

预览数据

wifi	e_pay	package_type	leave	predict_label
0	0	1	1	4
0	0	3	0	2
0	0	3	0	2
0	0	3	0	4
0	0	2	0	1
0	1	1	1	4
0	0	3	0	2
0	0	1	0	2

共 796 条 25 条/页 < 1 2 3 4 5 6 ... 32 > 前往 1 页

图 285

3.4.5.3 KNN

图标:



描述: KNN 算法的核心思想是如果一个样本在特征空间中的 k 个最相邻的样本中的大多数属于某一个类别, 则该样本也属于这个类别, 并具有这个类别上样本的特性。

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, 一般为字符型数据。

参数设置:

最近邻个数 K : 整数型, 通常不大于 20, 默认 5。

投票权重类型: 权重相等或权重与距离成反比, 默认权重相等。

计算最近邻的算法: 包括 自动、BallTree、KDTree、暴力搜索法, 默认自动。

输出

表结果: KNN 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 KNN 算法分类:

- 选择特征数列, 标签列, 如图 286 所示。

- 保留默认参数，最近邻个数 K 为 5，投票权重类型为权重相等，暴力搜索为自动，如图 287 所示。
- KNN 运行成功后，可选择查看结果，如图 288 所示。
- KNN 运行成功后，可选择查看报告，如图 289 所示。
- 模型评估配置如图 290 所示。
- 模型评估运行成功后，选择查看数据，如图 291 所示。
- 模型评估运行成功后，选择查看报告，如所示。
- 模型预测配置如图 292 所示。
- 模型预测结果如图 292 所示。

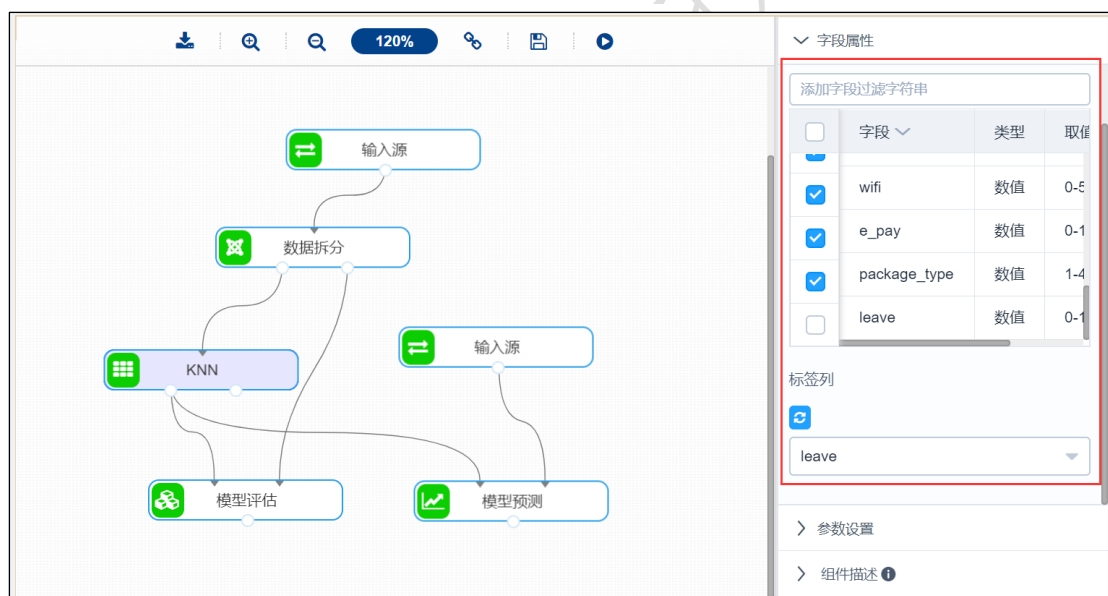


图 286

> 字段属性

√ 参数设置

K值

5 - +

投票权重类型

权重相等 ▼

计算最近邻的算法

自动 ▼

图 287

预览数据 ×

	wifi	e_pay	package_type	leave	predict_label
	0	1	2	1	0
	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	0
	0	0	2	0	0

