

ALIBABA CLOUD

阿里云

机器学习PAI
PAI-DSW Notebook建模

文档版本：20200820

 阿里云

法律声明

阿里云提醒您在阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。
2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。
3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。
4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的“现状”、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。
5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含“阿里云”、“Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。
6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

通用约定

格式	说明	样例
 危险	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	 危险 重置操作将丢失用户配置数据。
 警告	该类警示信息可能会导致系统重大变更甚至故障，或者导致人身伤害等结果。	 警告 重启操作将导致业务中断，恢复业务时间约十分钟。
 注意	用于警示信息、补充说明等，是用户必须了解的内容。	 注意 权重设置为0，该服务器不会再接受新请求。
 说明	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	 说明 您也可以通过按Ctrl+A选中全部文件。
>	多级菜单递进。	单击设置> 网络> 设置网络类型。
粗体	表示按键、菜单、页面名称等UI元素。	在结果确认页面，单击确定。
<code>Courier</code> 字体	命令或代码。	执行 <code>cd /d C:/window</code> 命令，进入Windows系统文件夹。
<i>斜体</i>	表示参数、变量。	<code>bae log list --instanceid</code> <i>Instance_ID</i>
[] 或者 [a b]	表示可选项，至多选择一个。	<code>ipconfig [-all -t]</code>
{ } 或者 {a b}	表示必选项，至多选择一个。	<code>switch {active stand}</code>

目录

1. PAI-DSW V2专业版概述	05
2. PAI-DSW V1入门版概述	08
3. 授权	11
4. 购买	14

1. PAI-DSW V2专业版概述

PAI-DSW V2专业版基于阿里云Docker和Kubernetes等云原生技术，为您提供灵活且开放的AI开发环境。

新功能DSW镜像关联角色基础能力

功能特性

PAI-DSW V2专业版于2020年5月7日上线，在V1入门版的基础上，新增如下功能：

- 提供便捷的购买方式，可以选择更多的计算资源类型。
- 支持实例随时停止和启动、镜像一键保存、开发环境恢复及VPC访问。
- 提供集成式AI开发环境：
 - 预装常用大数据开发包和算法包，且开放Sudo权限，从而允许安装第三方库。
 - 预装JupyterLab插件，可以提高开发效率。例如Git及TensorBoard。
 - 提供官方镜像，可以覆盖多版本主流计算框架。例如TensorFlow及PyTorch。
 - 嵌入WebIDE，可以安装任意插件。
- 预制PAI的基础能力，包括视觉类算法工具EasyVision、自动调参工具AutoML、编译优化及读取MaxCompute表的CommonIO组件。

实例管理

您可以对PAI-DSW V2专业版实例进行以下操作：

- 启动

如果实例处于停止或启动失败状态，您可以重新启动实例。启动实例后，系统可以自动加载保存的镜像，恢复开发环境。对于按量计费（后付费）的实例，启动成功后，系统开始计费。
- 停止

实例停止支持以下方式：

 - 直接停止：释放ECS实例，停止PAI-DSW实例。对于按量计费（后付费）的实例，系统停止计费。
 - 保存系统再停止：先保存镜像，再停止ECS实例（ECS停机不收费模式）。重新启动实例时，可以恢复开发环境。通常停止时间大于直接停止的时间。
- 删除

如果在PAI-DSW控制台删除实例，则系统释放PAI-DSW实例、ECS实例及系统盘，从而导致数据无法恢复。对于自动创建的VPC、交换机及安全组，系统继续保留。如果需要完全清理资源，则在容器服务ACK控制台，删除名称以DSW_开头的集群。

实例环境

PAI-DSW V2专业版支持多种开发环境：

- JupyterLab

预置TensorBoard可视化工具及Git插件等，便于调试、优化及管理代码。
- WebIDE

不仅支持工程化的代码管理，而且支持在线开发、调试、运行代码。同时，您可以根据实际需要，自主安装任意插件。
- Terminal

适用于命令行编程。为了提供类似本地开发的体验，PAI-DSW V2专业版开放了Sudo权限，可以安装任意插件。

预置的PAI基础能力

PAI-DSW V2专业版内置了PAI的基础能力，您可以直接在实例内使用PAI视觉类算法工具EasyVision进行图像分类训练的评估和预测，或通过自动调参AutoML进行算法超参数自动调优。同时，PAI-DSW V2专业版提供CommonIO组件，您可以使用其直接读取MaxCompute表数据。

