

# 《语音识别》实验手册 1

## 目录

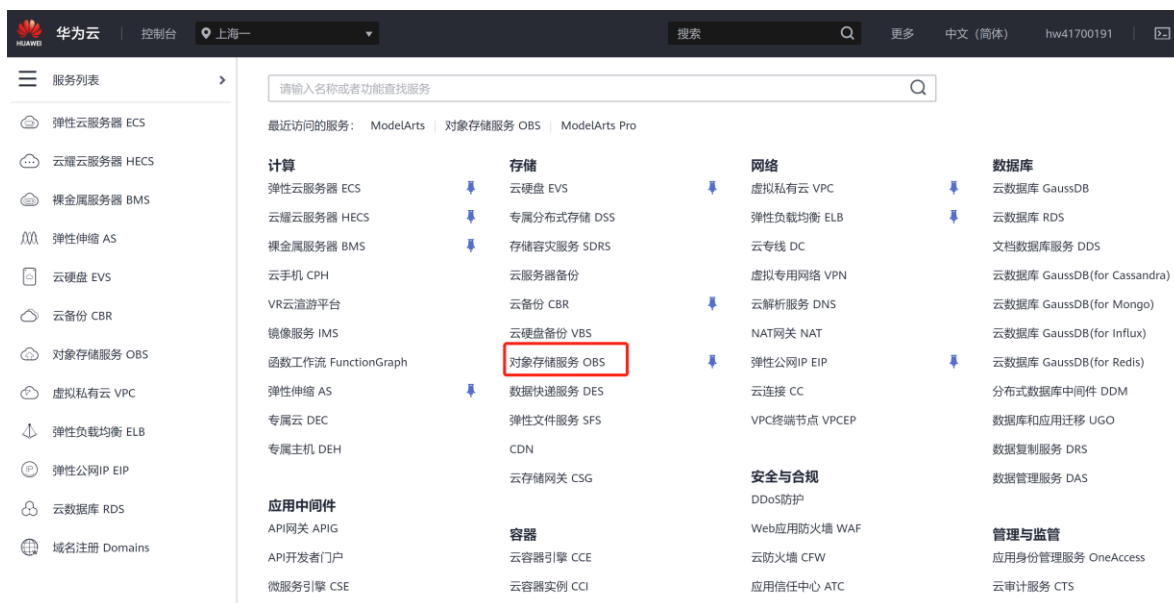
1 华为云使用指导.....	2
1.1 对象存储服务 OBS.....	2
1.2 ModelArts.....	5
2 语音分类.....	8
2.1 实验介绍.....	8
2.1.1 实验背景.....	8
2.1.2 实验目的.....	8
2.1.3 实验准备.....	8
2.2 实验步骤.....	8
2.3 实验小结.....	21
1. 连续语音识别.....	22

# 1 华为云使用指导

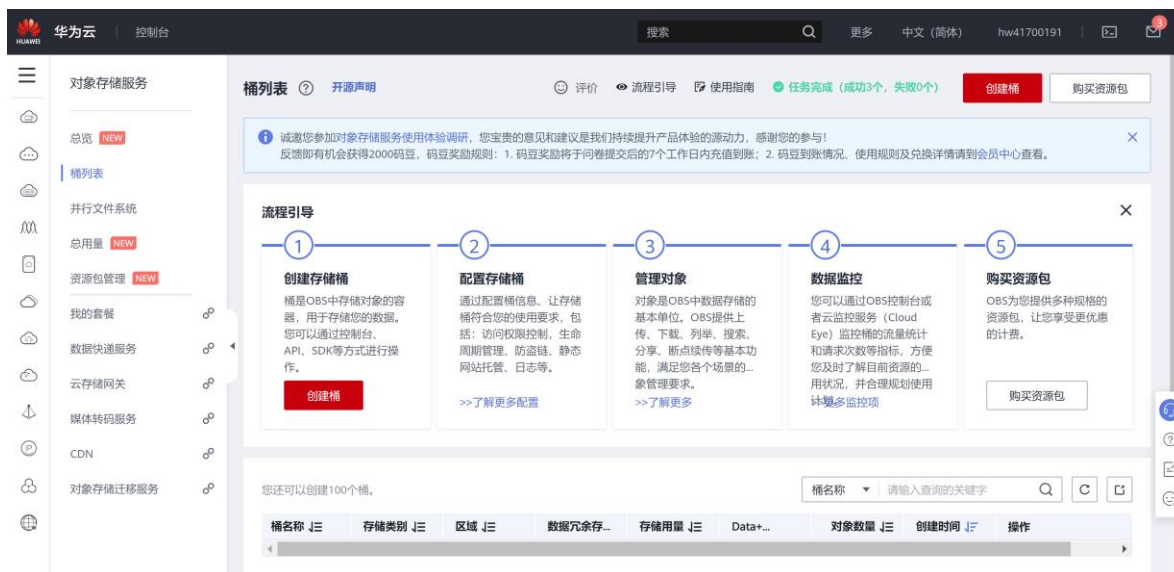
## 1.1 对象存储服务 OBS

对象存储服务（Object Storage Service）是一款稳定、安全、高效、易用的云存储服务，具备标准 Restful API 接口，可存储任意数量和形式的非结构化数据。

打开华为云官网（<https://console.huaweicloud.com/>）并点击 OBS 服务



OBS 服务是华为云中最常用的服务，通常都会搭配 ModelArts 服务一起使用。点击链接进入服务。



点击右上角的创建桶，来创建自己的 OBS 桶。



桶的具体配置情况，可以由自己设定，区域自行选择，桶名称自定义，储存类别为标准储存。桶策略请自行定义，如果需要别人也读取你的文件，请选择“公共读”，如果只供自己使用，建议选择“私有”，之后点击“立即创建”。



默认加密 ☐ 开启默认加密 免费 建议开启默认加密，密钥管理全免费，核心数据更安全。

归档数据直读   ?

关闭归档直读，归档存储类别的数据要先恢复才能访问。归档存储数据恢复和访问会收取相应的费用。[价格详情](#)

标签

如果您需要使用同一标签标识多种云资源，即所有服务均可在标签输入框下拉选择同一标签。建议在TMS中创建预定义标签。 [查看预定义标签](#) ?

标签键  标签值

您还可以添加10个标签。

**存储包超值购**

标准存储包 (多AZ)		标准存储包 (多AZ)		标准存储包 (多AZ)	
1 TB	1 年	1 PB	1 年	100 GB	1 个月

创建阶段 使用阶段

OBS桶: **创建免费** **按需/资源包计费** [OBS计费说明](#)

**立即创建**

之后，点击进入新创的桶，点击“对象”，进行数据的存储。

华为云 控制台

搜索 更多 中文 (简体) hw41700191

对象存储服务

总览 NEW

桶列表

并行文件系统

总用量 NEW

资源包管理 NEW

我的套餐

桶列表 ? [开源声明](#)

评价 流程引导 使用指南 任务完成 (成功3个, 失败0个) **创建桶** [购买资源包](#)

您还可以参加对象存储服务使用体验调研。您的意见和建议是我们持续提升产品体验的源动力，感谢您的参与！  
反馈即有机会获得2000码豆，码豆奖励规则：1. 码豆奖励将于问卷提交后的7个工作日内充值到账；2. 码豆到账情况、使用规则及兑换详情请到会员中心查看。

您还可以创建99个桶。

桶名称 存储类别 区域 数据冗余... 存储用量 Data+... 对象数量 创建时间 操作

speechclassi...	--	华东-上海一	--	--	--	--	2021/11/21 ...	<a href="#">修改存储类别</a> <a href="#">删除</a>
-----------------	----	--------	----	----	----	----	----------------	---

