

数据挖掘-关联规则生成

 在V9.7版本中，数据挖掘新增关联规则生成节点。

- 1. 功能概述
- 2. 功能输出说明
- 3. 应用示例
 - 3.1 场景说明
 - 3.2 操作步骤

关联规则致力于挖掘隐藏在海量数据中有趣联系，已被广泛应用于各个商业领域。本文将介绍如何基于FP-Growth关联算法输出关联规则和相关分析指标，以此满足用户不同的关联分析场景，支持企业进行科学地营销决策。

1. 功能概述

关联规则生成，基于FP-Growth 训练后的模型输入后，能够输出详尽的关联规则分析指标： 频繁项集、支持度、频率、 置信度、提升度、前项、后项及其数量。



2. 功能输出说明

当前显示 18 条 / 总共有 18 条数据

频繁项集	频率	支持度	前项	后项	置信度	提升度	前项数量	后项数量
<i>A</i> Items	# freq	# support	<i>A</i> antecedent	<i>A</i> consequent	# confidence	# lift	# antecedent_quantity	# consequent_quantity
冻肉,啤酒	120	0.182370820668693	冻肉	啤酒	0.5529953917050692	1.7578307620383358	1	1
冻肉,啤酒	120	0.182370820668693	啤酒	冻肉	0.5797101449275363	1.7578307620383358	1	1
蔬菜罐头,冻肉,啤酒	107	0.16261398176291794	蔬菜罐头,冻肉	啤酒	0.823076923076923	2.6163507989594943	2	1
蔬菜罐头,冻肉,啤酒	107	0.16261398176291794	蔬菜罐头,啤酒	冻肉	0.8770491803278688	2.6594394500264413	2	1
蔬菜罐头,冻肉,啤酒	107	0.16261398176291794	冻肉,啤酒	蔬菜罐头	0.8916666666666667	2.6790715372907155	2	1
蔬菜罐头,冻肉,啤酒	107	0.16261398176291794	蔬菜罐头	冻肉,啤酒	0.4885844748858447	2.6790715372907155	1	2
蔬菜罐头,冻肉,啤酒	107	0.16261398176291794	冻肉	蔬菜罐头,啤酒	0.4930875576036866	2.659439450026441	1	2
蔬菜罐头,冻肉,啤酒	107	0.16261398176291794	啤酒	蔬菜罐头,冻肉	0.5169082125603864	2.6163507989594943	1	2
蔬菜罐头,啤酒	122	0.18541033434650456	蔬菜罐头	啤酒	0.5570776255707762	1.7708071382877815	1	1
蔬菜罐头,啤酒	122	0.18541033434650456	啤酒	蔬菜罐头	0.5893719806763285	1.7708071382877815	1	1
鱼类,水果	100	0.1519756838905775	鱼类	水果	0.4830917874396135	1.5736356244320082	1	1
鱼类,水果	100	0.1519756838905775	水果	鱼类	0.49504950495049505	1.5736356244320084	1	1
蔬菜罐头,冻肉	130	0.19756838905775076	蔬菜罐头	冻肉	0.593607305936073	1.7999705405803506	1	1
蔬菜罐头,冻肉	130	0.19756838905775076	冻肉	蔬菜罐头	0.5990783410138248	1.7999705405803503	1	1
糖果,白酒	94	0.14285714285714285	糖果	白酒	0.48704663212435234	1.6605009530457193	1	1
糖果,白酒	94	0.14285714285714285	白酒	糖果	0.48704663212435234	1.6605009530457193	1	1
蔬菜罐头,白酒	72	0.1094224924012158	蔬菜罐头	白酒	0.3287671232876712	1.1208744410533038	1	1
蔬菜罐头,白酒	72	0.1094224924012158	白酒	蔬菜罐头	0.37305699481865284	1.120874441053304	1	1

表头真名 ☒ 表头别名 提示：点击单元格可查看超出的内容。注意：表头中  表示特征列，* 表示标签列

下载预览数据

序号	名称	说明		业务意义
1	关联规则	关联规则形如X→Y的表达式（其中X，Y是不相交的项集），例如{牛奶，尿布}→{啤酒}。		关联规则暗示两个物品之间可能存在很强的关系，通俗来说，就是如果 X 发生了，那么 Y 也很可能会发生。
2	前项 Antecedent 后项 Consequent	对于关联规则X→Y，X称为前项，Y称为后项。	前项数量 Antecedent_quantity 后项数量 Consequent_quantity	例如在{蔬菜罐头，冻肉}→{啤酒}关联规则中，有2个前项，1个后项，属于多对一的关联形式。