预置能力的使用方法请参见如下文档：

- [如何使用PAI-DSW SQLMagic写ODPS SQL](#)
- [如何使用PyODPS读写MaxCompute表数据](#)
- [如何使用PAI-DSW读取OSS Bucket中的数据](#)
- [如何使用tensorflow_io.oss直读OSS数据](#)
- [在PAI-DSW中实现NLP词向量word2vec](#)

官方镜像

PAI-DSW V2专业版提供以下官方镜像。

镜像	描述
py27_cpu_tf1.12_ubuntu	支持TensorFlow 1.12版本。
py27_cuda90_tf1.12_ubuntu	支持TensorFlow 1.12版本。
py36_cuda101_tf2.1_torch1.4_ubuntu	支持TensorFlow 2.1和PyTorch 1.4。
py36_cpu_tf2.1_torch1.4_ubuntu	支持TensorFlow 2.1和PyTorch 1.4 (CPU) 版本。

关联角色

创建PAI-DSW V2专业版实例之前，需要为以下服务关联角色授权，授权方法请参见[授权](#)。

角色	描述
AliyunPAIDSWDefaultRole	PAI-DSW使用该角色访问您的云资源。
AliyunCSDefaultRole	容器服务进行集群操作时，使用该角色访问您的云资源。
AliyunCSManagedLogRole	容器服务Kubernetes版的集群日志组件，使用该角色访问您的云资源。
AliyunCSManagedCmsRole	容器服务的集群CMS组件，使用该角色访问您的云资源。
AliyunCSClusterRole	容器服务在应用运行期，使用该角色访问您的云资源。
AliyunCSKubernetesAuditRole	容器服务Kubernetes版的审计功能，使用该角色访问您的云资源。
AliyunCSManagedNetworkRole	容器服务的集群网络组件，使用该角色访问您的云资源。

角色	描述
AliyunCSManagedKubernetesRole	容器服务Managed Kubernetes版，使用该角色访问您的云资源。
AliyunCSKubernetesAuditRole	容器服务Kubernetes版的审计功能，使用该角色访问您的云资源。
AliyunESSDefaultRole	弹性伸缩ESS使用该角色访问您的云资源。

