

数据挖掘-自助机器学习

概述

- 概述
- 创建方法
 - 第一步：选择数据源
 - 第二步：算法设置
 - 第三步：特征设置
- 示例

机器学习的应用往往需要大量拥有专业知识的人进行人工干预，这些人工干预表现在：特征提取、模型选择、参数调节等机器学习的各个方面。AutoML可以将这些与特征、模型、优化、评价有关的重要步骤进行自动化地学习，使得机器学习模型无需人工干预即可被应用，也让更多的人可以在没有专业知识的情况下轻松使用。

创建方法

使用AutoML来创建包含算法的实验分为三个步骤：

第一步：选择数据源

选择实验需要使用的数据源、SCHEMA、表名。

第二步：算法设置

⌚ 设置
×

✔
②
③

选择数据源
算法设置
特征设置

算法选择 请选择 ▼

启用自动调参

上一步
下一步
取消

参数名称	说明
算法选择	分类算法 回归算法 聚类算法
启用自动调参	勾选该项，则系统自动调参数，不需要用户手工设置参数。 聚类算法没有此设置项。

第三步：特征设置

⌚ 设置
✕

✓
✓
③

选择数据源
算法设置
特征设置

特征选择

<input type="checkbox"/>	field	字段
<input type="checkbox"/>	airborne	INTEGER
<input checked="" type="checkbox"/>	animal_name	STRING
<input type="checkbox"/>	aquatic	INTEGER
<input type="checkbox"/>	backbone	INTEGER
<input type="checkbox"/>	breathes	INTEGER
<input type="checkbox"/>	catsize	INTEGER

标签列选择 ②

输出的特征数

上一步
完成
取消

参数名称	说明	备注
特征选择	选择需要的特征列，必须是数值列	必填（特征列中不能含有null）
标签列选择	选择做为标签列的字段	必填
输出的特征数	从待选择的特征列中输出特征列的数量	默认值为1，范围是[1, 已选择特征的数量]的整数

示例

使用AutoML快速创建分类算法实验，如下：

实验管理 | 服务管理 | 模型管理 | 模型自学习 | 模型批量预测 | 自定义模块管理
□

机器学习-可视化、零编程的挖掘建模，从小白到专家

实验
案例
刷新

名称	类型	描述	修改时间	状态	常用操作
数据挖掘	公有文件夹	数据挖掘	2020-12-10 14:54:52		
0_demo_演示	公有文件夹		2020-09-04 11:24:34		
6月7月新功能演示	公有文件夹		2020-07-30 15:51:40		
HDJ	公有文件夹		2020-04-03 09:26:07		
KK_testSmoke8	公有文件夹		2019-11-06 19:18:28		
SMS-2561	公有文件夹		2020-01-03 15:23:32		
carmen	公有文件夹		2020-12-02 10:48:19		
cp	公有文件夹		2020-11-23 18:49:16		
czd	公有文件夹		2020-12-03 11:38:03		