华为云 控制台

搜索 更多 中文 (简体) hw41700191

speechclassification

多版本控制 未启用 存储类别 标准存储 任务完成 (成功3个, 失败0个)

概览

**对象**

用量统计 NEW

访问权限控制

基础配置

域名管理

跨区域复制

数据回源

数据处理

桶清单

Data+

基本信息

桶名称 speechclassification

存储类别 标准存储

桶版本号 3.0

区域 华东-上海一

存储用量 0 byte

对象数量 0

帐号ID 0bb3aabf400010950f9ec00eb044c4c0

创建时间 2021/11/21 14:37:29 GMT+08:00

多版本控制 ? 未启用 [编辑](#)

Endpoint ? obs.cn-east-3.myhuaweicloud.com

基础配置

生命周期规则 ☐ 未配置

静态网站托管 ☐ 未配置

CORS规则 ☐ 未配置

防盗链 ☐ 未配置

事件通知 ☐ 未配置

标签 ☐ 未配置

日志记录 ☐ 未配置

默认加密 ☐ 未配置

归档数据直读 ☐ 未配置

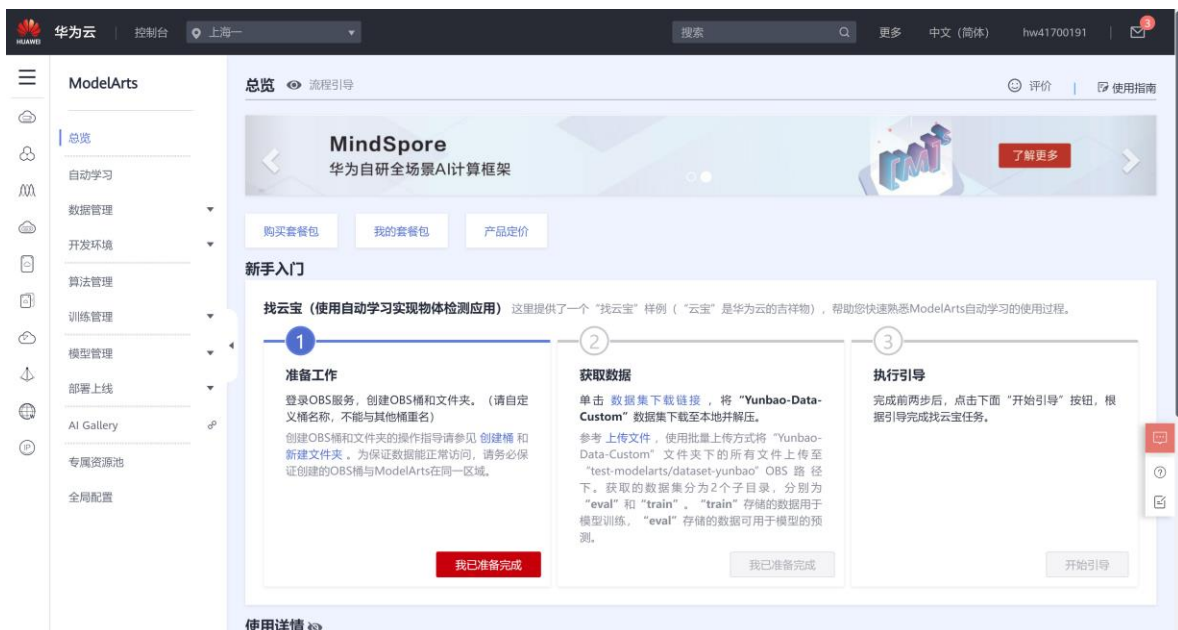
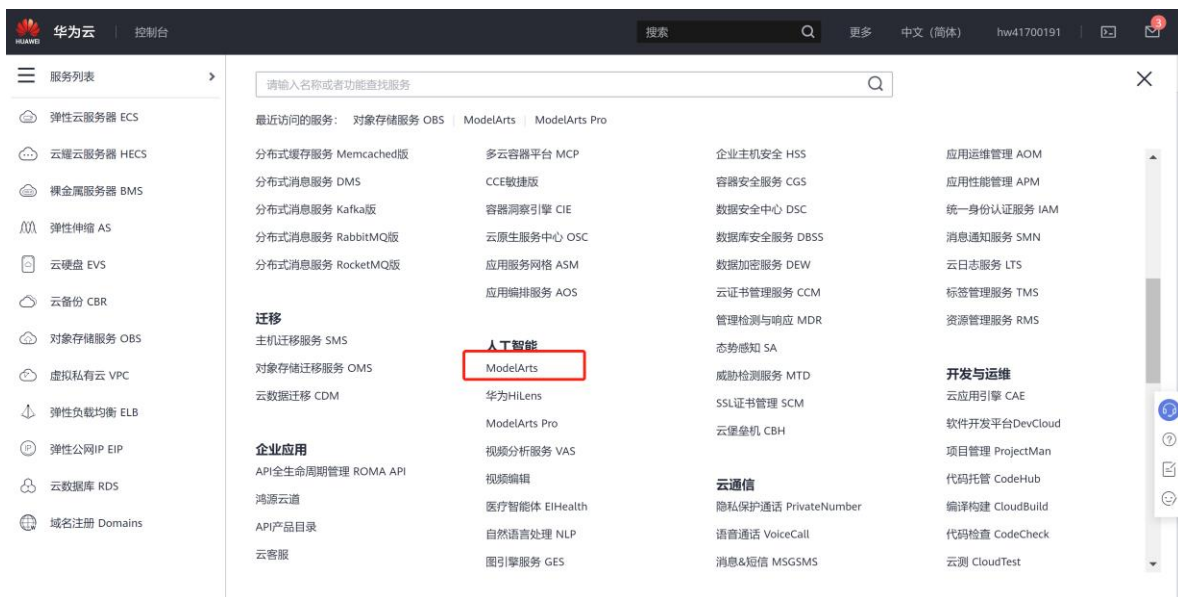
接下来就可以上传文件或者创建新的文件夹（暂时先不用）。



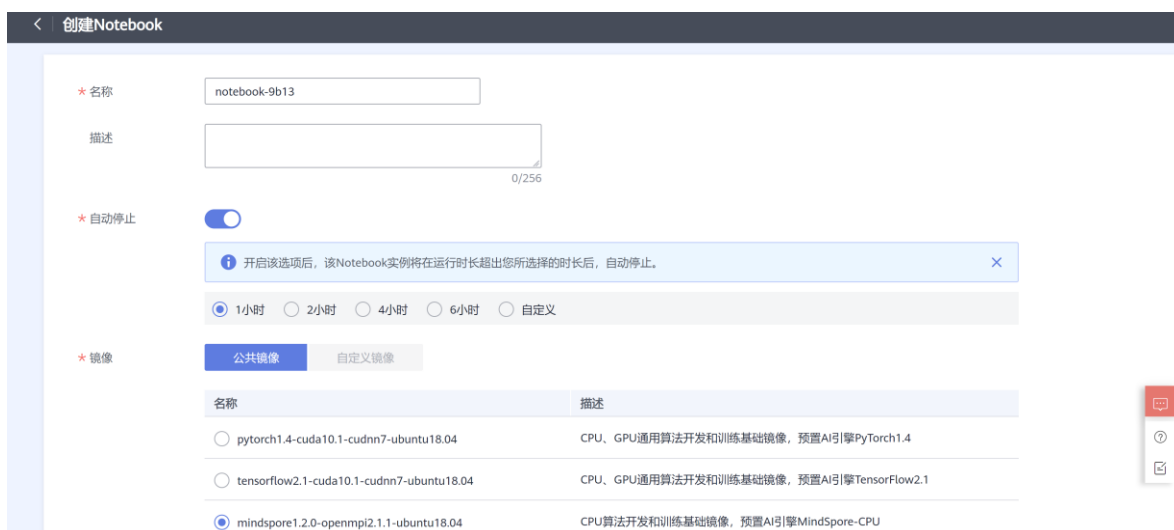
## 1.2 ModelArts

ModelArts 是面向开发者的一站式 AI 开发平台，为机器学习与深度学习提供海量数据预处理及半自动化标注、大规模分布式 Training、自动化模型生成，及端-边-云模型按需部署能力，帮助用户快速创建和部署模型，管理全周期 AI 工作流。

