Error is not recoverable: exiting now
[17:53:24.804] >
> ERROR: tar exited with non-0 exit code: 0
> Already attempted local download, failing
> d3aeabcaa9c5: start
> exitCode==37==
```

Possible cause: The disk space of `/home/ma-user/work` is insufficient.

Solution: Delete unnecessary files in `/home/ma-user/work`.

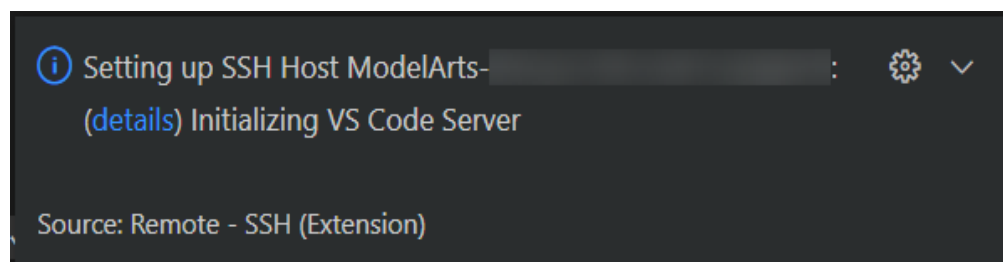
3.8.17 What Do I Do If Error Message "XHR failed" Is Displayed When a Remote Notebook Instance Is Accessed Through VS Code?

Possible cause: The network of the environment may be faulty.

Solution: Rectify the fault by referring to [Troubleshooting Failed XHR](#).

3.8.18 What Do I Do If It Takes a Long Time to Set Up a Remote Connection After VS Code Is Automatically Upgraded?

Symptom



Possible Cause

VS Code is automatically upgraded. As a result, download the new VS Code server to set up a new connection.

Solution

Disable automatic VS Code upgrade. To do so, click **Settings** in the lower left corner, search for **Update: Mode**, and set it to **none**.

Figure 3-11 Settings

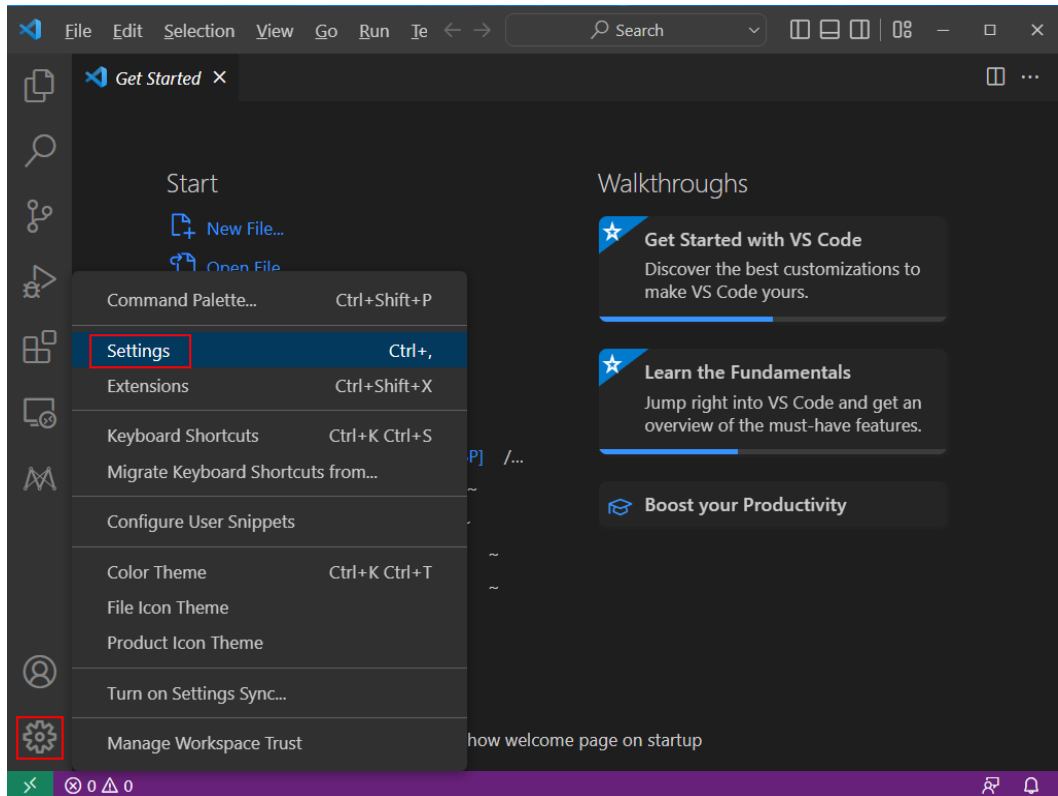
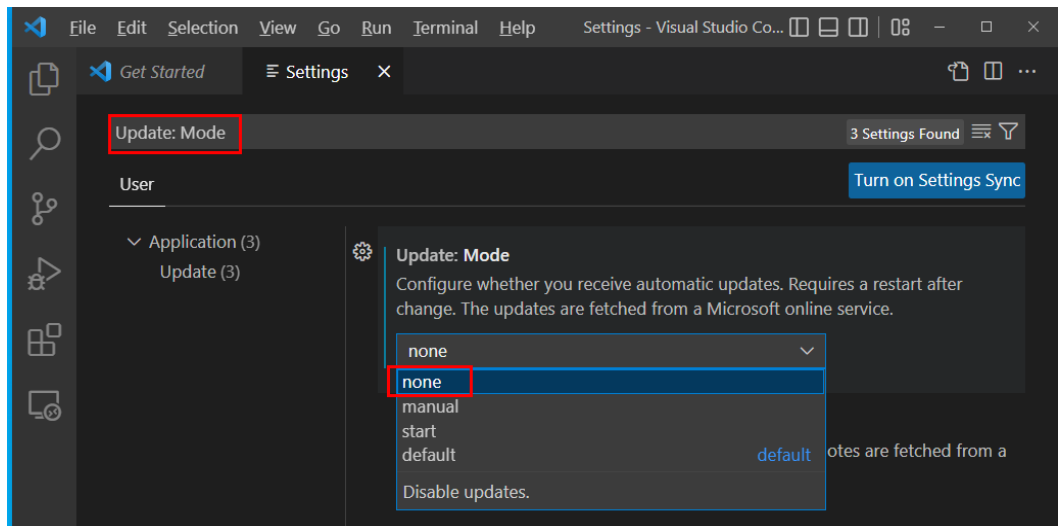


Figure 3-12 Setting the update mode to none



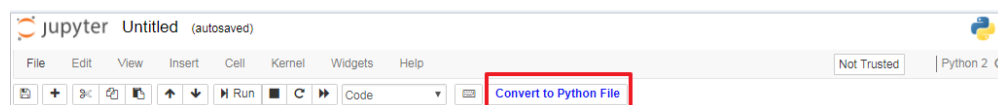
3.9 Others

3.9.1 How Do I Use Training Code from a Training Jobs After Debugging the Code in a Notebook Instance?

After the training code is debugged in a notebook instance, if you need to use the training code for training jobs on ModelArts, convert the **IPYNB** file into a Python file.

On the Jupyter page, click **Convert to Python File** to convert the training code to a Python file (.py file). After the conversion is completed, manually upload the Python file to OBS. Then, select the OBS path in the ModelArts training job to use the training code.

Figure 3-13 Converting the .IPYNB file into a Python file



3.9.2 Why Is the Training Speed Similar When Different Notebook Flavors Are Used?

If your training job is single-process in code, the training speed is basically the same no matter when the notebook flavor of 8 vCPUs and 64 GB of memory or the flavor of 72 vCPUs and 512 GB of memory is used. For example, if your training job uses 2 vCPUs and 4 GB of memory, the training speed is similar no matter when you use the notebook flavor of 4 vCPUs and 8 GB of memory or the flavor of 8 vCPUs and 64 GB of memory.

If your training job is multi-process in code, the training speed backed by the notebook flavor of 72 vCPUs and 512 GB of memory is higher than that backed by the notebook flavor of 8 vCPUs and 64 GB of memory.

3.9.3 How Do I Perform Incremental Training When Using MoXing?

If you are not satisfied with training results when using MoXing to build a model, you can perform incremental training after modifying some data and label information.

Adding Incremental Training Parameters to `mox.run`

After modifying labeling data or datasets, you can modify the `log_dir` parameter in and add the `checkpoint_path` parameter to `mox.run`. Set `log_dir` to a new directory and `checkpoint_path` to the output path of the previous training results. If the output path is an OBS directory, set the path to a value starting with `obs://`.

If labels are changed for label data, perform operations in [If Labels Are Changed](#) before running `mox.run`.