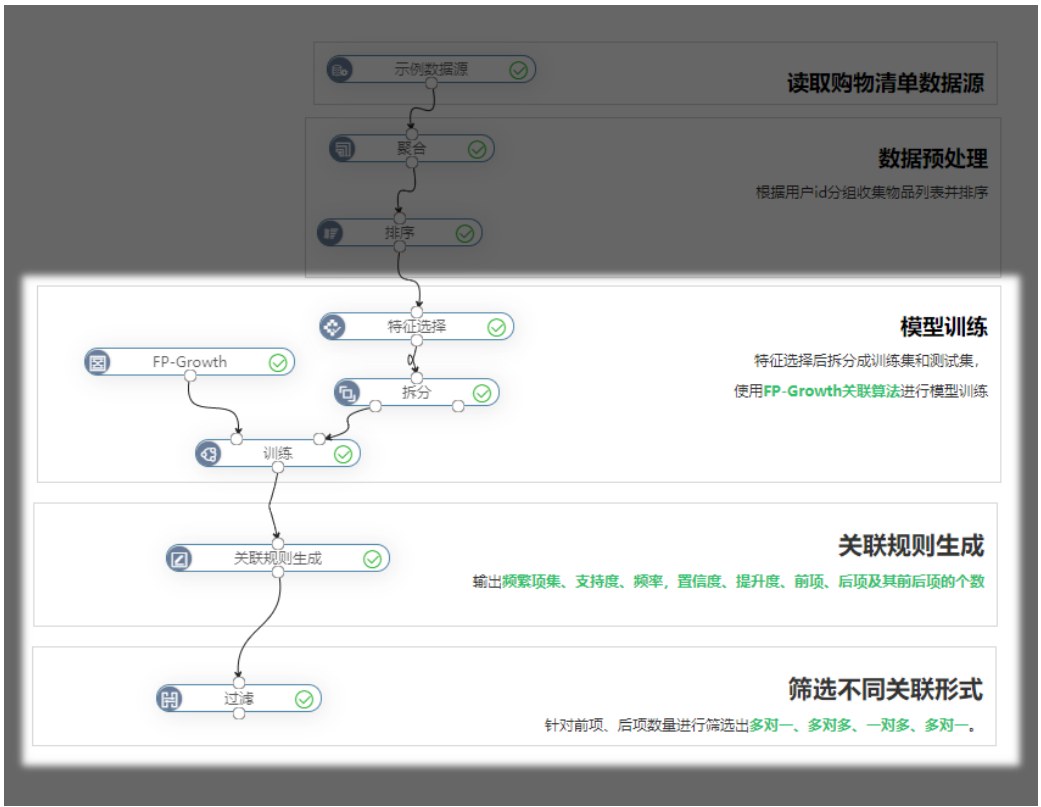
3	提升度 Lift	提升度是指含有X的条件下同时含有Y的概率，与无论是否含X含有Y的概率之比。	提升度>1且越高表明正相关性越高，提升度<1且越低表明负相关性越高，提升度=1表明没有相关性，即相互独立。	提升度反映了关联规则中的X与Y的相关性，提升度可用来判断商品组合方式是否具有实际价值。	在实际零售场景中，除了通过支持度、置信度来看待商品组合，发现高价值关联商品，同时也需要考虑提升度，若是提升度小于1，提升效果有限，商家便可以把精力花在提升度大于1的商品组合
---	----------	---------------------------------------	---	---	--

关于频繁项集、支持度、置信度的说明，点击 [此处](#) 可查阅。

3. 应用示例

3.1 场景说明

我们可以基于产品内置的购物篮分析案例，通过发现在交易数据库中不同商品之间的关联规则，有助发现顾客购买行为，指导企业进行科学的商业活动。



- 如下图的输出数据中，{冻肉，啤酒} → {蔬菜罐头} 这组关联商品的置信度、提升度最高，说明冻肉、啤酒和蔬菜罐头出现在同一购物篮的概率非常高（可信度达到89%），说明买了冻肉和啤酒的用户，很有可能购买蔬菜罐头。
- 商家可以考虑采取特殊的市场推广活动或手段（给买冻肉和啤酒、蔬菜罐头的客户打折）、货架设计（蔬菜罐头放在冻肉和啤酒相邻位置）等。
- 当然还有很多其他结论，可以帮助我们对商品的销售进行分析和处理，有利于提高商品销量，实现数据价值。