登录华为云，进入 ModelArts 控制台



## 创建 notebook 开发环境（暂时不用）



< 创建Notebook

产品名称	产品规格	计费模式	价格
notebook-9b13	描述	--	
	自动停止	1 小时	
	镜像	mindspore1.2.0-openmpi2.1.1-ubuntu18.04	
	资源池	公共资源池	
	规格	CPU: 2核 8GB	按需计费
	存储配置	默认存储	Notebook: ¥ 0.8/小时
	存储空间	50 GB	
	SSH远程开发	--	
	远程访问白名单	--	

配置费用: **¥0.80/小时**

先试后付费套餐用量

了解更多信息

上一步

提交

# 2 语音分类

## 2.1 实验介绍

### 2.1.1 实验背景

ModelArts 自动学习是帮助人们实现 AI 应用的低门槛、高灵活、零代码的定制化模型开发工具。自动学习功能根据标注数据自动设计模型、自动调参、自动训练、自动压缩和部署模型。当前自动学习支持快速创建图像分类、物体检测、预测分析和声音分类模型的定制化开发。可广泛应用在工业、零售安防等领域。

本实验通过 ModelArts 自动学习实现动物叫声分类，主要面向业务开发者，无需专业的开发基础和编码能力，只需上传数据，通过自动学习界面引导和简单操作即可完成模型训练和部署。具体流程如图 2-1 所示。

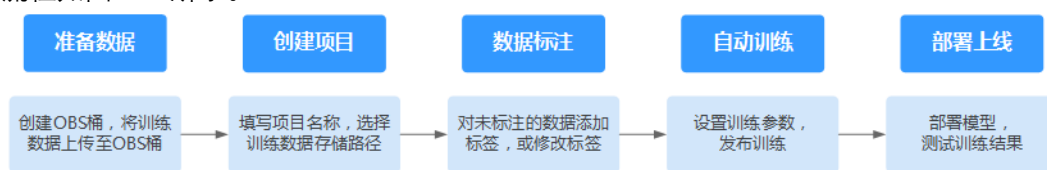


图2-1

### 2.1.2 实验目的

熟练使用 ModelArts 自动学习实现模型训练与部署

### 2.1.3 实验准备

(1) 实验环境：华为云 ModelArts 自动学习

(2) 实验所需数据集：

[https://modelarts-labs.obs.cn-north-1.myhuaweicloud.com/ExeML/ExeML\\_Sound\\_Classification/sound\\_classification.zip](https://modelarts-labs.obs.cn-north-1.myhuaweicloud.com/ExeML/ExeML_Sound_Classification/sound_classification.zip)

## 2.2 实验步骤

(1) 数据准备

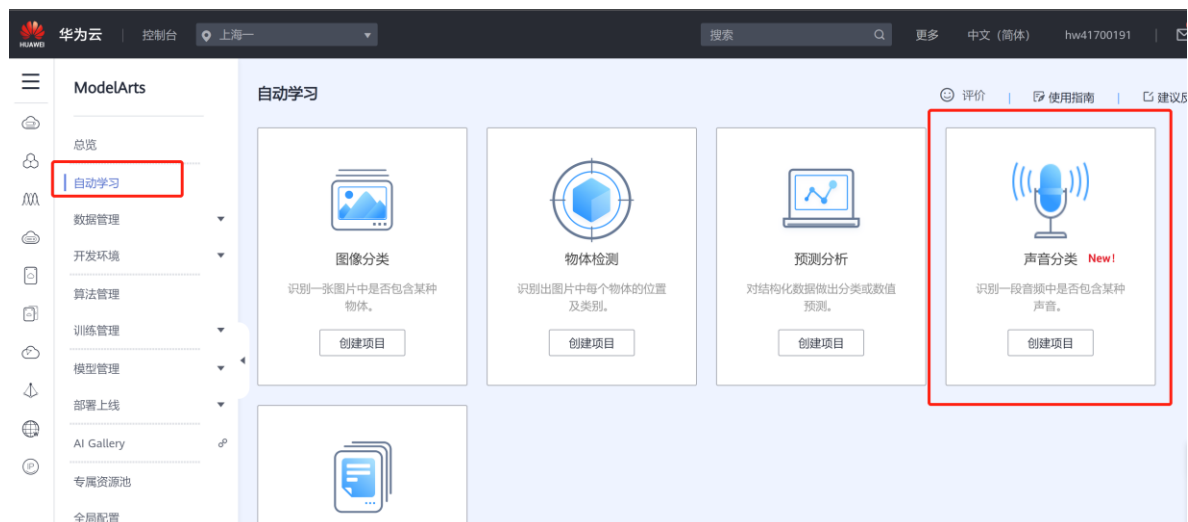
根据 2.1.3 中的数据集链接下载数据集压缩包至本地，解压得到文件夹 sound\_classification。获取的数据集包含“train”和“test”两个目录，其中训练集位于 sound\_classification\train 目录下，共 4 类动物叫声，分别是鸟 (bird)、猫 (cat)、狗 (dog) 和虎 (tiger)，每种动物 25 个叫声音频数据，共一百条音频数据。测试集位于 sound\_classification\test 目录下，每种动物提供 5 个叫声音