```
mox.run(input_fn=input_fn,
        model_fn=model_fn,
        optimizer_fn=optimizer_fn,
        run_mode=flags.run_mode,
        inter_mode=mox.ModeKeys.EVAL if use_eval_data else None,
```

```
log_dir=log_dir,  
batch_size=batch_size_per_device,  
auto_batch=False,  
max_number_of_steps=max_number_of_steps,  
log_every_n_steps=flags.log_every_n_steps,  
save_summary_steps=save_summary_steps,  
save_model_secs=save_model_secs,  
checkpoint_path=flags.checkpoint_url,  
export_model=mox.ExportKeys.TF_SERVING)
```

If Labels Are Changed

If the labels in a dataset have changed, execute the following statement. The statement must be executed before running **mox.run**.

In the statement, the **logits** variable indicates classification layer weights in different networks, and different parameters are configured. Set this parameter to the corresponding keyword.

```
mox.set_flag('checkpoint_exclude_patterns', 'logits')
```

If the built-in network of MoXing is used, the corresponding keyword needs to be obtained by calling the following API. In this example, the **Resnet_v1_50** keyword is the value of **logits**.

```
import moxing.tensorflow as mox  
  
model_meta = mox.get_model_meta(mox.NetworkKeys.RESNET_V1_50)  
logits_pattern = model_meta.default_logits_pattern  
print(logits_pattern)
```

You can also obtain a list of networks supported by MoXing by calling the following API:

```
import moxing.tensorflow as mox  
print(help(mox.NetworkKeys))
```

The following information is displayed:

```
Help on class NetworkKeys in module  
moxing.tensorflow.nets.nets_factory:  
  
class NetworkKeys(builtins.object)  
| Data descriptors defined here:  
|  
|  __dict__  
|      dictionary for instance variables (if defined)  
|  
|  __weakref__  
|      list of weak references to the object (if defined)  
|  
|-----  
| Data and other attributes defined here:  
|  
| ALEXNET_V2 = 'alexnet_v2'  
| CIFARNET = 'cifarnet'  
| INCEPTION_RESNET_V2 = 'inception_resnet_v2'  
| INCEPTION_V1 = 'inception_v1'  
| INCEPTION_V2 = 'inception_v2'  
| INCEPTION_V3 = 'inception_v3'  
| INCEPTION_V4 = 'inception_v4'
```

```
LENET = 'lenet'  
MOBILENET_V1 = 'mobilenet_v1'  
MOBILENET_V1_025 = 'mobilenet_v1_025'  
MOBILENET_V1_050 = 'mobilenet_v1_050'  
MOBILENET_V1_075 = 'mobilenet_v1_075'  
MOBILENET_V2 = 'mobilenet_v2'  
MOBILENET_V2_035 = 'mobilenet_v2_035'  
MOBILENET_V2_140 = 'mobilenet_v2_140'  
NASNET_CIFAR = 'nasnet_cifar'  
NASNET_LARGE = 'nasnet_large'  
NASNET_MOBILE = 'nasnet_mobile'  
OVERFEAT = 'overfeat'  
PNASNET_LARGE = 'pnasnet_large'  
PNASNET_MOBILE = 'pnasnet_mobile'  
PVANET = 'pvanet'  
RESNET_V1_101 = 'resnet_v1_101'  
RESNET_V1_110 = 'resnet_v1_110'  
RESNET_V1_152 = 'resnet_v1_152'  
RESNET_V1_18 = 'resnet_v1_18'  
RESNET_V1_20 = 'resnet_v1_20'  
RESNET_V1_200 = 'resnet_v1_200'  
RESNET_V1_50 = 'resnet_v1_50'  
RESNET_V1_50_8K = 'resnet_v1_50_8k'  
RESNET_V1_50_MOX = 'resnet_v1_50_mox'  
RESNET_V1_50_OCT = 'resnet_v1_50_oct'  
RESNET_V2_101 = 'resnet_v2_101'  
RESNET_V2_152 = 'resnet_v2_152'  
RESNET_V2_200 = 'resnet_v2_200'  
RESNET_V2_50 = 'resnet_v2_50'  
RESNEXT_B_101 = 'resnext_b_101'  
RESNEXT_B_50 = 'resnext_b_50'  
RESNEXT_C_101 = 'resnext_c_101'  
RESNEXT_C_50 = 'resnext_c_50'  
VGG_16 = 'vgg_16'
```

```
VGG_16_BN = 'vgg_16_bn'  
VGG_19 = 'vgg_19'  
VGG_19_BN = 'vgg_19_bn'  
VGG_A = 'vgg_a'  
VGG_A_BN = 'vgg_a_bn'  
XCEPTION_41 = 'xception_41'  
XCEPTION_65 = 'xception_65'  
XCEPTION_71 = 'xception_71'
```

3.9.4 How Do I View GPU Usage on the Notebook?

If you select GPU when creating a notebook instance, perform the following operations to view GPU usage:

1. Log in to the ModelArts management console, and choose **DevEnviron > Notebooks**.
2. In the **Operation** column of the target notebook instance in the notebook list, click **Open** to go to the **Jupyter** page.
3. On the **Files** tab page of the **Jupyter** page, click **New** and select **Terminal**. The **Terminal** page is displayed.
4. Run the following command to view GPU usage:
nvidia-smi
5. Check which processes in the current notebook instance use GPUs.
Open `/resource_info/gpu_usage.json` to view the processes that are using GPUs.

```
{  
  <notebook name>: {  
    <GPU0 UUID>: [  
      {  
        "pid": 2263,  
        "processName": "python",  
        "gpuMemoryUsage": "4935Mi"  
      },  
      {...}  
    ]  
    <GPU1 UUID>: [...]  
  }  
}
```

If no process uses GPUs, the file may not exist or be empty.

3.9.5 How Can I Obtain GPU Usage Through Code?

Run the shell or python command to obtain the GPU usage.

Using the shell Command

1. Run the **nvidia-smi** command.
This operation relies on CUDA NVCC.

```
watch -n 1 nvidia-smi
```

```
Every 1.0s: nvidia-smi
```

```
Mon Oct 25 15:20:11 2021
```

```

+-----+
| NVIDIA-SMI 440.33.01      Driver Version: 440.33.01   CUDA Version: 10.2   |
+-----+-----+-----+-----+-----+-----+
| GPU  Name          Persistence-M| Bus-Id        Disp.A | Volatile Uncorr. ECC |
| Fan  Temp   Perf    Pwr:Usage/Cap|      Memory-Usage | GPU-Util  Compute M. |
+-----+-----+-----+-----+-----+-----+
|    0   Tesla V100-SXM2...    On      | 00000000:5F:00.0 Off  |          0          |
| N/A   31C    P0     43W / 300W |  0MiB / 32510MiB |    0%      Default  |
+-----+-----+-----+-----+-----+-----+
|    1   Tesla V100-SXM2...    On      | 00000000:B5:00.0 Off  |          0          |
| N/A   34C    P0     44W / 300W |  0MiB / 32510MiB |    0%      Default  |
+-----+-----+-----+-----+-----+-----+

+-----+-----+-----+-----+-----+-----+
| Processes:                                                       GPU Memory |
|  GPU       PID    Type    Process name                               Usage      |
+-----+-----+-----+-----+-----+-----+
| No running processes found
+-----+-----+-----+-----+-----+-----+

```

2. Run the **gpustat** command.

```
pip install gpustat
gpustat -cp -i
```

```

notebook-6a654129-698e-4635-b6be-67aedbdd4c54 Mon Oct 25 15:19:11 2021 440.33.01
[0] Tesla V100-SXM2-32GB | 31'C, 0% | 0 / 32510 MB |
[1] Tesla V100-SXM2-32GB | 34'C, 0% | 0 / 32510 MB |

```

To stop the command execution, press **Ctrl+C**.

Using the python Command

1. Run the **nvidia-ml-py3** command (commonly used).

