

机器学习-销售预测（操作指导）

1 任务介绍

本次任务将介绍如何使用**MLS内置模板**进行销售预测。

当餐厅经营规模变大或者影响餐厅销售的因素多种多样，只是单纯地靠人为推算餐厅销售量，容易出现偏差，大大降低了餐厅选址准确性和经营寿命。

现在您可以使用华为云机器学习服务提供的销售预测模板，省时省力得到餐厅未来3个月内的销售结果。作为餐厅经营人员根据预测结果，可以更好判断在新地段开设哪种类型餐厅，并把预测出来的销售量较高时间段（例如每年5~7月是餐厅旺季）作为餐厅经营周期，由此缩短资金投入前期，提高餐厅纯利润收入。

2 任务执行

2.1 数据理解

数据集的具体字段如下：

字段名	含义	类型	描述
Id	餐厅id	Int	餐厅标识
OpenDate	开业时间	String	餐厅开业时间
City	城市名	String	餐厅所在城市
CityGroup	城市类型	String	餐厅所在城市类型（大型、其他）
Type	类型	String	餐厅类型（流动餐点、汽车餐厅、一般餐厅）
P1, P2 - P37	其他信息	Real	人口、房地产、商业数据（1-5评分）
revenue	金额	Real	餐厅营业额

数据示例：

Id	OpenDate	City	CityGroup	Type	P1	P2	...	P36	P37	revenue
0	07/17/1999	istanbul	Big Cities	IL	4	5	...	3	4	5653753
1	02/14/2008	Ankara	Big Cities	FC	4	5	...	0	0	6923131
2	03/09/2013	Diyarbakir	Other	IL	2	4	...	0	0	2055379

Id	OpenDate	City	CityGroup	Type	P1	P2	...	P36	P37	revenue
3	02/02/2012	Tokat	Other	IL	6	4.5	...	12	6	2675511
4	05/09/2009	Gaziantep	Other	IL	3	4	...	3	3	4316715

2.2 建模

使用销售预测的模板创建项目

- 1) 如《2.1 机器学习实战-销售预测-学习资料》最后所示，在“销售预测”模板上单击“+”，命名一个自己的项目名称。

创建项目

★ 项目名称

mySalePredict

描述

公开项目

☒

确定

取消

- 2) 单击“确定”后，直接进入项目的“train”工作流。

如果没有直接进入train工作流，是因为浏览器拦截了弹窗。

在实例主页左侧点击“项目”