2.PAI-DSW V1入门版概述

PAI-DSW是一款云端机器学习开发IDE，为您提供交互式编程环境，适用于不同水平的开发者。

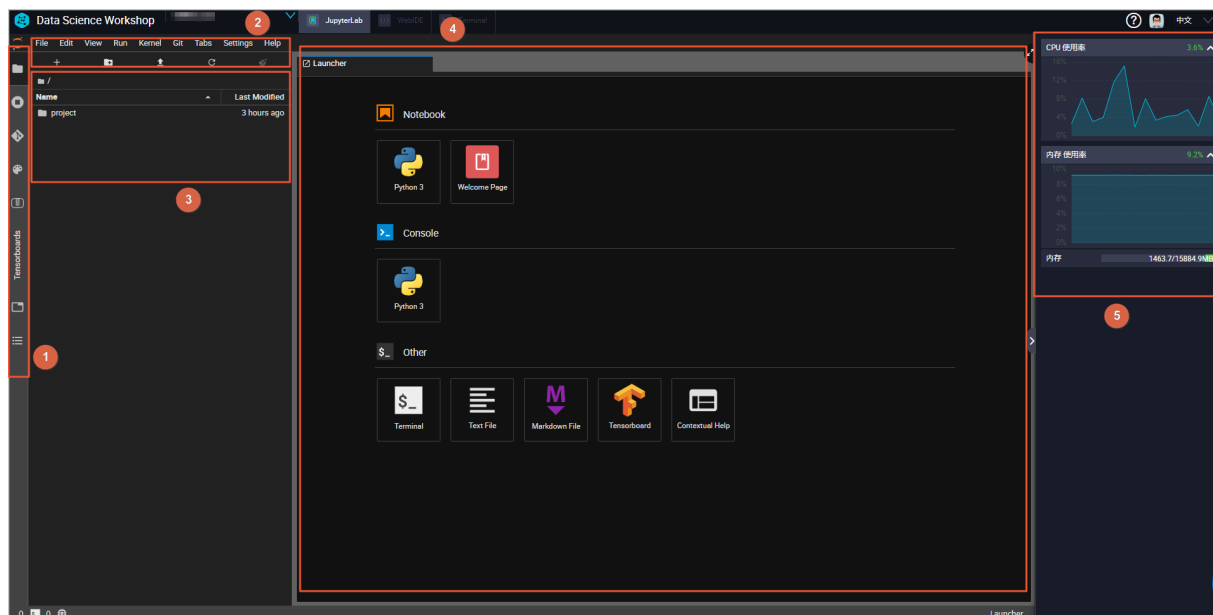
支持的框架资源功能页面介绍

PAI-DSW集成了开源JupyterLab，并以插件化的形式进行深度定制化开发。您无需任何运维配置，可以直接进行Notebook编写、调试及运行Python代码。同时，PAI-DSW提供丰富的计算资源，且对接多种数据源。通过EASCMD的方式，可以将PAI-DSW获得的训练模型部署为RESTful接口，对外提供模型服务，从而实现一站式机器学习。

功能特性

- 支持资源实时监控。算法开发时，可以显示CPU或GPU的使用情况。
- 支持多源数据接入，包括MaxCompute、OSS及NAS。
- 支持编写和运行SQL语句。
- 支持多种资源型号，包括纯CPU及多种GPU算力卡。
- 支持灵活切换各类资源，有效降低使用成本。
- 预置常用大数据开发包和算法库，且支持自定义安装第三方库。

实例环境



区域	描述
①	辅助工具栏
②	菜单栏和快捷工具
③	文件列表
④	代码编辑区域
⑤	资源使用情况

资源规格

资源类型	资源规格	可用地域
pai.medium.1xv100	GPU V100卡	<ul style="list-style-type: none"> 华北2（北京） 华东2（上海） 华东1（杭州） 华南1（深圳）
pai.medium.1xt4	GPU T4卡	<ul style="list-style-type: none"> 华东1（杭州） 华东2（上海） 华北2（北京） 华南1（深圳） 中国（香港） 新加坡（新加坡） 马来西亚（吉隆坡） 印度尼西亚（雅加达） 德国（法兰克福） 印度（孟买）
pai.medium.1xp100	GPU P100卡	<ul style="list-style-type: none"> 华东1（杭州） 华东2（上海） 华北2（北京） 华南1（深圳） 中国（香港） 新加坡（新加坡） 马来西亚（吉隆坡） 印度尼西亚（雅加达） 德国（法兰克福） 印度（孟买） <p>🔍 说明 华北2（北京）最大支持8卡。</p>
pai.medium.1xm40	GPU M40卡	华东2（上海）

资源类型	资源规格	可用地域
pai.large.2core4g	2 CPU Core+4 GB	<ul style="list-style-type: none"> • 华东1（杭州） • 华东2（上海） • 华北2（北京） • 华南1（深圳） • 中国（香港） • 新加坡（新加坡） • 马来西亚（吉隆坡） • 印度尼西亚（雅加达） • 德国（法兰克福） • 印度（孟买）
pai.xlarge.4core8g	4 CPU Core+8 GB	
pai.2xlarge.8core16g	8 CPU Core+16 GB	
pai.4xlarge.16core32g	16 CPU Core+32 GB	
pai.6xlarge.24core48g	24 CPU Core+48 GB	

计算框架

付费类型	框架	版本号
预付费	TensorFlow	2.0和1.12.3
	PyTorch	1.3.1
后付费	TensorFlow	2.0和1.14.0
	PyTorch	1.1.0

3. 授权

本文为您介绍如何为子账号和PAI-DSW关联角色授权。

子账号授权角色授权资源授权角色授权

背景信息

授权策略语言的结构和语法请参见[权限策略语法和结构](#)。

子账号授权

阿里云主账号可以授权子账号管理PAI-DSW实例，包括创建、启动、编辑及删除实例。

1. 登录[RAM控制台](#)。
2. 创建自定义权限策略。
 - i. 在左侧导航栏，选择权限管理 > 权限管理策略。
 - ii. 在权限管理策略页面，单击创建权限策略。
 - iii. 在新建自定义权限策略页面，配置参数。