频数据。音频数据格式均为 wav 格式，均可在各音频播放器直接播放试听。

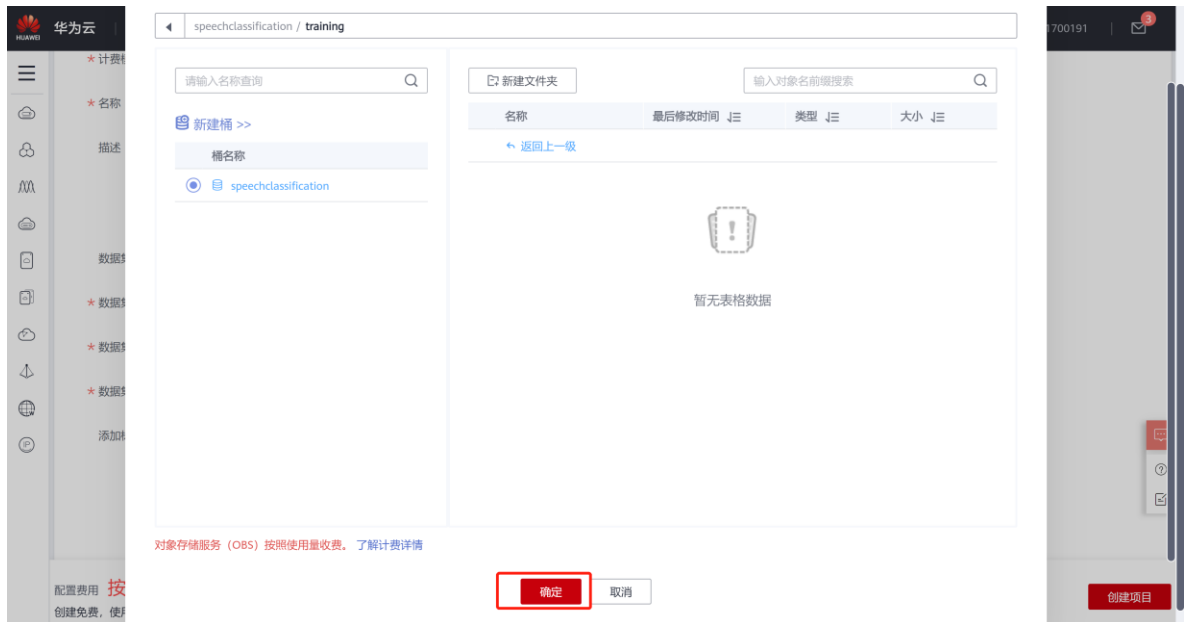
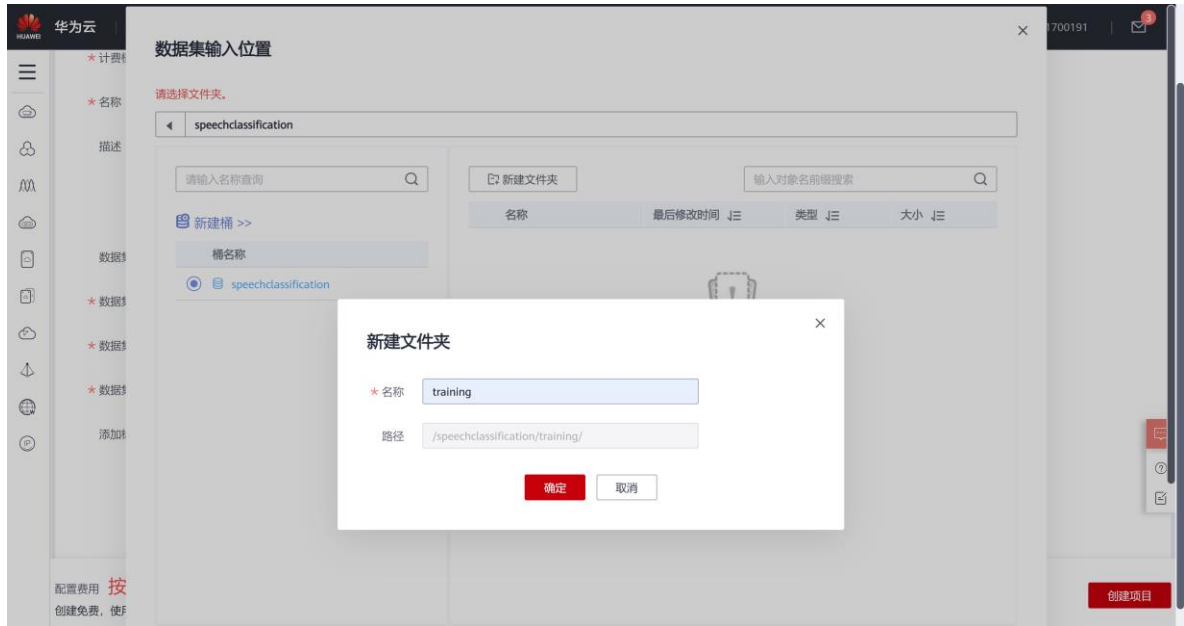
## (2) 创建声音分类项目

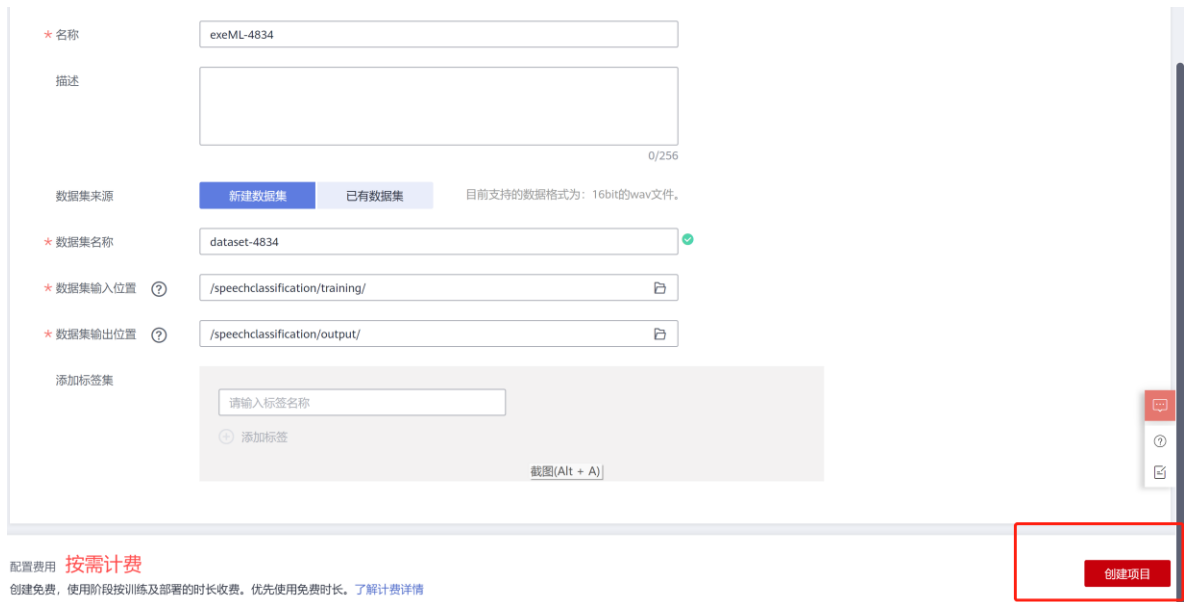
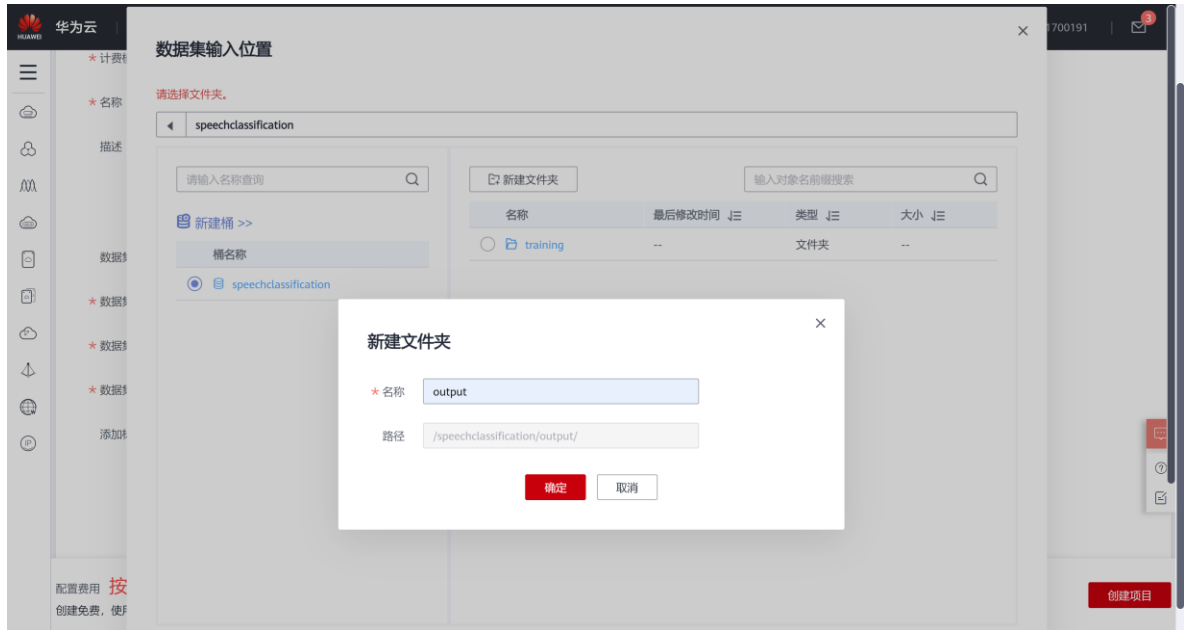
点击进入 ModelArts 自动学习界面，然后点击“声音分类”创建项目按钮：