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

图 288

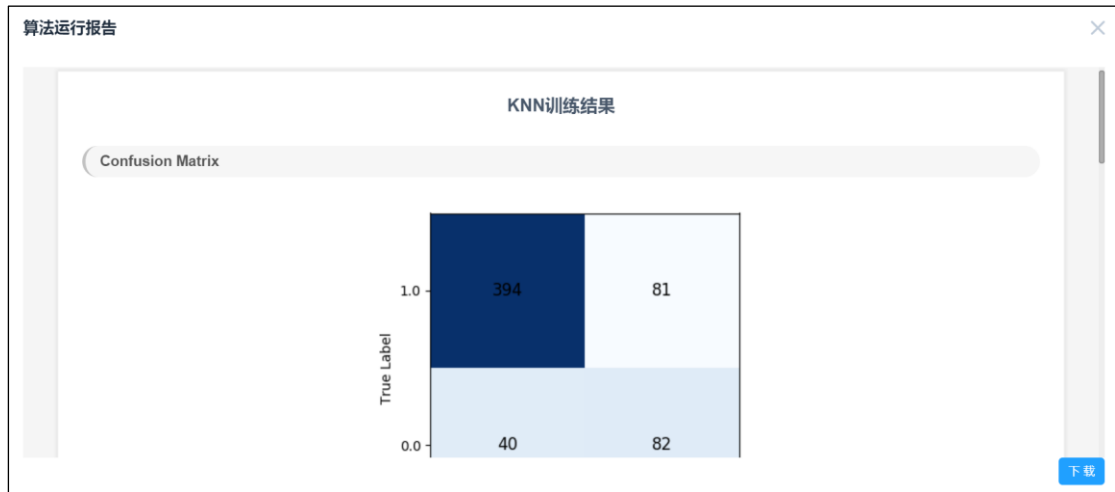


图 289

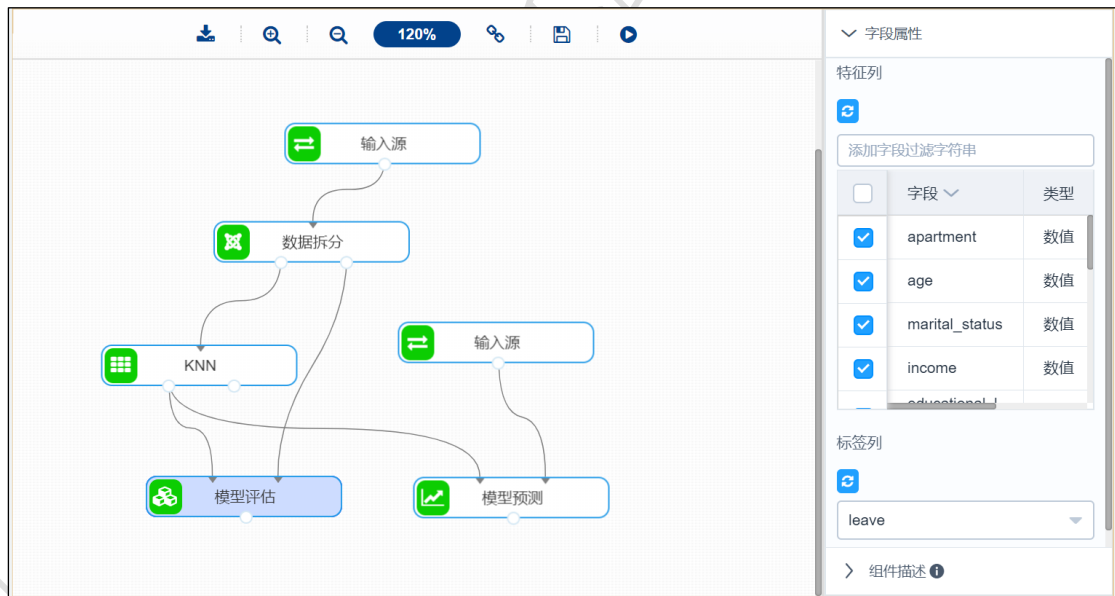


图 290

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	1
	33	0	4	0	1
	0	0	3	0	0
	28	1	1	0	1
	0	1	4	0	0
	0	1	2	0	0

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

图 291

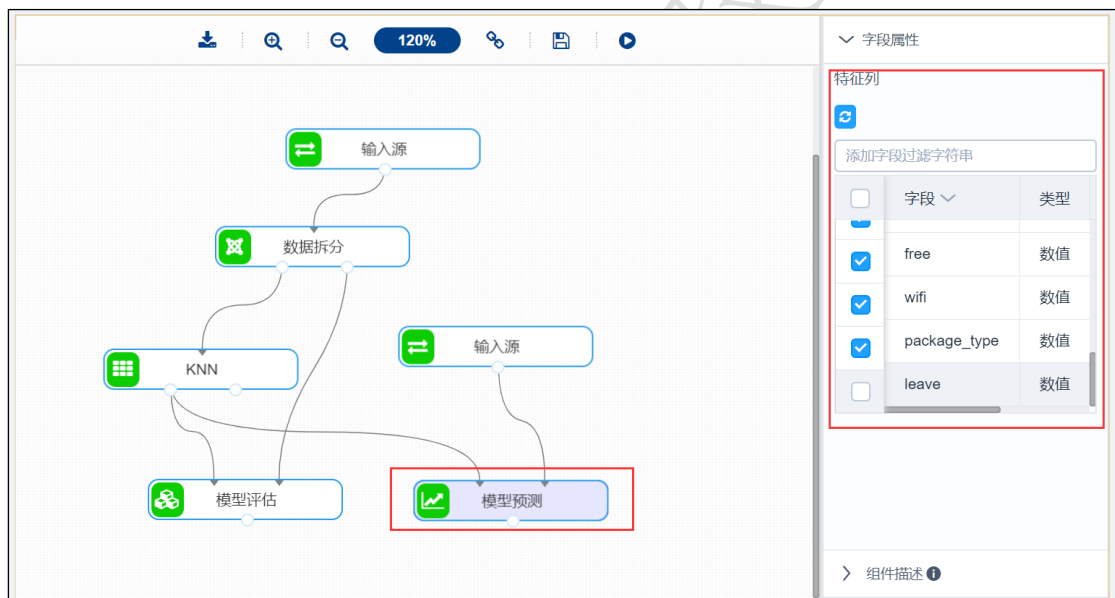


图 292

预览数据

onths	capital_cost	free	wifi	package_type	predict_label
	4	0	0	1	0
	6	0	0	3	1
	12	19	0	3	1
	10	29	0	3	1
	24	0	0	2	0
	5	0	0	1	1
	7	22	0	3	1
	9	0	0	1	1

共 796 条 25 条/页 < 1 2 3 4 5 6 ... 32 > 前往 1 页

图 293