⌵ 当前显示 18 条 / 总共有 18 条数据

A_1 items	# freq	# support	A_2 antecedent	A_2 consequent	# confidence	# lift	# antecedent_quantity	# consequent_quantity
蔬菜罐头,白酒	72	0.1094224924012158	蔬菜罐头	白酒	0.3287671232876712	1.1208744410533038	1	1
蔬菜罐头,白酒	72	0.1094224924012158	白酒	蔬菜罐头	0.37305699481865284	1.120874441053304	1	1
鱼类,水果	100	0.1519756838905775	鱼类	水果	0.4830917874396135	1.5736356244320082	1	1
鱼类,水果	100	0.1519756838905775	水果	鱼类	0.49504950495049505	1.5736356244320084	1	1
糖果,白酒	94	0.14285714285714285	糖果	白酒	0.48704663212435234	1.6605009530457193	1	1
糖果,白酒	94	0.14285714285714285	白酒	糖果	0.48704663212435234	1.6605009530457193	1	1
冻肉,建酒	120	0.182370820668693	冻肉	建酒	0.5529953917050692	1.7578307620383358	1	1
冻肉,建酒	120	0.182370820668693	建酒	冻肉	0.5797101449275363	1.7578307620383358	1	1
蔬菜罐头,建酒	122	0.18541033434650456	蔬菜罐头	建酒	0.5570776255707762	1.7708071382877815	1	1
蔬菜罐头,建酒	122	0.18541033434650456	建酒	蔬菜罐头	0.5893719806763285	1.7708071382877815	1	1
蔬菜罐头,冻肉	130	0.19756838905775076	冻肉	蔬菜罐头	0.5990783410138248	1.7999705405803503	1	1
蔬菜罐头,冻肉	130	0.19756838905775076	蔬菜罐头	冻肉	0.593607305936073	1.7999705405803506	1	1
蔬菜罐头,冻肉,建酒	107	0.16261398176291794	建酒	蔬菜罐头,冻肉	0.5169082125603864	2.6163507989594943	1	2
蔬菜罐头,冻肉,建酒	107	0.16261398176291794	蔬菜罐头,冻肉	建酒	0.823076923076923	2.6163507989594943	2	1
蔬菜罐头,冻肉,建酒	107	0.16261398176291794	冻肉	蔬菜罐头,建酒	0.4930875576036866	2.659439450026441	1	2
蔬菜罐头,冻肉,建酒	107	0.16261398176291794	蔬菜罐头,建酒	冻肉	0.8770491803278688	2.6594394500264413	2	1
蔬菜罐头,冻肉,建酒	107	0.16261398176291794	蔬菜罐头	冻肉,建酒	0.4885844748858447	2.6790715372907155	1	2
蔬菜罐头,冻肉,建酒	107	0.16261398176291794	冻肉,建酒	蔬菜罐头	0.8916666666666667	2.6790715372907155	2	1

3.2 操作步骤

(1) 在**实验管理**中，选择产品内置的实验案例 “ **购物篮分析** ”，参考如下GIF演示的步骤，拖拽 **关联规则生成** 后，配置 **最小提升度**，便可轻松输出关联规则的相关指标。



(2) 结合“**过滤**”功能，在过滤器中定义针对前项、后项的数量，以便筛选出一对一、一对多、多对一、多对多等关联形式。如下图所示，通过SQL语句定义筛选条件：前项数量大于1、后项数量为1。

🔍 当前显示 3 条 / 总共有 3 条数据

Items	# freq	# support	Antecedent	Consequent	# confidence	# lift	Antecedent quantity	Consequent quantity
蔬菜罐头,冻肉,啤酒	107	0.16261398176291794	蔬菜罐头,冻肉	啤酒	0.823076923076923	2.6163507989594943	2	1
蔬菜罐头,冻肉,啤酒	107	0.16261398176291794	蔬菜罐头,啤酒	冻肉	0.8770491803278688	2.6594394500264413	2	1
蔬菜罐头,冻肉,啤酒	107	0.16261398176291794	冻肉,啤酒	蔬菜罐头	0.8916666666666667	2.6790715372907155	2	1