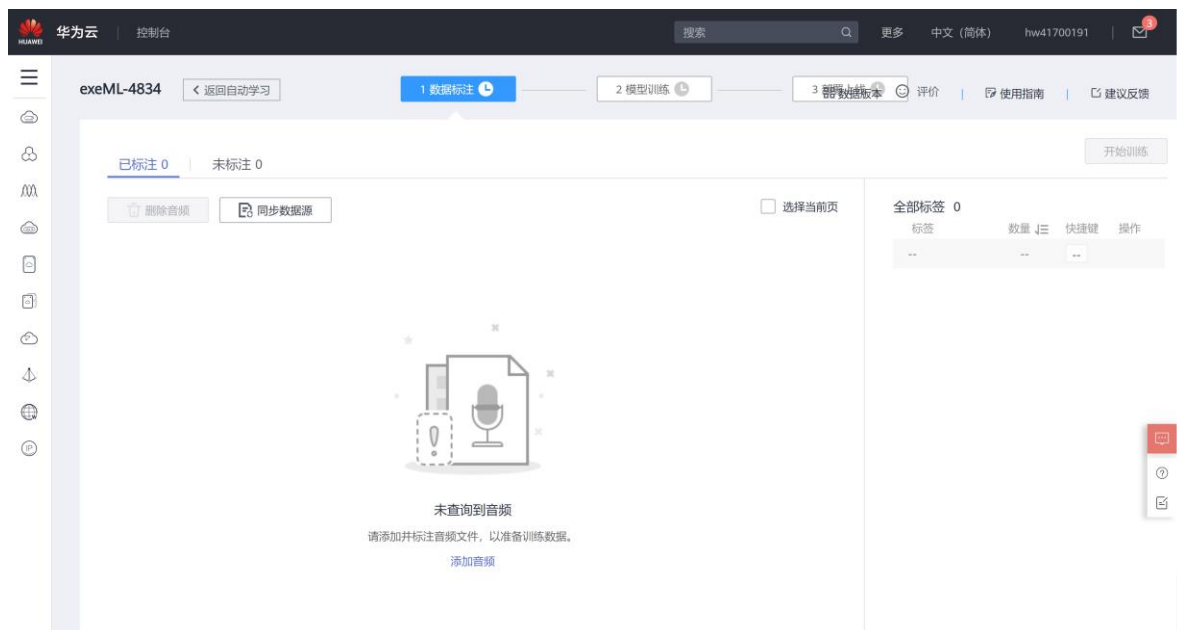


按照如下示例填写参数：名称、数据集名称、数据集输入位置、数据集输出位置。最后点击“创建项目”完成图像分类项目创建。







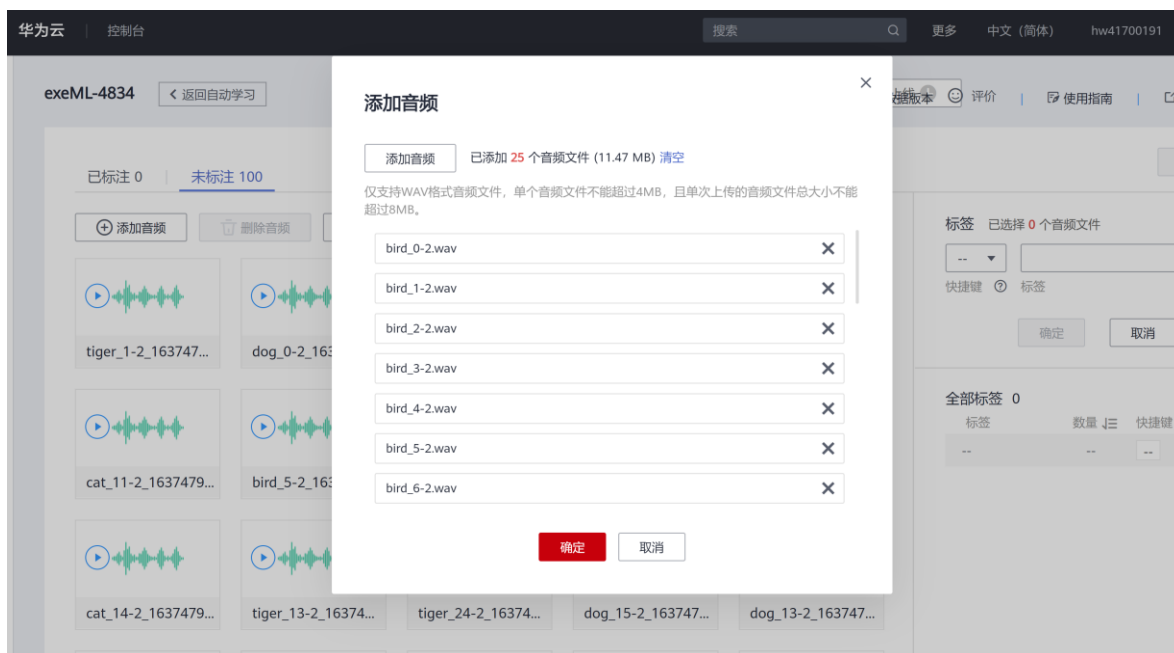


### (3) 动物叫声音频数据上传

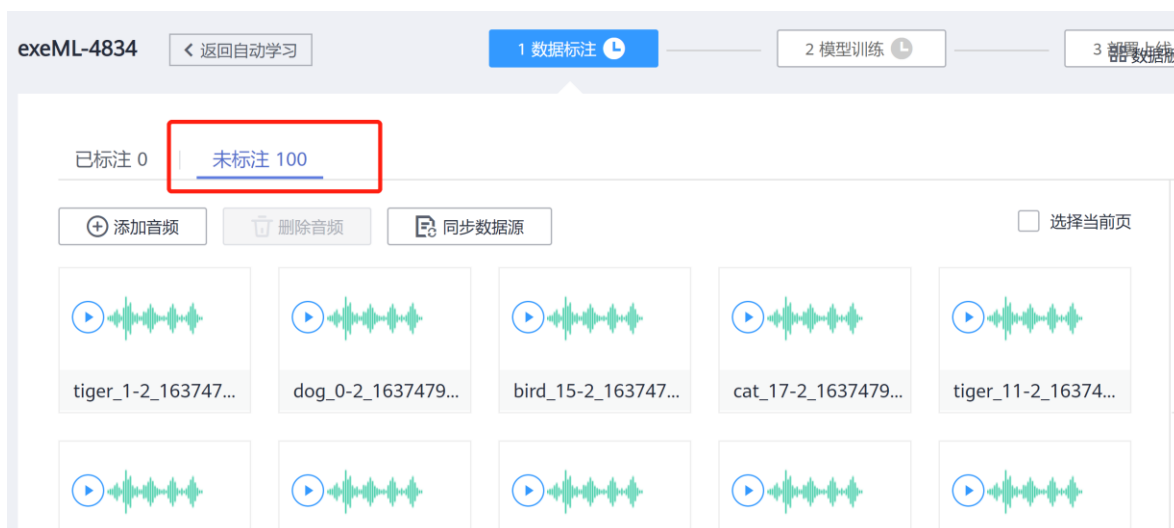
进入声音分类项目后，首先进入数据标注界面。点击下图中的“添加音频”按钮：



添加 sound\_classification\train 目录下任一个动物文件夹下数据，注意要求“仅支持 WAV 格式音频文件，单个音频文件不能超过 4MB，且单次上传的音频文件总大小不能超过 8MB。”，点击确定：

