

# 机器学习PAI

## 上传数据方法说明

# 上传数据方法说明

## PAI上传数据说明

PAI底层支持两种数据源，一种是Maxcompute存储数据，另一种是OSS存储数据。Maxcompute主要用来存储表结构数据，支持稀疏与稠密两种格式的数据，支持PAI IDE上传以及命令行工具上传两种方式，这部分的数据对应平台中除了深度学习以外的算法组件。OSS数据源主要针对深度学习相关算法组件，可用来存储结构化或非结构化数据。

**注意：使用Maxcompute作为存储，建议当数据小于20M使用PAI IDE环境上传数据，当数据大于20M使用命令行工具上传。**

---

### 目录

IDE上传数据到Maxcompute

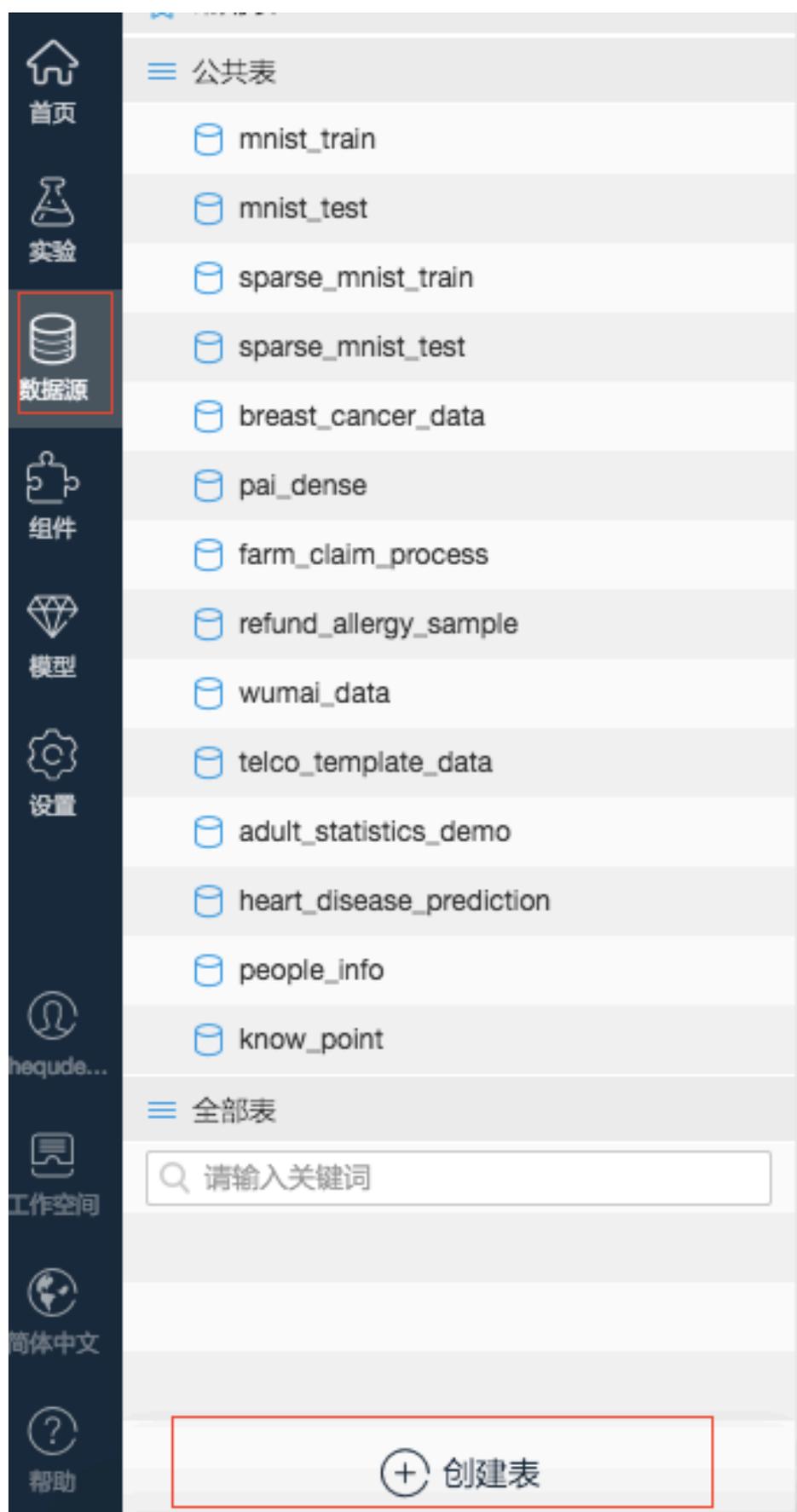
命令行工具上传数据到Maxcompute

OSS上传数据

---

## IDE端上传数据到Maxcompute

1.进入机器学习PAI，点击数据源，创建表：



2.选择相应的数据源，并且创建与之匹配的字段，建议使用txt格式上传，csv格式易出现特殊字符。

创建表 🔍 ✕

表结构 数据

注意：上传的数据将追加搭到原表，请上传 .txt 或 .csv 文件

选择文件

确定 关闭

对于稀疏格式数据，请下载以下文件供参考：libsvm数据使用

## 命令行工具上传数据到Maxcompute

Maxcompute提供多种数据上传功能，请下载以下参考资料，选择最合适的方式：数据迁移到Maxcompute的N种方式

## OSS上传数据

OSS上传数据，详情参见：<https://promotion.aliyun.com/ntms/act/ossdoclist.html>