找到之前创建的销售预测项目，点击名称进入。

点击“工作流”，再点击“train”进入train工作流。

训练模型的工作流如下图所示：



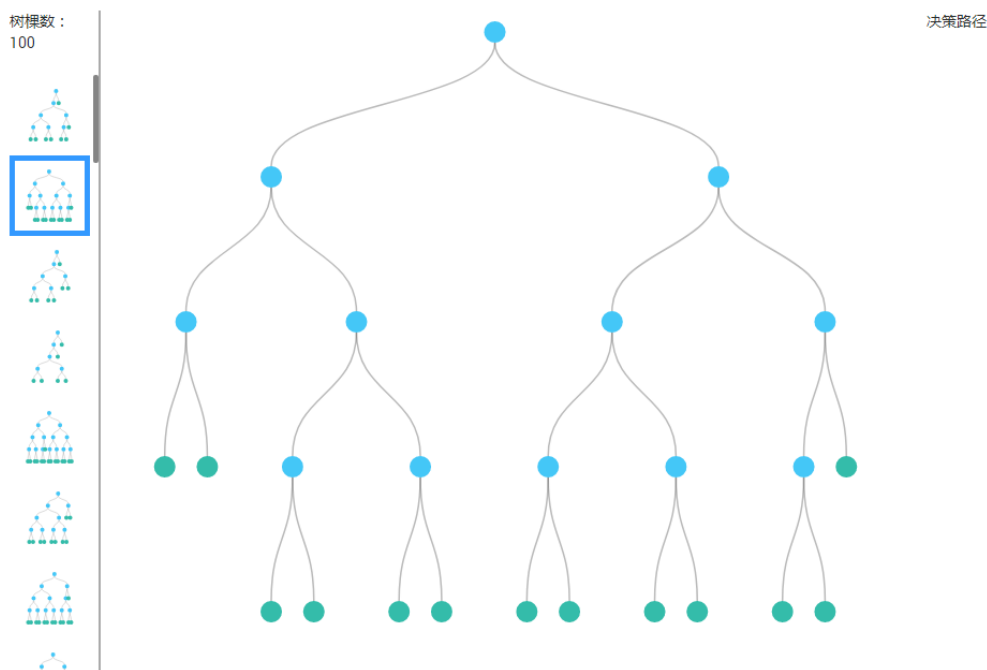
- 3) 单击 运行工作流，在下方“运行日志”出查看运行状态。

运行日志

2018/09/13 15:38:54 GMT+08:00 工作流运行开始...
2018/09/13 15:38:54 GMT+08:00 节点[读取餐厅训练数据]开始运行...
2018/09/13 15:39:04 GMT+08:00 节点[读取餐厅训练数据]运行成功。
2018/09/13 15:39:04 GMT+08:00 节点[修改元数据]开始运行...
2018/09/13 15:39:09 GMT+08:00 节点[修改元数据]运行成功。
2018/09/13 15:39:09 GMT+08:00 节点[拆分]开始运行...

- 4) 当“运行日志”显示工作流运行成功之后，单击“随机森林回归”节点，右键选择“查看模型”，查看模型可视化界面。

查看模型



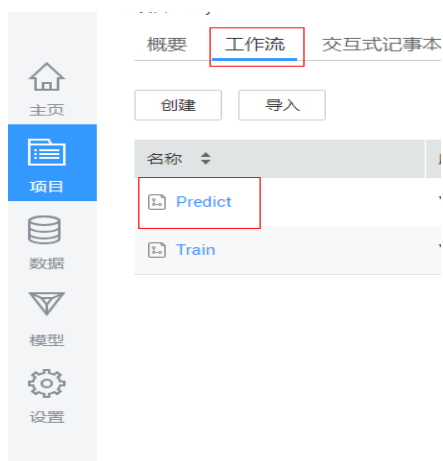
- 5) 单击“回归模型评估”，右键选择““输出数据集”数据预览”，可以

查看模型的评估结果。回归模型的评估值为MAE（平均绝对误差）、MSE（均方误差）、RMSE（均方根误差）。上述3个误差值是指评分真实值与预测值之间的误差，在多次建模的过程中，每一次建模结果都会产生一组误差值。评判一个回归模型好坏的方法就是看这三个误差值是否变小或者变大，误差值越小表示回归模型越好。