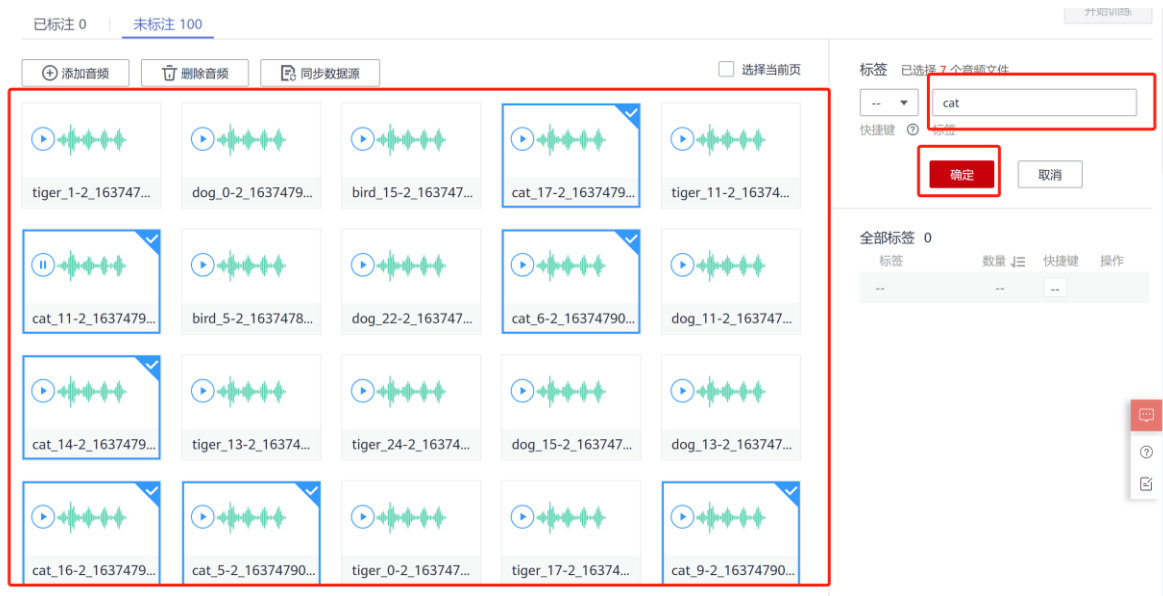


可以看到上一步选取的音频文件已全部上传并显示未标注，此处可以点击播放音频文件：

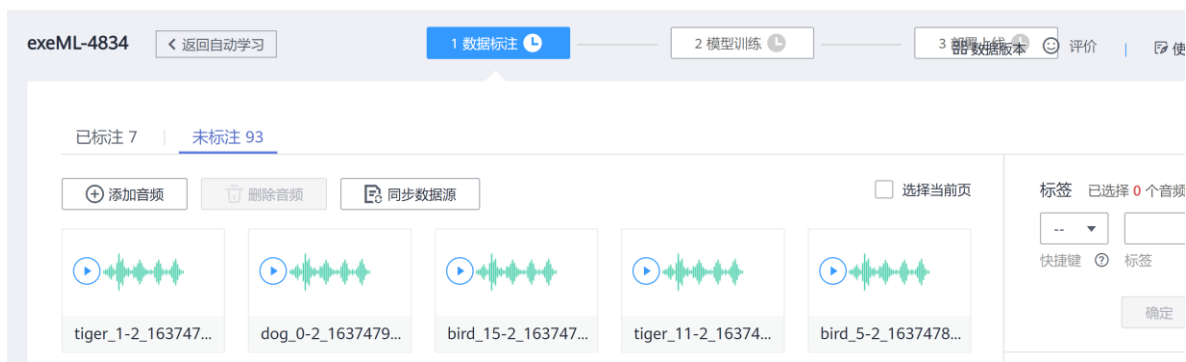


#### (4) 动物叫声音频数据标注

在“未标注”页面，批量选中相同类别的音频文件，在右侧添加标签名称，最后点击“确定”按钮。



随后可以看到刚才选中的音频文件已标注完成



按照同样的方法，将所有训练数据上传并进行标注，直到 100 个音频数据文件全部标注完，确认每类动物的数量为 25，可以点击右侧的标签来查看每个标签对应的文件。



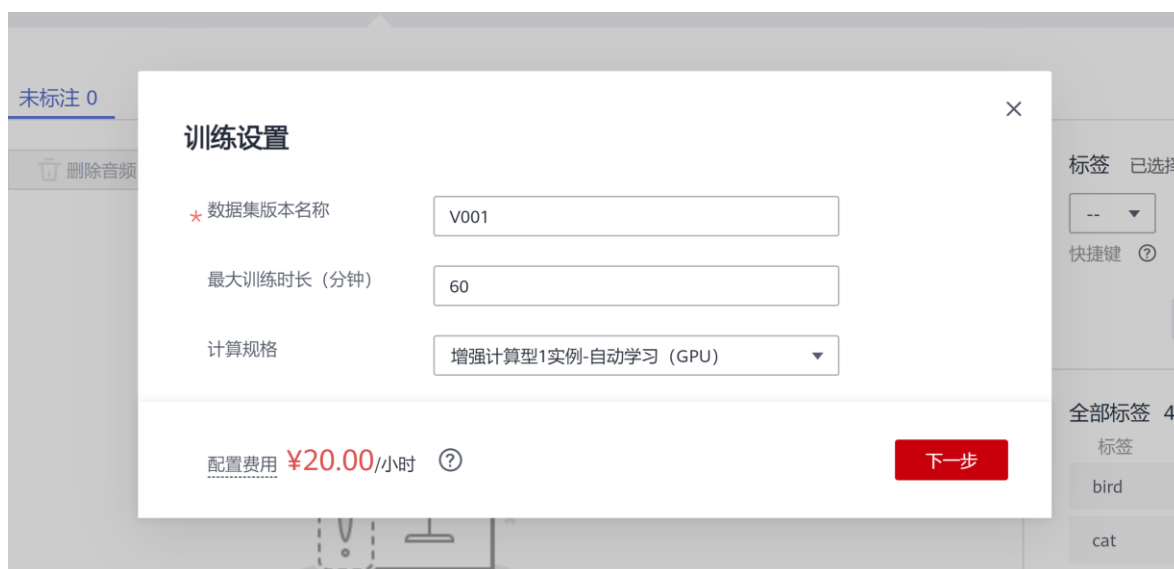
如果发现标签不正确，可以选中音频数据，重新选择标签。

### (5) 模型训练

点击右上角“开始训练”按钮，设置数据集版本名称，可以选择设置训练时长（不同任务类型和数据量需要的训练时长不一样），点击“确定”即可开始训练。