```

!pip install nvidia-ml-py3
import nvidia_smi
nvidia_smi.nvmlInit()
deviceCount = nvidia_smi.nvmlDeviceGetCount()
for i in range(deviceCount):
    handle = nvidia_smi.nvmlDeviceGetHandleByIndex(i)
    util = nvidia_smi.nvmlDeviceGetUtilizationRates(handle)
    mem = nvidia_smi.nvmlDeviceGetMemoryInfo(handle)
    print(f"|Device {i}| Mem Free: {mem.free/1024**2:5.2f}MB / {mem.total/1024**2:5.2f}MB | gpu-util:
{util.gpu:3.1%} | gpu-mem: {util.memory:3.1%} |")

```

```

Output:
|Device 0| Mem Free: 32510.44MB / 32510.50MB | gpu-util: 0.0% | gpu-mem: 0.0% |
|Device 1| Mem Free: 32510.44MB / 32510.50MB | gpu-util: 0.0% | gpu-mem: 0.0% |

```

2. Run the **nvidia_smi**, **wapper**, and **prettytable** commands.

Use the decorator to obtain the GPU usage in real time during model training.

```

def gputil_decorator(func):
    def wrapper(*args, **kwargs):
        import nvidia_smi
        import prettytable as pt

        try:
            table = pt.PrettyTable(['Devices','Mem Free','GPU-util','GPU-mem'])
            nvidia_smi.nvmlInit()
            deviceCount = nvidia_smi.nvmlDeviceGetCount()
            for i in range(deviceCount):
                handle = nvidia_smi.nvmlDeviceGetHandleByIndex(i)
                res = nvidia_smi.nvmlDeviceGetUtilizationRates(handle)
                mem = nvidia_smi.nvmlDeviceGetMemoryInfo(handle)

```

```

        table.add_row([i, f"{mem.free/1024**2:5.2f}MB/{mem.total/1024**2:5.2f}MB",
f"{res.gpu:3.1%}", f"{res.memory:3.1%}"])

    except nvidia_smi.NVMLError as error:
        print(error)

    print(table)
    return func(*args, **kwargs)
    return wrapper

```

Output:

Devices	Mem Free	GPU-util	GPU-mem
0	32510.44MB/32510.50MB	0.0%	0.0%
1	32510.44MB/32510.50MB	0.0%	0.0%

3. Run the **pynvml** command.

Run **nvidia-ml-py3** to directly obtain the nvml c-lib library, without using **nvidia-smi**. Therefore, this command is recommended.

```

from pynvml import *
nvmlInit()
handle = nvmlDeviceGetHandleByIndex(0)
info = nvmlDeviceGetMemoryInfo(handle)
print("Total memory:", info.total)
print("Free memory:", info.free)
print("Used memory:", info.used)

```

Output:

```

Total memory: 34089730048
Free memory: 34089664512
Used memory: 65536

```

4. Run the **gputil** command.

```

!pip install gputil
import GPUUtil as GPU
GPU.showUtilization()

```

Output:

ID	GPU	MEM
0	0%	25%
1	0%	0%

```

import GPUUtil as GPU
GPUs = GPU.getGPUs()
for gpu in GPUs:
    print("GPU RAM Free: {0:0f}MB | Used: {1:0f}MB | Util {2:3.0f}% | Total
{3:0f}MB".format(gpu.memoryFree, gpu.memoryUsed, gpu.memoryUtil*100, gpu.memoryTotal))

```

Output:

```

GPU RAM Free: 32510MB | Used: 0MB | Util 0% | Total 32510MB
GPU RAM Free: 32510MB | Used: 0MB | Util 0% | Total 32510MB

```

When using a deep learning framework such as PyTorch or TensorFlow, you can also use the APIs provided by the framework for query.

3.9.6 Which Real-Time Performance Indicators of an Ascend Chip Can I View?

The real-time performance indicator that can be viewed is **npu-smi**, which is similar to **nvidia-smi** of a GPU chip.

3.9.7 What Are the Relationships Between Files Stored in JupyterLab, Terminal, and OBS?

- Files stored in JupyterLab are the same as those in the work directory on the **Terminal** page. That is, the files are created on your notebook instances or synchronized from OBS.
- Notebook instances with OBS storage mounted can synchronize files from OBS to JupyterLab using the Sync OBS function. The files on the **Terminal** page are the same as those in JupyterLab.
- Notebook instances with EVS storage mounted can read files from OBS to JupyterLab using the MoXing API or SDKs. The files on the **Terminal** page are the same as those in JupyterLab.

3.9.8 How Do I Migrate Data from an Old-Version Notebook Instance to a New-Version One?

The notebook of the old version will be discontinued. Use the new version. This section describes how to migrate data from a notebook instance of the old version to a notebook instance of the new version.

Storage Differences Between the Old and New Versions

Table 3-1 Storage supported by notebook of the old and new versions

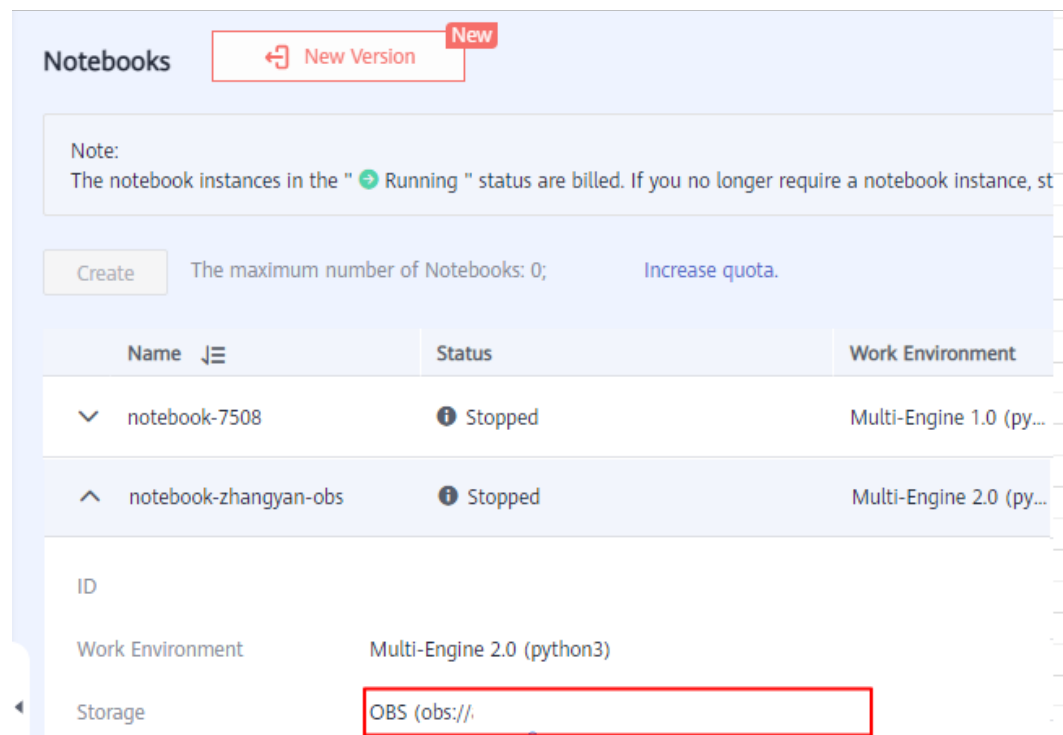
Storage	Old-Version Notebook	New-Version Notebook	Description
OBS	Supported	Not supported	OBS is a storage system, not a file system. In old-version notebook, remote replication and local replication of OBS data may be confused, leading to issues in controlling operations on data. Therefore, OBS mounting is removed from notebook of the new version. You can flexibly obtain and operate OBS data using code.

Storage	Old-Version Notebook	New-Version Notebook	Description
OBS parallel file system	Not supported	Supported	The new-version notebook allows dynamic mounting of OBS parallel file systems. You can mount storage on the details page of a running notebook instance. Data migration from the old version to the new version is not involved.
EVS	Supported	Supported	EVS disks can be attached to notebook instances of both the old and new versions. Data stored in the old version needs to be migrated to the new version.
SFS	Not supported	Supported	SFS is used in dedicated resource pools. This function has been discontinued in notebook of the old version. Therefore, data migration is not involved.
EFS	Not supported	Supported	EFS is used in notebook of the new version only.

OBS Used in Notebook of the Old Version

When notebook instances of the old version use OBS for storage, data is stored in OBS and does not need to be migrated. After a new-version notebook instance is created, directly use the data in the OBS directory. For details, see [How Do I Read and Write OBS Files in a Notebook Instance?](#)

Figure 3-14 OBS used in notebook of the old version



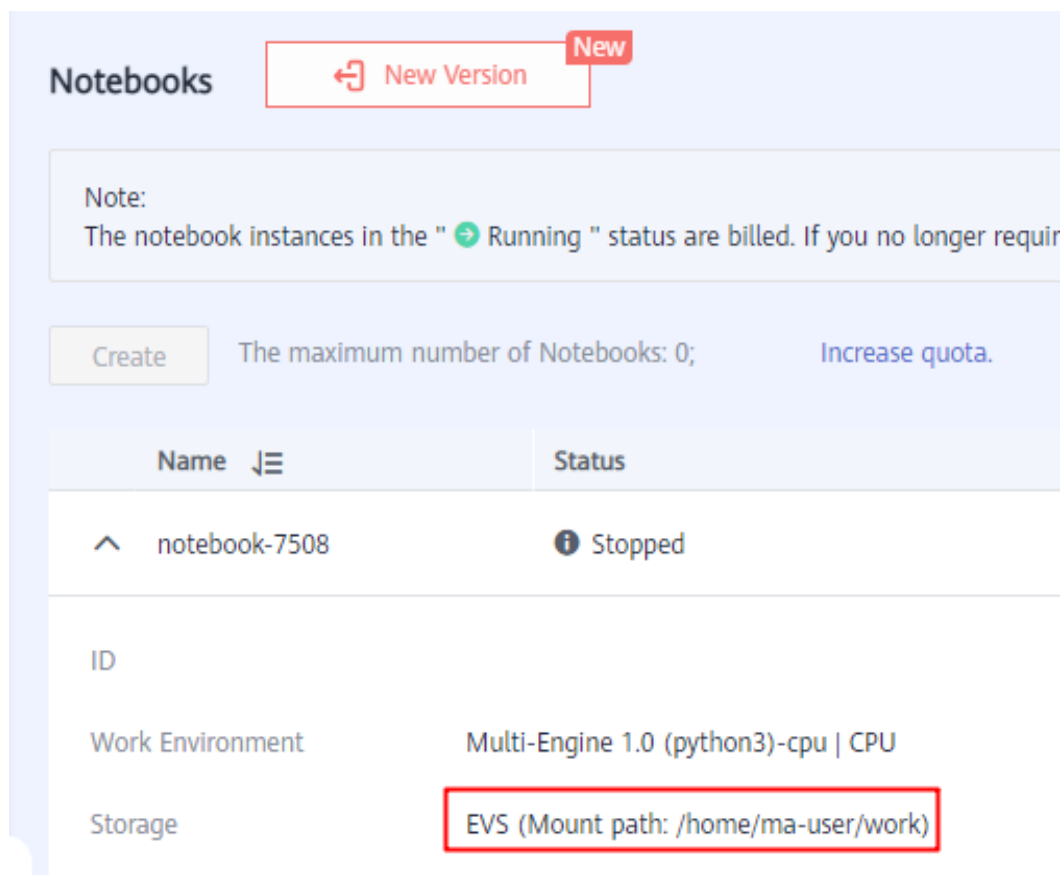
EVS Used in Notebook of the Old Version

If EVS disks are attached to a notebook instance of the old version for storing data, back up and migrate the EVS data to a notebook instance of the new version.

- If the volume of data stored in EVS is small, download the data to a local directory, create a notebook instance of the new version, and upload the data to the new notebook instance.
- If a large amount of data is stored in EVS, upload the data to an OBS bucket. After a notebook instance of the new version is created, read data from the the OBS bucket.

For more details, see .

Figure 3-15 EVS storage used in notebook of the old version



3.9.9 How Do I Use the Datasets Created on ModelArts in a Notebook Instance?

Datasets created on ModelArts are stored in OBS. To use these datasets in a notebook instance, download them from OBS to the notebook instance.

For details, see [How Do I Upload a File from a Notebook Instance to OBS or Download a File from OBS to a Notebook Instance?](#)

3.9.10 pip and Common Commands

pip is a common Python package management tool. It allows you to search for, download, install, and uninstall Python packages.

Common pip commands:

```
pip --help # Obtain help information.
pip install SomePackage==XXXX # Install a specified version.
pip install SomePackage # Install the latest version.
pip uninstall SomePackage # Uninstall a software version.
```

For other commands, run the **pip --help** command.

3.9.11 What Are Sizes of the /cache Directories for Different Notebook Specifications in DevEnviron?

When creating a notebook instance, you can select CPUs, GPUs, or Ascend based on the data volume.

ModelArts mounts disks to **/cache**. You can use this directory to store temporary files. The **/cache** directory shares resources with the code directory. The directory size varies depending on resource specifications.

No disks can be mounted to **/cache** for CPUs. When only one GPU or Ascend card is used, the **/cache** directory size is limited to 500 GB. If multiple GPUs or Ascend cards are used, the **/cache** directory size is limited to 3 TB and calculated using the following formula: **/cache** directory size = Number of cards x 500 GB. For details, see [Table 3-2](#).

Table 3-2 /cache directory sizes for different notebook specifications

Specification	/cache Directory Size
GPU, 0.25 cards	500 GB x 0.25
GPU, 0.5 cards	500 GB x 0.5
GPU, 1 card	500 GB
GPU, dual cards	500 GB x 2
GPU, four cards	500 GB x 4
GPU, eight cards	3 TB
Ascend, single card	500 GB
Ascend, dual cards	500 GB x 2
Ascend, four cards	500 GB x 4
Ascend, eight cards	3 TB
CPU	N/A

4 Training Jobs

4.1 Functional Consulting

4.1.1 What Are the Solutions to Underfitting?

1. Increasing model complexity
 - For an algorithm, add more high-order items to the regression model, improve the depth of the decision tree, or increase the number of hidden layers and hidden units of the neural network to increase model complexity.
 - Discard the original algorithm and use a more complex algorithm or model. For example, use the neural network to replace the linear regression, and use the random forest to replace the decision tree.
2. Adding more features to make input data more expressive
 - Feature mining is very important. Specifically, features with strong expression capabilities can outperform a large number of features with weak expression capabilities.
 - Feature quality is the focus.
 - To explore features with strong expression capabilities, you must have an in-depth understanding of data and application scenarios, which depends on experience.
3. Adjusting parameters and hyperparameters
 - Neural network: learning rate, learning attenuation rate, number of hidden layers, number of units in a hidden layer, β_1 and β_2 parameters in the Adam optimization algorithm, and `batch_size`
 - Other algorithms: number of trees in the random forest, number of clusters in k -means, and regularization parameter λ
4. Adding training data (not recommended)

Underfitting is usually caused by weak model learning capabilities. Adding data cannot significantly increase the training effect.
5. Reducing regularization constraints

Regularization aims to prevent model overfitting. If a model is underfitting instead of overfitting, reduce the regularization parameter λ or directly remove the regularization item.

4.1.2 How Do I Obtain a Trained ModelArts Model?

Models generated using ModelArts ExeML can be deployed only on ModelArts and cannot be downloaded to your local PC.

Models trained using a custom or subscription algorithm are stored in specified OBS paths for you to download.

4.1.3 How Do I Set the Runtime Environment of the AI Engine Scikit_Learn 0.18.1?

On the **Algorithm Management** page of ModelArts, click **Create**. On the **Create Algorithm** page, select **Show Old Engines** for **AI Engine**. Then, select XGBoost-Sklearn.

For details about how to create an algorithm in ModelArts, see .

For details about how to create a training job, see .

4.2 Reading Data During Training

4.2.1 How Do I Configure the Input and Output Data for Training Models on ModelArts?

ModelArts allows you to upload a custom algorithm for creating training jobs. and upload it to an OBS bucket. For details about how to create an algorithm, see . For details about how to create a training job, see .

Parsing Input and Output Paths

When a ModelArts model reads data stored in OBS or outputs data to a specified OBS path, perform the following operations to configure the input and output data:

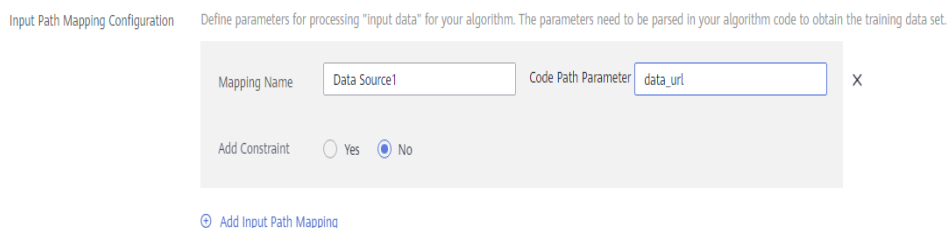
1. Parse the input and output paths in the training code. The following method is recommended:

```
import argparse
# Create a parsing task.
parser = argparse.ArgumentParser(description="train mnist",
                                formatter_class=argparse.ArgumentDefaultsHelpFormatter)
# Add parameters.
parser.add_argument('--train_url', type=str,
                    help='the path model saved')
parser.add_argument('--data_url', type=str, help='the training data')
# Parse the parameters.
args, unknown = parser.parse_known_args()
```

After the parameters are parsed, use **data_url** and **train_url** to replace the paths to the data source and the data output, respectively.

2. When using a preset image to create a custom algorithm, configure the input and output parameters on the **Create Algorithm** page based on code settings.
 - Training data is a must for algorithm development. By default, the input data is **Data Source** and the code path parameter is **data_url** (customizable).

Figure 4-1 Parsing the input path parameter **data_url**



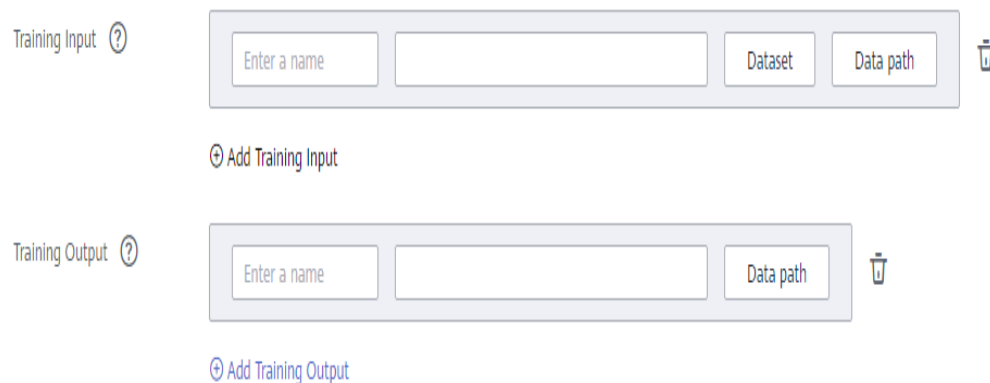
- After model training is complete, the trained model and the output information must be stored in an OBS path. By default, the output data is **Output Data** and the code path parameter is **train_url** (customizable).

Figure 4-2 Parsing the output path parameter **train_url**



3. When creating a training job, set the input and output paths. Select the OBS path or dataset path as the training input, and the OBS path as the output.

Figure 4-3 Setting training input and output



4.2.2 How Do I Improve Training Efficiency While Reducing Interaction with OBS?

Scenario Description

When ModelArts is used for custom deep learning training, training data is usually stored in OBS. If the volume of training data is large (for example, greater than

200 GB), a GPU resource pool is required for training each time, resulting in low training efficiency.

To improve training efficiency while reducing interaction with OBS, perform the following operations for optimization.

Optimization Principles

For the GPU resource pool provided by ModelArts, 500 GB NVMe SSDs are attached to each training node for free. The SSDs are attached to the **/cache** directory. The lifecycle of data in the **/cache** directory is the same as that of a training job. After the training job is complete, all content in the **/cache** directory is cleared to release space for the next training job. Therefore, you can copy data from OBS to the **/cache** directory during training so that data can be read from the **/cache** directory each time until the training is complete. After the training is complete, content in the **/cache** directory will be automatically cleared.

Optimization Methods

TensorFlow code is used as an example.

The following is code before optimization:

```
...
tf.flags.DEFINE_string('data_url', '', 'dataset directory.')
FLAGS = tf.flags.FLAGS
mnist = input_data.read_data_sets(FLAGS.data_url, one_hot=True)
```

The following is an example of the optimized code. Data is copied to the **/cache** directory.

```
...
tf.flags.DEFINE_string('data_url', '', 'dataset directory.')
FLAGS = tf.flags.FLAGS
import mox as mox
TMP_CACHE_PATH = '/cache/data'
mox.file.copy_parallel(FLAGS.data_url, TMP_CACHE_PATH)
mnist = input_data.read_data_sets(TMP_CACHE_PATH, one_hot=True)
```

4.2.3 Why the Data Read Efficiency Is Low When a Large Number of Data Files Are Read During Training?

If a dataset contains a large number of data files (massive small files) and data is stored in OBS, files need to be repeatedly read from OBS during training. As a result, the training process is waiting for reading files, resulting in low read efficiency.

Solution

1. Compress the massive small files into a package on your local PC, for example, a .zip package.
2. Upload the package to OBS.
3. During training, directly download this package from OBS to the **/cache** directory of your local PC. Perform this operation only once.

For example, you can use `mox.file.copy_parallel` to download the .zip package to the **/cache** directory, decompress the package, and then read files for training.

```
...
tf.flags.DEFINE_string('<obs_file_path>/data.zip', '', 'dataset directory.')
FLAGS = tf.flags.FLAGS
import os
import moxing as mox
TMP_CACHE_PATH = '/cache/data'
mox.file.copy_parallel(FLAGS.data_url, TMP_CACHE_PATH)
zip_data_path = os.path.join(TMP_CACHE_PATH, '*.zip')
unzip_data_path = os.path.join(TMP_CACHE_PATH, 'unzip')
# You can also decompress .zip Python packages.
os.system('unzip '+ zip_data_path + ' -d ' + unzip_data_path)
mnist = input_data.read_data_sets(unzip_data_path, one_hot=True)
```

4.3 Compiling the Training Code

4.3.1 How Do I Create a Training Job When a Dependency Package Is Referenced by the Model to Be Trained?

Store the **pip-requirements.txt** file in the training code directory.

NOTE

Any one of the following file names can be used. This section uses **pip-requirements.txt** as an example.

- pip-requirement.txt
- pip-requirements.txt
- requirement.txt
- requirements.txt

Before the training boot file is executed, the system automatically runs the following command to install the specified Python packages:

```
pip install -r pip-requirements.txt
```

- For details about the code directory, see [Storing the Installation File in the Code Directory](#).
- For details about the specifications of **pip-requirements.txt**, see [Installation File Specifications](#).

Storing the Installation File in the Code Directory

ModelArts allows you to install third-party dependency packages during model training in either of the following ways:

- Method 1 (recommended): , [store the required files or installation packages](#) in the code directory.

Figure 4-4 Creating an algorithm

- Method 2: Before using a common framework to create a training job, **store the required files or installation packages** in the code directory. (This function will be unavailable soon.)

Figure 4-5 Using a common framework to create an algorithm

Installation File Specifications

The installation file varies depending on the dependency package type.

- **WHL packages**
If the training background does not support the download of open source installation packages or use of user-compiled WHL packages, the system cannot automatically download and install the package. In this case, place the

WHL package in the code directory, create a file named **pip-requirements.txt**, and specify the name of the WHL package in the file. The dependency package must be a **.whl** file.

Take for example, an OBS path specified by **Code Dir** that contains model files, the **.whl** file, and the **pip-requirements.txt** file. The code directory structure would be as follows:

```
|---OBS path to the model boot file
|   |--model.py          #Model boot file
|   |--XXX.whl          #Dependency package. If multiple dependencies are required, place multiple
|                       #dependency packages here.
|   |--pip-requirements.txt #Defined configuration file, which specifies the name of the dependency
|                           #package
```

The following shows the content of the **pip-requirements.txt** file:

```
numpy-1.15.4-cp36-cp36m-manylinux1_x86_64.whl
tensorflow-1.8.0-cp36-cp36m-manylinux1_x86_64.whl
```

4.3.2 What Is the Common File Path for Training Jobs?

The path to the training environment and the code directory in the container are generally obtained using the environment variable **MA_JOB_DIR**, which is **/home/ma-user/modelarts/user-job-dir**.

4.3.3 How Do I Install a Library That C++ Depends on?

A third-party library may be used during job training. The following uses C++ as an example to describe how to install a third-party library.

1. Download source code to a local PC and upload it to OBS. For details about how to upload a file using OBS Browser, see .
2. Use MoXing to copy the source code uploaded to OBS to a notebook instance in the development environment.

The following is a code example for copying data to a notebook instance in a development environment running on an EVS:

```
import moxing as mox
mox.file.make_dirs('/home/ma-user/work/data')
mox.file.copy_parallel('obs://bucket-name/data', '/home/ma-user/work/data')
```

3. On the **Files** tab page of the **Jupyter** page, click **New** and select **Terminal**. Run the following command to go to the target path, and check whether the source code has been downloaded, that is, whether the **data** file exists.

```
cd /home/ma-user/work
ls
```

4. Compile code in **Terminal** based on service requirements.
5. Use MoXing to copy the compilation results to OBS. The following is a code example.

```
import moxing as mox
mox.file.make_dirs('/home/ma-user/work/data')
mox.file.copy_parallel('/home/ma-user/work/data', 'obs://bucket-name/file')
```

6. During training, use MoXing to copy the compilation result from OBS to the container. The following is a code example.

```
import moxing as mox
mox.file.make_dirs('/cache/data')
mox.file.copy_parallel('obs://bucket-name/data', '/cache/data')
```

4.3.4 How Do I Check Whether a Folder Copy Is Complete During Job Training?

In the script for training job boot file, run the following commands to obtain the sizes of the copied folders and the folders to be copied. Then determine whether folder copy is complete based on the command output.

```
import moxing as mox
mox.file.get_size('obs://bucket_name/obs_file',recursive=True)
```

get_size indicates the size of the file or folder to be obtained. **recursive=True** indicates that the type is folder. **True** indicates that the type is folder, and **False** indicates that the type is file.

If the command output is consistent, the folder copy is complete. If the command output is inconsistent, the folder copy is not complete.

4.3.5 How Do I Load Some Well Trained Parameters During Job Training?

During job training, some parameters need to be loaded from a pre-trained model to initialize the current model. You can use the following methods to load the parameters:

1. View all parameters by using the following code.

```
from moxing.tensorflow.utils.hyper_param_flags import mox_flags
print(mox_flags.get_help())
```
2. Specify the parameters to be restored during model loading. **checkpoint_include_patterns** is the parameter that needs to be restored, and **checkpoint_exclude_patterns** is the parameter that does not need to be restored.

```
checkpoint_include_patterns: Variables names patterns to include when restoring checkpoint. Such as:
conv2d/weights.
checkpoint_exclude_patterns: Variables names patterns to include when restoring checkpoint. Such as:
conv2d/weights.
```
3. Specify a list of parameters to be trained. **trainable_include_patterns** is a list of parameters that need to be trained, and **trainable_exclude_patterns** is a list of parameters that do not need to be trained.

```
--trainable_exclude_patterns: Variables names patterns to exclude for trainable variables. Such as:
conv1,conv2.
--trainable_include_patterns: Variables names patterns to include for trainable variables. Such as:
logits.
```

4.3.6 How Do I Obtain Training Job Parameters from the Boot File of the Training Job?

Training job parameters can be automatically generated in the background or you can enter them manually. To obtain training job parameters:

1. When a training job is created, **train_url** in the running parameters of the training job indicates where the training results are output to, and **data_url** indicates a data source. The **test** parameter is entered manually.

Figure 4-6 Creating a training job of the new version

Training Input

data_url Dataset Data path

⊕ Add Training Input

Training Output [?]

train_url Data path

⊕ Add Training Output

Figure 4-7 Creating a training job of the old version

* Algorithm Source

Algorithm Management **Frequently-used** Custom [MoXing Documentation](#)

Frequently-used AI engines used to create training jobs.

* AI Engine TensorFlow TF-1.13.1-python2.7

* Code Directory [?] /modelarts-test07/code/ 选择

* Boot File [?] /modelarts-test07/code/train_iceberg.py 选择

* Data Source [?] Dataset **Data path**

* Data Path /modelarts-test07/dataset-yolo/ Select ④ ⑤

* Training Output Path [?] /modelarts-test07/output/ Select

We recommend you select an empty directory as the output path.

Running Parameter [?]

train_url = /modelarts-test07/output/

data_url = /modelarts-test07/dataset-yolo/

⊕ Add Running Parameter

Job Log Path [?] Select Clear

By default, logs are stored in the service and will be deleted irregularly. Select a path for storing logs.

2. After the training job is executed, you can click the job name in the training job list to view its details. You can obtain the parameter input mode from logs, as shown in [Figure 4-8](#).

Figure 4-8 Viewing logs

```
[ModelArts Service Log]modelarts-pipe: will create log file /tmp/log/trainjob-4bac.log
* Restarting DNS forwarder and DHCP server dnsmasq
...done.
[ModelArts Service Log]user: uid=1101(work) gid=1101(work) groups=1101(work)
[ModelArts Service Log]pwd: /home/work
[ModelArts Service Log]app_url: s3://donotdel-modelarts-test/AI/code/PyTorch/
[ModelArts Service Log]boot_file: PyTorch/PyTorch.py
[ModelArts Service Log]log_url: /tmp/log/trainjob-4bac.log
[ModelArts Service Log]command: PyTorch/PyTorch.py --data_url=s3://donotdel-modelarts-test/AI/data/PyTorch/ --init_method=tcp://job1f00a54e-job-trainjob-4bac-0:6666 --test=test --  
train_url=s3://donotdel-modelarts-test/out/
```

3. To obtain the values of **train_url**, **data_url**, and **test** during training, add the following code to the boot file of the training job:

```
import argparse
parser = argparse.ArgumentParser()
parser.add_argument('--data_url', type=str, default=None, help='test')
parser.add_argument('--train_url', type=str, default=None, help='test')
parser.add_argument('--test', type=str, default=None, help='test')
```

4.3.7 Why Can't I Use `os.system('cd xxx')` to Access the Corresponding Folder During Job Training?

If you cannot access the corresponding folder by using `os.system('cd xxx')` in the boot script of the training job, you are advised to use the following method:

```
import os
os.chdir('/home/work/user-job-dir/xxx')
```

4.3.8 How Do I Invoke a Shell Script in a Training Job to Execute the `.sh` File?

ModelArts enables you to invoke a shell script, and you can use Python to invoke `.sh`. The procedure is as follows:

1. Upload the `.sh` script to an OBS bucket. For example, upload the `.sh` script to `/bucket-name/code/test.sh`.
2. Create the `.py` file on a local PC, for example, `test.py`. The background automatically downloads the code directory to the `/home/work/user-job-dir/` directory of the container. Therefore, you can invoke the `.sh` file in the `test.py` boot file as follows:

```
import os
os.system('bash /home/work/user-job-dir/code/test.sh')
```
3. Upload `test.py` to OBS. Then the file storage path is `/bucket-name/code/test.py`.
4. When creating a training job, set the code directory to `/bucket-name/code/`, and the boot file directory to `/bucket-name/code/test.py`.

After the training job is created, you can use Python to invoke the `.sh` file.

4.3.9 How Do I Obtain the Path for Storing the Dependency File in Training Code?

The code developed locally must be uploaded to the ModelArts backend. In training code, it is error-prone to set the path for storing the dependency file.

The following general solution is recommended: Use the OS API to obtain the absolute path of the dependency file.

Example:

```
|--project_root      # Root directory for code
| |--bootfile.py    # Boot file
| |--otherfileDirectory # Directory of other dependency files
| |--otherfile.py   # Other dependency files
```

Do as follows to obtain the path of the dependency file, `otherfile_path` in this example, in the boot file:

