

| 预览数据 | | | | | |
|------|----|----|----|----|---------------|
| | qs | wl | jb | cg | predict_label |
| | 浊响 | 清晰 | 凹陷 | 硬骨 | 是 |
| | 沉闷 | 清晰 | 凹陷 | 硬骨 | 是 |
| | 浊响 | 清晰 | 凹陷 | 硬骨 | 是 |
| | 沉闷 | 清晰 | 凹陷 | 硬骨 | 是 |
| | 浊响 | 清晰 | 凹陷 | 硬骨 | 是 |
| | 浊响 | 清晰 | 稍凹 | 软粘 | 是 |
| | 浊响 | 稍糊 | 稍凹 | 软粘 | 否 |
| | 浊响 | 清晰 | 稍凹 | 硬骨 | 是 |

共 17 条 25 条/页 < 1 > 前往 1 页

图 309

3.4.5.6 CART 决策树

图标:



描述: CART(Classification And Regression Tree)算法是一种决策树分类方法。它采用一种二分递归分割的技术,分割方法采用基于最小距离的基尼指数估计函数,将当前的样本集分为两个子样本集,使得生成的每个非叶子节点都有两个分支。因此,CART 算法生成的决策树是结构简洁的二叉树。

字段属性

特征列: 通过勾选的方式选择特征所在列。

标签列: 选择分类标签所在的列,请选择字符型数据。

参数设置

切分时的评价准则: 包括 Gini 系数、熵,默认 Gini 系数。

切分原则: 包括选择最优的切分、随机切分,默认选择最优的切分。

输出

表结果: CART 决策树算法分类结果。

报告: Confusion Matrix、Receiver Operating Characteristic(ROC)、Precision-Recall

示例

下列对某数据进行 CART 决策树算法分类:

- 选择特征数列，标签列。如图 310 所示。
- 保留默认参数，切分时的评价准则为 Gini 系数，切分原则为最优的切分，如图 311 所示。
- CART 决策树运行成功后，可选择查看结果，如图 312 所示。
- CART 决策树运行成功后，可选择查看报告，如图 313 所示。
- 模型评估配置如图 314 所示。
- 模型评估运行成功后，选择查看数据，如图 315 所示。
- 模型评估运行成功后，选择查看报告，如图 316 所示。
- 模型预测配置如图 317 所示。
- 模型预测运行成功后，选择查看数据，如图 318 所示。



图 310

> 字段属性

∨ 参数设置

切分时的评价准则 ?

Gini系数 ▾

切分原则 ?

选择最优的切分 ▾

图 311

预览数据

| wifi | e_pay | package_type | leave | predict_label |
|------|-------|--------------|-------|---------------|
| 0 | 1 | 2 | 1 | 1 |
| 28 | 1 | 4 | 1 | 1 |
| 0 | 0 | 3 | 1 | 1 |
| 0 | 0 | 3 | 0 | 0 |
| 0 | 0 | 3 | 0 | 0 |
| 0 | 1 | 1 | 0 | 0 |
| 0 | 1 | 2 | 1 | 1 |
| 0 | 0 | 2 | 0 | 0 |

共 597 条 25 条/页 < 1 2 3 4 5 6 ... 24 > 前往 1 页

图 312

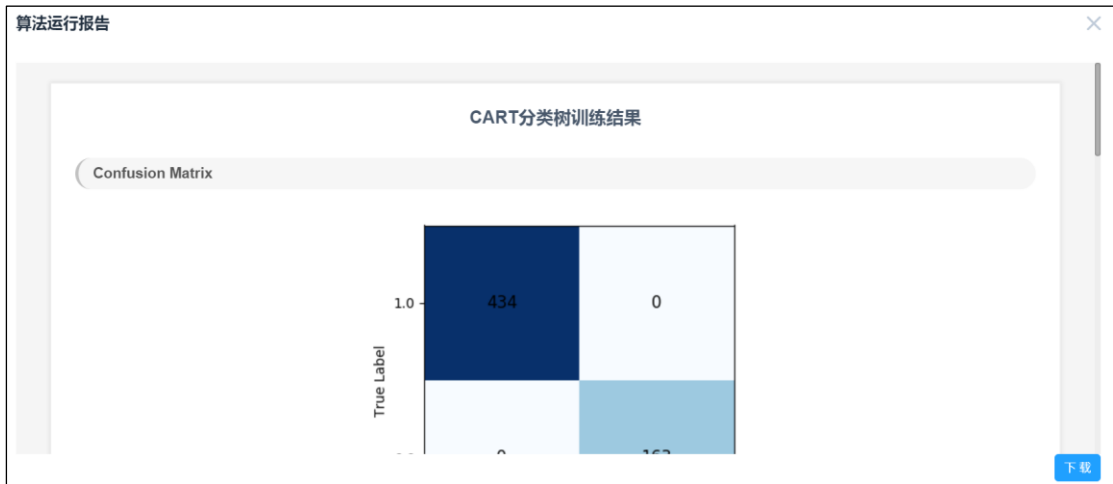


图 313



图 314

预览数据

| | wifi | e_pay | package_type | leave | predict_label |
|--|------|-------|--------------|-------|---------------|
| | 0 | 0 | 3 | 0 | 0 |
| | 0 | 0 | 1 | 0 | 0 |
| | 0 | 1 | 2 | 1 | 0 |
| | 33 | 0 | 4 | 0 | 0 |
| | 0 | 0 | 3 | 0 | 1 |
| | 28 | 1 | 1 | 0 | 0 |
| | 0 | 1 | 4 | 0 | 1 |
| | 0 | 1 | 2 | 0 | 0 |