参数	描述
策略名称	输入DSW_Notebook_Access。
备注	输入PAI-DSW访问策略。
配置模式	单击脚本配置。

参数	描述
策略内容	<p>将策略内容修改为如下内容。</p> <pre data-bbox="592 327 1385 1021"> { "Statement": [{ "Action": ["notebook:CreateInstance", "notebook:StartInstance", "notebook:StopInstance", "notebook>EditInstance"], "Effect": "Allow", "Resource": "*" }], "Version": "1" } </pre> <p>其中Action表示赋予的操作权限，可以包括以下权限：</p> <ul style="list-style-type: none"> ■ notebook:CreateInstance：创建PAI-DSW实例。 ■ notebook:StartInstance：开启PAI-DSW实例。 ■ notebook:StopInstance：停止PAI-DSW实例。 ■ notebook>EditInstance：编辑PAI-DSW实例。 ■ notebook>ListInstance：查看所有PAI-DSW实例。 <p>Resource表示资源权限，配置方式包括：</p> <ul style="list-style-type: none"> ■ 指定实例的地域权限 <pre data-bbox="619 1406 1385 1476">"Resource": "acs:notebook:cn-beijing:*:notebook/*"</pre> ■ 为特定实例（例如hhdemo）赋予PAI-DSW的使用权限 <pre data-bbox="619 1547 1385 1617">"Resource": "acs:notebook:*:*:notebook/hhdemo"</pre> ■ 为所有实例赋予PAI-DSW的使用权限 <pre data-bbox="619 1688 1385 1758">"Resource": "*"</pre> <p>如果需要配置更多权限，请参见权限策略基本元素。</p>

- iv. 单击确定。
3. 为子账号授权。

- i. 在左侧导航栏，选择人员管理 > 用户。
- ii. 在用户页面，单击操作列下的添加权限。
- iii. （可选）在添加权限面板，单击自定义策略。
- iv. 在选择权限下的文本框，输入DSW_Notebook_Access。
- v. 单击权限策略名称下的DSW_Notebook_Access，使其显示在已选择列表中。
- vi. 单击确定。

为关联角色授权

首次使用PAI-DSW，需要对相关资源进行访问授权。

1. 进入Notebook建模服务页面。
 - i. 登录PAI控制台。
 - ii. 在PAI控制台首页，选择模型开发和训练 > DSW-Notebook建模。
2. 单击创建实例。
3. 在角色授权对话框，单击去授权。
4. 在云资源访问授权页面，单击同意授权。在云资源访问授权页面，系统自动配置PAI-DSW需要的关联角色，无需手动配置。

4. 购买

本文为您介绍如何购买PAI-DSW V2专业版。

购买DSW开通DSW2.0

前提条件

授权容器服务ACK及ESS，并开通ESS，详情请参见[首次使用容器服务Kubernetes版](#)和[开通ESS并授权](#)。

背景信息

每次创建PAI-DSW V2专业版实例时，系统都会跳转至购买页面。使用PAI-DSW V2专业版会产生ECS、EIP及SLB相关费用。


PAI-DSW V1入门版的购买方式请参见[开通](#)。

操作步骤

1. 进入PAI-DSW V2专业版购买页面。
 - i. 登录[PAI控制台](#)。
 - ii. 在左侧导航栏，选择模型开发和训练 > DSW-Notebook建模。
 - iii. 在页面左上方，选择一个PAI-DSW V2专业版支持的地域。PAI-DSW V2专业版支持的地域包括华东1（杭州）和华南1（深圳）。
 - iv. 单击创建实例。
2. 在购买页面，配置参数。

参数	描述	
实例名称	只能包含英文字母、数字及下划线（_），长度不能超过27个字符。	
实例版本	默认为DSW2.0，不支持修改。	
地域及可用区	在不同地域的实例之间，网络不相通。选择距离较近的地域，可以降低网络时延，从而提高访问速度。	
付费模式	仅支持按量付费模式。	
网络配置	默认为专有网络，不支持修改。系统自动创建一组VPC、交换机及安全组。	
实例资源	架构	支持X86计算和 异构计算 GPU架构。
	分类	默认为所有类型，不支持修改。系统支持多种规格族，您可以根据实际需要，选择合适的规格族。规格族的说明请参见 实例规格族 。
实例镜像	支持的镜像覆盖Python、TensorFlow及PyTorch多个版本。	
存储	ECS系统盘配置	用于安装实例基础镜像，可以指定系统盘的云盘类型（SSD云盘、ESSD云盘或高效云盘）和容量。
	共享盘NAS	支持挂载已创建的文件系统。

参数	描述
自动停止	是否自动停止实例的开关。如果打开该开关，当PAI-DSW实例运行时长超过配置的自动停止时间，则系统自动停止该实例。

 **说明** 实例创建后，除实例名称以外，其他参数不支持修改。

3. 单击**确认订单**。
4. 核对订单信息，选中《**机器学习PAI DSW服务条款**》复选框，并单击**创建实例**。