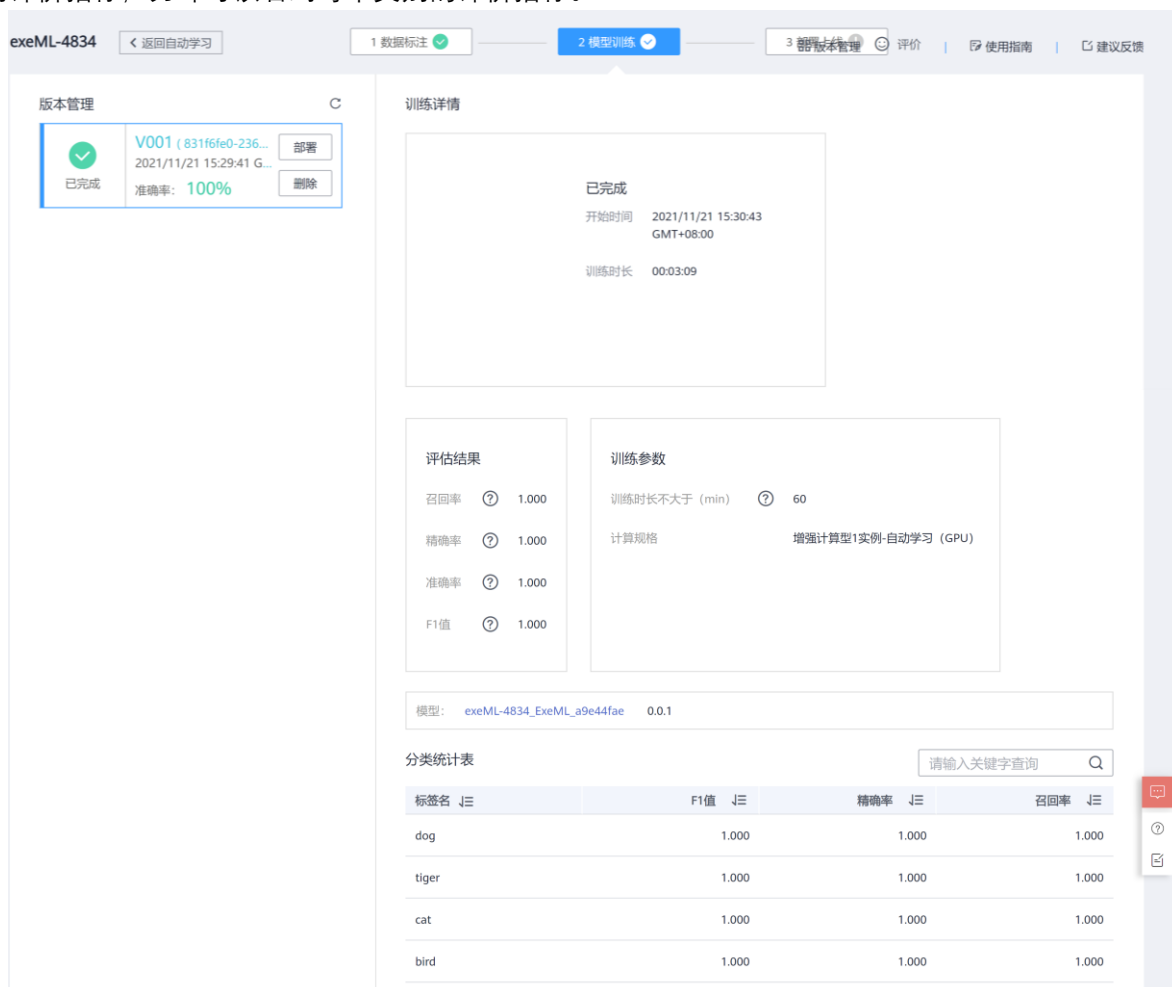
训练设置保持默认即可。



(3) 训练任务提交成功，自动跳转模型训练的界面。



待模型训练完成后，可以看到训练时长和训练结果，包括准确率、召回率、精确率、F1 值等常用评价指标，另外可以看到每个类别的评价指标。



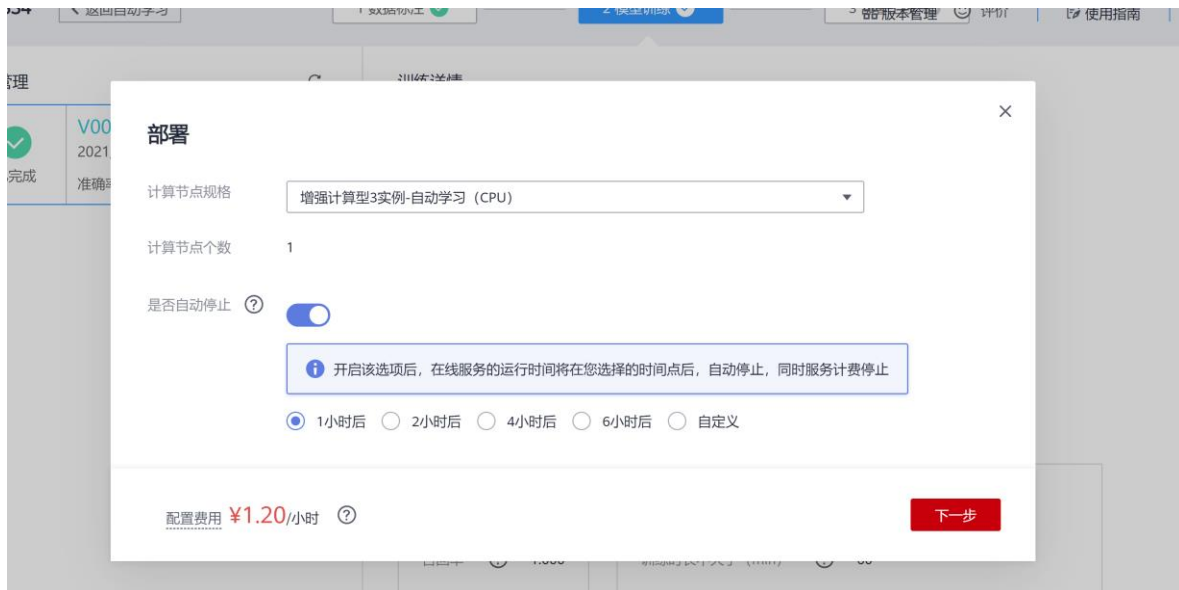
## (6) 模型部署

点击左侧“部署”按钮，将模型部署为一个在线服务：





选择节点计算规格，是否自动停止，点击确定开始部署模型。



部署任务提交成功，自动跳转部署上线的界面。

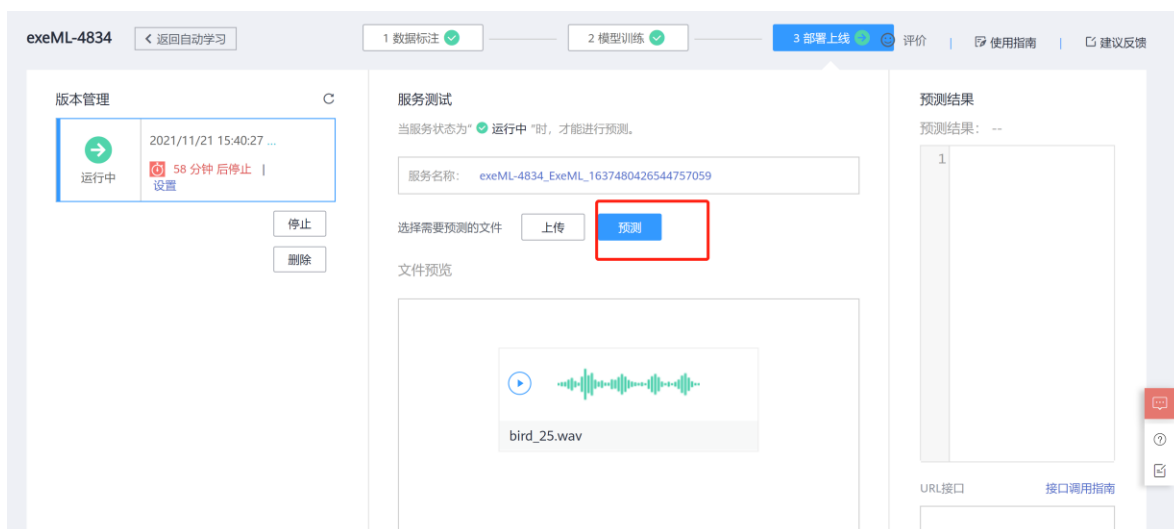
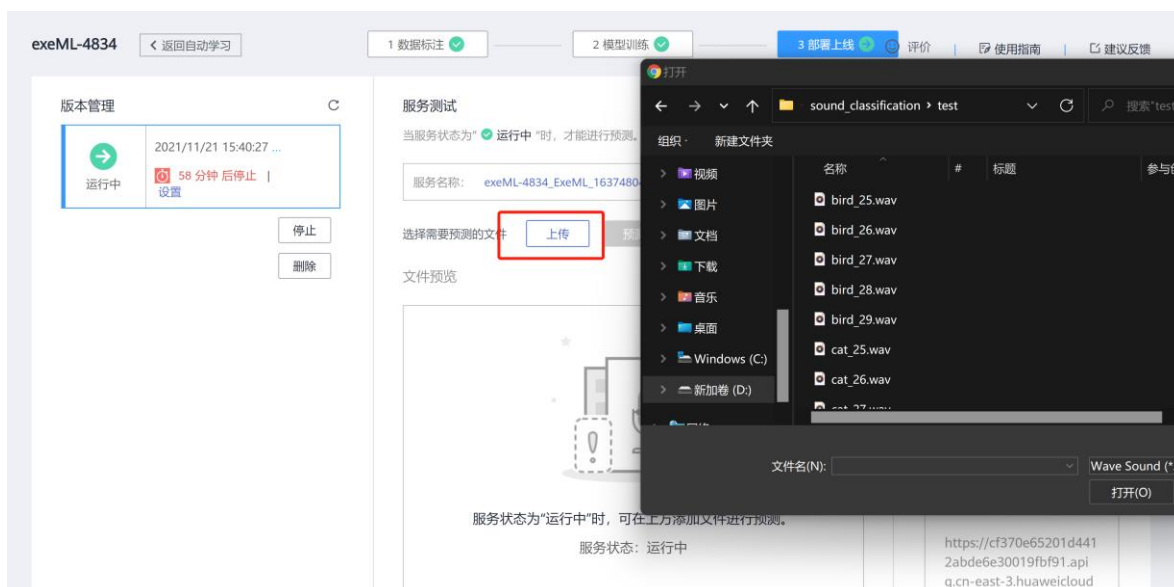