```
import os
current_path = os.path.dirname(os.path.realpath(__file__)) # Directory where the boot file is located
```

```
project_root = os.path.dirname(current_path) # Root directory of the project, which is the code directory set
on the ModelArts training console
otherfile_path = os.path.join(project_root, "otherfileDirectory", "otherfile.py")
```

4.4 Creating a Training Job

4.4.1 What Can I Do If the Message "Object directory size/ quantity exceeds the limit" Is Displayed When I Create a Training Job?

Issue Analysis

The code directory for creating a training job has limits on the size and number of files.

Solution

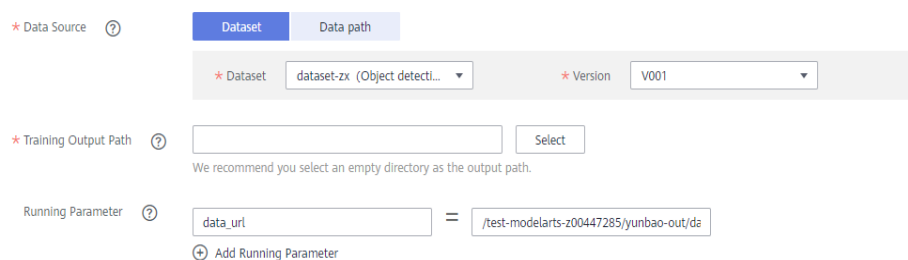
Delete the files except the code from the code directory or save the files in other directories. Ensure that the size of the code directory does not exceed 128 MB and the number of files does not exceed 4,096.

4.4.2 What Are Precautions for Setting Training Parameters?

Pay attention to the following when setting training parameters:

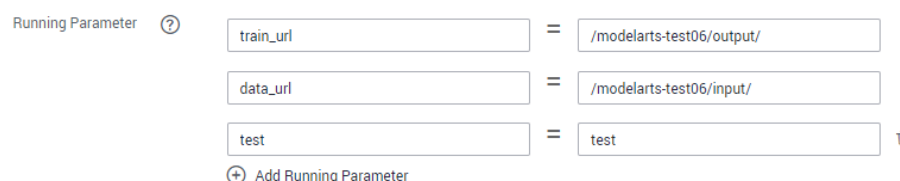
- If the algorithm source and data source have been configured, the **data_url** parameter is automatically set based on the selected object and cannot be directly modified by changing the running parameters.

Figure 4-9 Running parameters automatically set



- When setting running parameters for creating a training job, you only need to set the corresponding parameter names and values. See [Figure 4-10](#).

Figure 4-10 Setting running parameters



- If a parameter value is an OBS bucket path, use the path (starting with **obs://**) to the data. See [Figure 4-11](#).

Figure 4-11 Configuring an OBS path



- When creating an OBS folder in code, call a MoXing API as follows:


```
import moxing as mox
mox.file.make_dirs('obs://bucket_name/sub_dir_0/sub_dir_1')
```

4.4.3 What Are Sizes of the /cache Directories for Different Resource Specifications in the Training Environment?

When creating a training job, you can select CPU, GPU, or Ascend resources based on the size of the training job.

ModelArts mounts a disk to **/cache**. You can use this directory to store temporary files. The **/cache** directory shares resources with the code directory. The directory has different capacities for different resource specifications.

- GPU resources

Table 4-1 Capacities of the cache directories for GPU resources

GPU Specifications	cache Directory Capacity
V100	800 GB
8*V100	3 TB
P100	800 GB

- CPU resources

Table 4-2 Capacities of the cache directories for CPU resources

CPU Specifications	cache Directory Capacity
2 vCPUs 8 GiB	50 GB
8 vCPUs 32 GiB	50 GB

- Ascend resources

Table 4-3 Capacities of the cache directories for Ascend resources

Ascend Specifications	cache Directory Capacity
Ascend 910	3 TB

4.4.4 Is the /cache Directory of a Training Job Secure?

The program of a ModelArts training job runs in a container. The address of a directory to which the container is mounted is unique, and can be accessed only by the running container. Therefore, the **/cache** directory of the training job is secure.

4.4.5 Why Is a Training Job Always Queuing?

If the training job is always queuing, the selected resources are limited in the resource pool, and the job needs to be queued. In this case, wait for resources. To speed up resource obtaining, do as follows:

1. If you use a public resource pool:

Resources in a public resource pool are limited. During peak hours, resources may be insufficient if service traffic is heavy. Try to take the following measures:

- The less number of cards in the selected flavor leads to the lower queuing probability. For example, the probability of queuing when selecting a 1-card flavor is much less than that of queuing when selecting an 8-card flavor.
- If resources will be used for a long term, purchase a dedicated resource pool.

2. If you use a dedicated resource pool:

- If there are multiple available dedicated resource pools, switch to an idle one.
- Release resources in the current resource pool, for example, stop notebook instances that are not used for a long time.
- Submit a training job during off-peak hours.
- Contact the account administrator of the dedicated resource pool to expand the resource pool based on the resource usage.

4.5 Managing Training Job Versions

4.5.1 Does a Training Job Support Scheduled or Periodic Calling?

ModelArts training jobs do not support scheduled or periodic calling. When your job is in the **Running** state, you can call the job based on service requirements.

4.6 Viewing Job Details

4.6.1 How Do I Check Resource Usage of a Training Job?

In the left navigation pane of the ModelArts management console, choose **Training Management > Training Jobs** to go to the **Training Jobs** page. In the

training job list, click a job name to view job details. You can view the following metrics on the **Resource Usages** tab page.

- **CPU**: CPU usage (cpuUsage) percentage (Percent)
- **MEM**: Physical memory usage (memUsage) percentage (Percent)
- **GPU**: GPU usage (gpuUtil) percentage (Percent)
- **GPU_MEM**: GPU memory usage (gpuMemUsage) percentage (Percent)

4.6.2 How Do I Access the Background of a Training Job?

ModelArts does not support access to the background of a training job.

4.6.3 Is There Any Conflict When Models of Two Training Jobs Are Saved in the Same Directory of a Container?

Storage directories of ModelArts training jobs do not affect each other. Environments are isolated from each other, and data of other jobs cannot be viewed.

4.6.4 Only Three Valid Digits Are Retained in a Training Output Log. Can the Value of loss Be Changed?

In a training job, only three valid digits are retained in a training output log. When the value of **loss** is too small, the value is displayed as **0.000**. Log content is as follows:

```
INFO:tensorflow:global_step/sec: 0.382191
INFO:tensorflow:step: 81600(global step: 81600) sample/sec: 12.098 loss: 0.000
INFO:tensorflow:global_step/sec: 0.382876
INFO:tensorflow:step: 81700(global step: 81700) sample/sec: 12.298 loss: 0.000
```

Currently, the value of **loss** cannot be changed. You can multiply the value of **loss** by 1000 to avoid this problem.

4.6.5 Can a Trained Model Be Downloaded or Migrated to Another Account? How Do I Obtain the Download Path?

You can download the model trained by a training job and upload the downloaded model to OBS in the region corresponding to the target account.

Obtaining a Model Download Path

1. Log in to the ModelArts console. In the left navigation pane, choose **Training Management > Training Jobs**. The **Training Jobs** page is displayed.
2. In the training job list, click a job name to view job details.
3. On the **Configurations** tab page, obtain the path specified for **Training Output Path**, that is, the download path of the training model.

Migrating the Model to Another Account

There are two ways to migrate a trained model to another account:

- Download the trained model and then upload it to the OBS bucket in the region corresponding to the target account.
- Configure a policy for the folder or bucket where the model is stored to authorize other accounts to perform read and write operations. For details, see .

5 Model Management

5.1 Importing Models

5.1.1 How Do I Import the .h5 Model of Keras to ModelArts?

ModelArts does not support the import of models in .h5 format. You can convert the models in .h5 format of Keras to the TensorFlow format and then import the models to ModelArts.

For details about how to convert the Keras format to the TensorFlow format, see the [Keras official website](#).

6 Service Deployment

6.1 Functional Consulting

6.1.1 What Types of Services Can Models Be Deployed as on ModelArts?

Models can be deployed as real-time services or batch services.

6.1.2 What Are the Differences Between Real-Time Services and Batch Services?

- **Real-Time Services**
Models are deployed as web services. You can access the services through the management console or APIs.
- **Batch Services**
A batch service performs inference on batch data and automatically stops after data processing is completed.

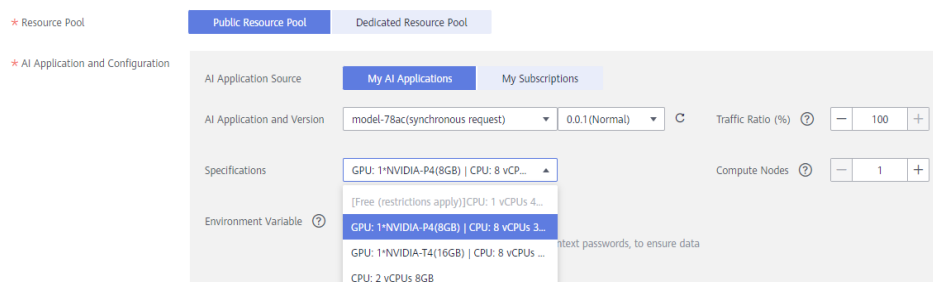
A batch service processes batch data at a time. A real-time service provides APIs for you to call.

6.1.3 How Do I Select Compute Node Specifications for Deploying a Service?

Before deploying a service, specify node specifications. The node specifications displayed on the GUI are calculated by ModelArts based on the target AI application and the node specifications available in the resource pool. You can select the specifications provided by ModelArts or customize the specifications (supported only in dedicated resource pools).

Selecting compute node specifications based on the resources required by your AI application. For example, if an AI application requires 3 CPUs and 10 GB of memory, select compute node specifications higher than 3 CPUs and 10 GB of memory. This ensures that the service can be successfully deployed and run properly.

Figure 6-1 Compute node specifications



When using compute node specifications, pay attention to the following:

Permission control

General-purpose compute node specifications are available publicly, for example, **modelarts.vm.cpu.2u**. You can select the specifications as long as there are idle resources in the resource pool. ModelArts provides two specifications by default, CPU-powered **modelarts.vm.cpu.2u** and GPU-powered **modelarts.vm.gpu.p4**.

For some special specifications, contact the system administrator to request for permissions.

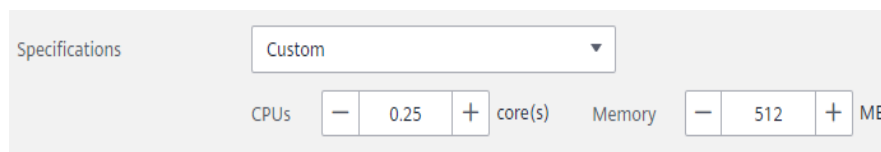
Specifications sold out in a public resource pool

Resources in a public resource pool are limited. If a specification is displayed as sold out, resources of the current specification have been used up. In this case, select other specifications or create your own dedicated resource pool.

Custom specifications

You can customize resource specifications only when a dedicated resource pool is used. Specifications cannot be customized in public resource pools.

Figure 6-2 Custom specifications



6.1.4 What Is the CUDA Version for Deploying a Service on GPUs?

CUDA 10.2 is supported by default. If a later version is required, submit a service ticket to apply for technical support.

6.2 Real-Time Services

6.2.1 What Do I Do If a Conflict Occurs in the Python Dependency Package of a Custom Prediction Script When I Deploy a Real-Time Service?

Before importing a model, save the inference code and configuration file in the model folder. When coding with Python, import custom packages in relative import (Python import) mode.

If there are packages with duplicate names in the ModelArts inference framework code and they are imported not in relative import mode, a conflict will occur, leading to a service deployment or prediction failure.

6.2.2 What Is the Format of a Real-Time Service API?

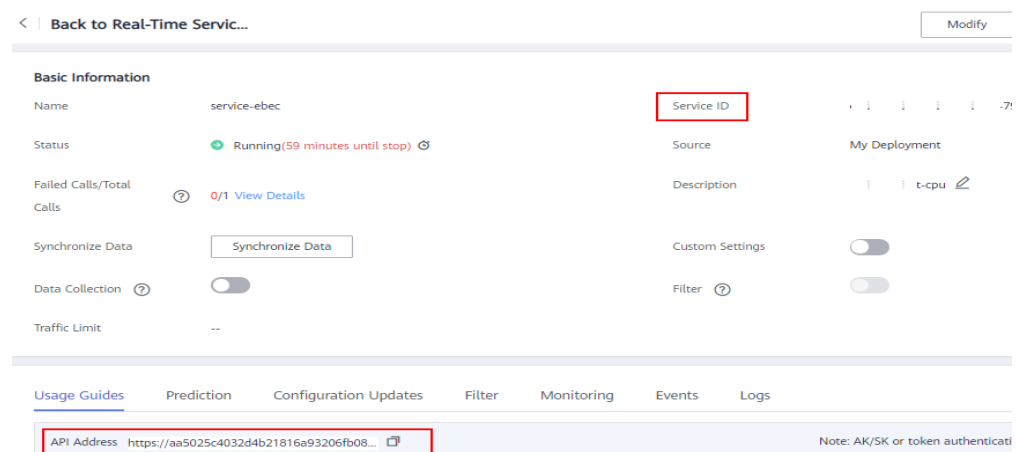
After a model is deployed as a real-time service, you can use the API for inference.

`https://Domain name/Version/infer/service ID`

Example:

`https://6ac81cdfac4f4a30be95xxxbb682.apig.xxx.xxx.com/v1/infers/468d146d-278a-4ca2-8830-0b6fb37d3b72`

Figure 6-3 API



6.2.3 What Do I Do If an Image Fails to Be Pulled When a Real-Time Service Is Deployed, Started, Upgraded, or Modified?

Possible Cause

The available disk space of the node is smaller than the image size.

Solution

1. Reduce the image size.
2. If the problem persists after the image size is reduced, contact the system administrator.

6.2.4 What Do I Do If an Image Restarts Repeatedly When a Real-Time Service Is Deployed, Started, Upgraded, or Modified?

Possible Cause

There is a bug in the container image code.

Solution

Debug the container image code based on container logs, create the AI application again, and deploy the application as a real-time service.

6.2.5 What Do I Do If a Container Health Check Failed When a Real-Time Service Is Deployed, Started, Upgraded, or Modified?

Possible Cause

Calling the container health check API failed.

Solution

Locate the fault based on container logs, debug the code, create an AI application again, and deploy the application as a real-time service.

6.2.6 What Do I Do If Resources Are Insufficient When a Real-Time Service Is Deployed, Started, Upgraded, or Modified?

Possible Cause

The configured instance specifications are beyond the specifications provided by the resource pool.

Solution

When resources are insufficient, ModelArts retries for three times. If resources are released during these retries, the service can be successfully deployed.

If resources are still insufficient after three retries, the service deployment fails. In this case, perform the following operations to resolve this issue:

- If the service is to be deployed in a public resource pool, wait until other users release resources.
- If the service is to be deployed in a dedicated resource pool, select lower container specifications or custom specifications to deploy the service on the premise that the model requirements are met.
- Expand the capacity of the current resource pool before deploying the service.

7 API/SDK

7.1 Can ModelArts APIs or SDKs Be Used to Download Models to a Local PC?

ModelArts APIs or SDKs cannot be used to download models to a local PC. However, the output models of training jobs are stored in OBS. You can use OBS APIs or SDKs to download the models. For details, see .

7.2 What Installation Environments Do ModelArts SDKs Support?

ModelArts SDKs can run in notebook or local environments. However, the supported environments vary depending on architectures. For details, see [Table 7-1](#).

Table 7-1 SDK installation environments

Development Environment	Architecture	Supported
Notebook	Arm	Yes
	x86	Yes
Local environment	Arm	No
	x86	Yes

7.3 Does ModelArts Use the OBS API to Access OBS Files over an Intranet or the Internet?

In the same region, ModelArts uses the OBS API to access files stored in OBS over an intranet and does not consume public network traffic.

If you download data from OBS through the Internet, you will be charged for the OBS public network traffic. For details about OBS billing, see .