共 199 条 25 条/页 < 1 2 3 4 5 6 ... 8 > 前往 1 页

图 315

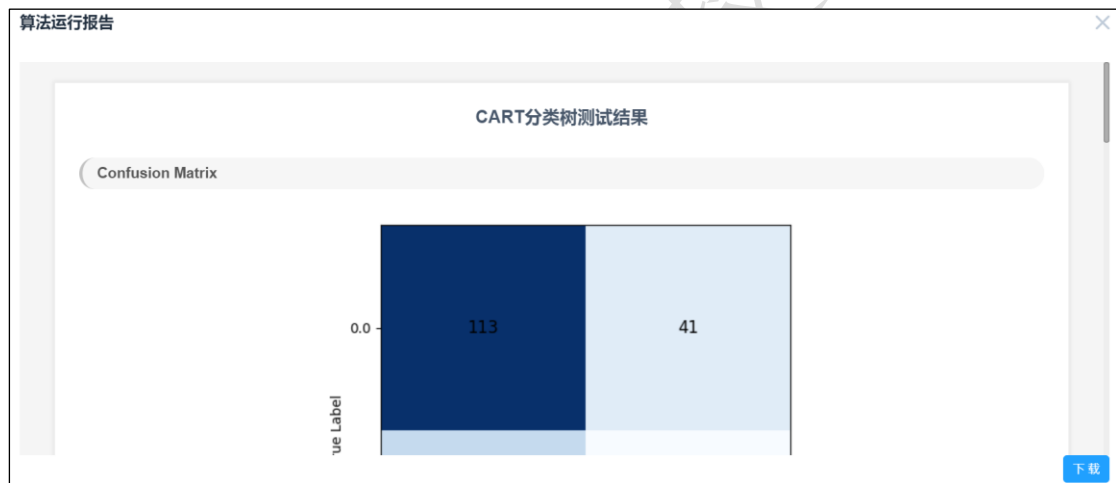


图 316

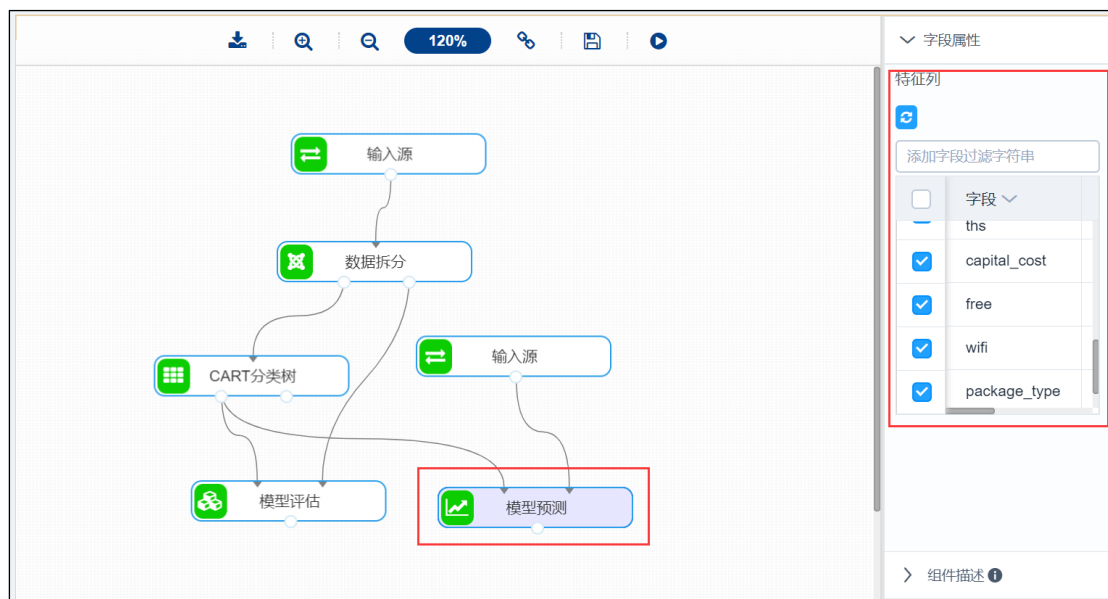


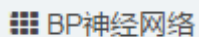
图 317

| onths | capital_cost | free | wifi | package_type | predict_label |
|-------|--------------|------|------|--------------|---------------|
| 5 | 5 | 0 | 0 | 1 | 0 |
| 7 | 7 | 22 | 0 | 3 | 0 |
| 9 | 9 | 0 | 0 | 1 | 0 |
| 16 | 16 | 46 | 61 | 4 | 0 |
| 4 | 4 | 0 | 0 | 1 | 0 |
| 5 | 5 | 0 | 0 | 2 | 0 |
| 16 | 16 | 30 | 0 | 3 | 0 |
| 7 | 7 | 0 | 38 | 4 | 0 |

图 318

3.4.5.7 BP 神经网络

图标:



描述: BP(back propagation)神经网络是一种按照误差逆向传播算法训练的多层前馈神经网络。

字段属性

特征列: 必选。通过勾选的方式选择特征所在列。

标签列: 必选。选择分类标签所在的列，仅支持字符型数据。