## (7) 服务测试

部署成功后，左侧显示服务状态为“运行中”，此时可以进行在线预测。



点击“上传”按钮，上传一张本地的测试文件（从测试集`test`目录中挑选）。上传的音频文件可以实时播放。点击“预测”按钮进行在线预测，右侧可以看到预测结果和 URL 接口。



版本管理

运行中

2021/11/21 15:40:27 ...

57 分钟后停止 | 设置

停止

删除

服务测试

当服务状态为“运行中”时，才能进行预测。

服务名称: exeML-4834\_ExeML\_1637480426544757059

选择需要预测的文件 上传 重新预测

文件预览

bird\_25.wav

预测结果

预测结果: bird

1

2 cted\_label": "bird",

3 s": 0.880055

URL接口 接口调用指南

https://cf370e65201d4412abde6e30019fbf91.apig.cn-east-3.huaweicloudapis.com/v1/infers/Saa7be0b-d55b-439a-90fa-08d4861858ba

同时，自动学习部署的服务可以在 ModelArts 控制台的“部署上线” -> “在线服务”中调用。

华为云 控制台 上海

搜索

更多 中文 (简体) hwy41700191

ModelArts

总览

自动学习

数据管理

开发环境

算法管理

训练管理

模型管理

部署上线

在线服务

批量服务

边缘服务

AI Gallery

专属资源池

全局配置

在线服务

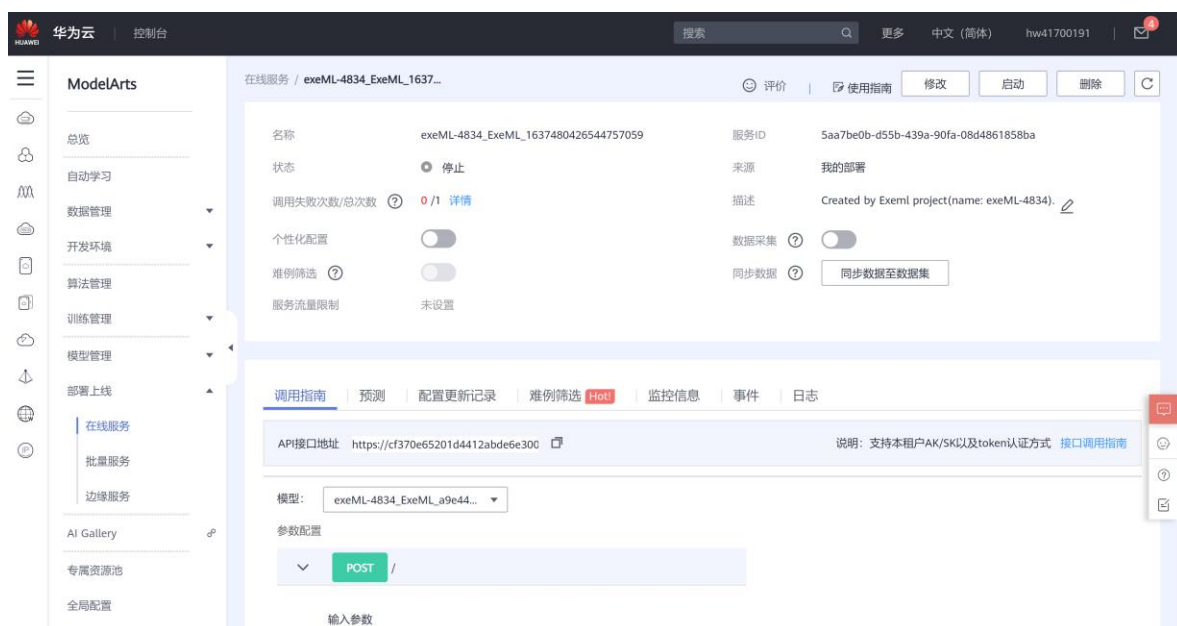
评价 使用指南 建议反馈

温馨提示: 状态为“运行中”、“告警”的服务正在产生费用，不使用时，请及时停止。若您已经设置了自动停止，请注意服务的剩余运行时间。

部署 删除 授权管理

全部 请输入名称查询

名称	状态	调用失败次数/总次数	来源	创建时间	描述	操作
exeML-4834_ExeML_...	停止	0 / 1	我的部署	2021/11/21 15:40:26...	Created by Exem...	修改 预测 启动 更多
exeML-add6_ExeML_...	停止	0 / 4	我的部署	2021/11/19 10:01:35...	Created by Exem...	修改 预测 启动 更多



进入“预测”，可以进行测试文件上传，上传的音频文件可以播放，进行在线预测，右侧可以看到预测结果。



下面是另外两类动物的在线预测结果，可以看到模型的预测效果比较好。

调用指南

预测

配置更新记录

难例筛选 Hot!

监控信息

事件

日志

请求路径: /

选择预测声音文件

上传

重新预测

难例反馈

预测结果显示

✓ 预测成功

```
1 {
2   "predicted_label": "tiger",
3   "score": 0.846516
4 }
```

tiger\_27.wav

调用指南

预测

配置更新记录

难例筛选 Hot!

监控信息

事件

日志

请求路径: /

选择预测声音文件

上传

重新预测

难例反馈

预测结果显示

✓ 预测成功

```
1 {
2   "predicted_label": "cat",
3   "score": 0.981295
4 }
```

cat\_29.wav

### 关闭在线服务

在使用完部署的服务后，需关闭服务，否则会持续收费。可以通过以下两种方式关闭服务：

1. 在“自动学习” -> “部署上线”中停止在线服务；
2. 在“部署上线” -> “在线服务”中停止在线服务。

当需要使用该在线服务的时候，可以重新启动该在线服务。

## 2.3 实验小结

本实验通过自动学习中的声音分类项目实现动物叫声分类，并将模型部署为在线服务进行预测，使学员零编码实现并熟悉了数据准备、创建项目、数据标注、自动训练、部署上线的全流程。

## 1. 连续语音识别