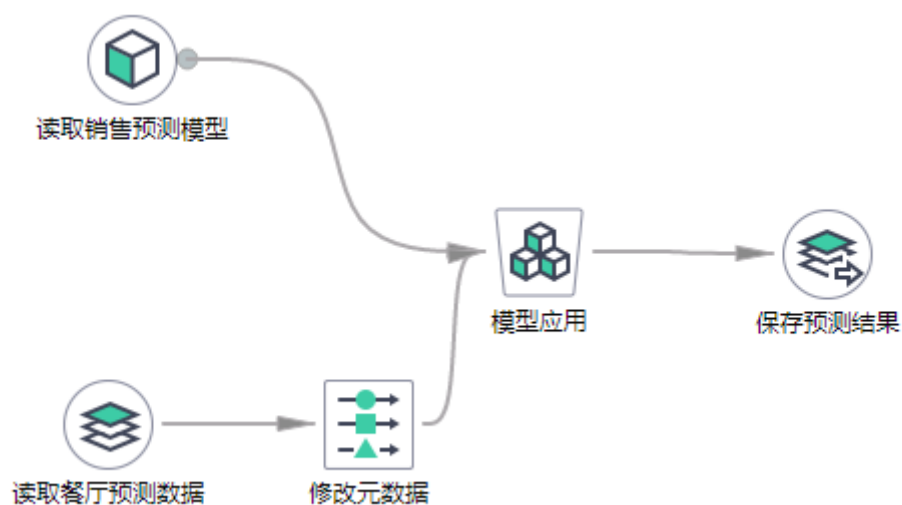
2.3 预测


模型建立完成后，使用已经保存好的模型和某商城的现阶段数据，即可帮您预测销售额。

- 1) 回到MLS主页，在页面左侧单击“项目”，单击自己的销售预测项目，再单击“工作流”。



单击“predict”，打开预测工作流，如下图。



- 2) 单击“读取销售预测模型”节点，在窗口右侧的节点配置中单击，在弹出的“选择模型”窗口，单击以“salesFore”命名的模型，单击“确定”。

选择模型

已选项：

通过名称搜索




名称 ▾	版本号 ▾	格式	修改时间 ▴
salesFore	20180913_153919_929	pmml	2018/09/13 15:39:22 GMT+08...

读取销售预测模型

* 模型名称:

salesFore.20180601_1

...

- 3) 单击 运行工作流，并查看日志。

运行日志

2018/09/13 15:48:21 GMT+08:00 工作流运行开始...

2018/09/13 15:48:21 GMT+08:00 节点[读取销售预测模型]开始运行...

2018/09/13 15:48:26 GMT+08:00 节点[读取销售预测模型]运行成功。

2018/09/13 15:48:26 GMT+08:00 节点[读取餐厅预测数据]开始运行...

2018/09/13 15:48:46 GMT+08:00 节点[读取餐厅预测数据]运行成功。

2018/09/13 15:48:46 GMT+08:00 节点[修改元数据]开始运行

- 4) 单击“模型应用”节点，右键选择““输出数据集”数据预览”，即可在“数据预览”中查看预测结果。MLS的数据可视化还提供数据统计的功能。

数据预览

列：43

p33	p34	p35	p36	p37	predictioncol
5	5	0	5	3	7200013.311016648
0	0	0	0	0	4153434.665205604
0	0	0	0	0	5882232.514092522
6	18	5	0	0	5887182.993633262
0	0	0	0	0	3922526.7142258347
1	1	5	3	1	6973487.705486281
0	3	0	0	0	4687527.70180244
0	0	0	0	0	4921617.176737852
3	0	4	0	0	4600490.178568352
0	0	0	0	0	6040536.827410154
2	3	4	4	3	5858004.469768684
0	0	0	0	1	5886116.454459011
2	3	4	4	1	4198100.728197841
0	0	4	0	0	4580390.544032702

^元数据

类型： Real 角色： Input

测量尺度： Continuous 值： (.)

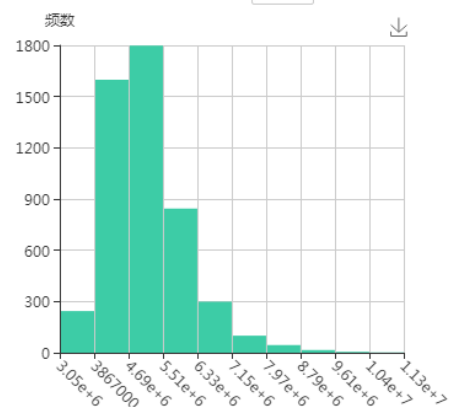
^统计

最大值： 1.1259e7 最小值： 3045961.4397

平均值： 5097065.8376 值个数： 4951

标准偏差： 937136.2935

^可视化

☐ 频数 ☐ 对数尺度变换 区间数 10


3 打卡任务

predict工作流运行成功后进行截图：

- 1、 右上角是用户名
- 2、 工作流的样式和图示相同
- 3、 下方有“工作流运行成功”字样



运行日志

2018/09/13 15:48:51 GMT+08:00 节点[模型应用]开始运行...
2018/09/13 15:48:51 GMT+08:00 节点[模型应用]运行成功...
2018/09/13 15:49:11 GMT+08:00 节点[保存预测结果]开始运行...
2018/09/13 15:49:36 GMT+08:00 节点[保存预测结果]运行成功...
2018/09/13 15:49:36 GMT+08:00 工作流运行成功。