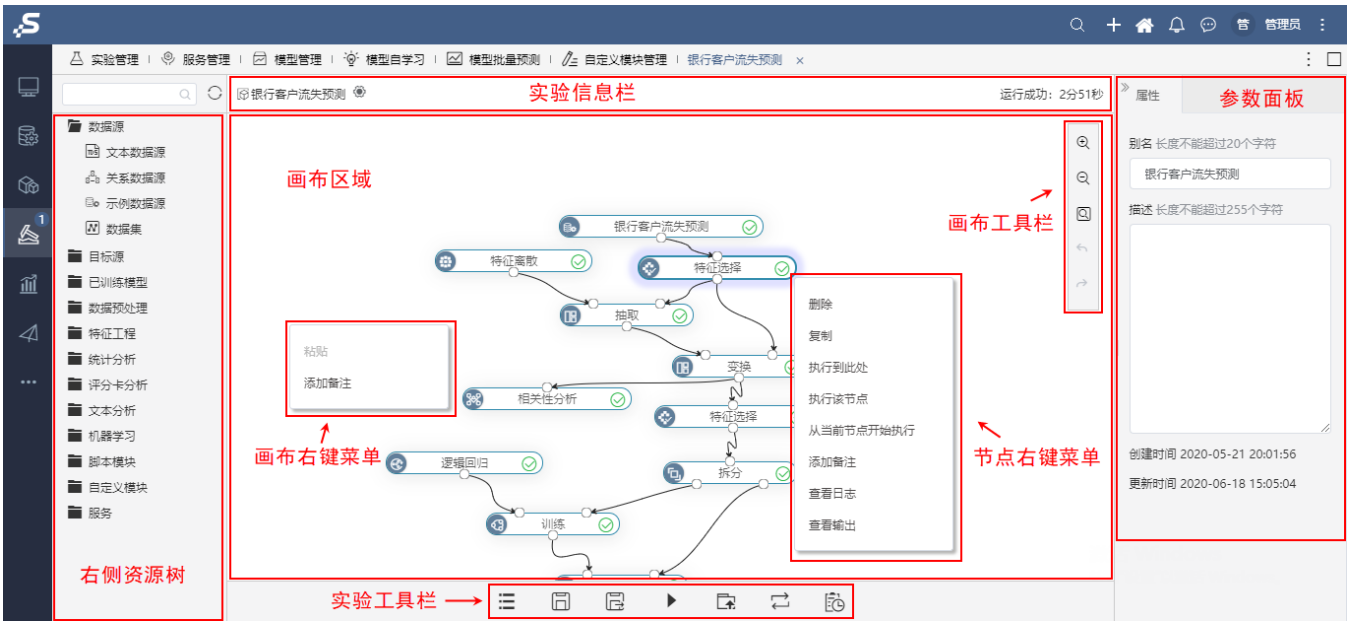
# 实验界面介绍

> “引擎按照节点粒度进行调度”为V9.7版本功能

数据挖掘主要页面为实验界面，其中也存在很多工具按钮，接下来为大家做详细介绍。

**实验界面**

实验界面除了存放节点的‘右侧资源树’，节点右键菜单还有‘画布区域’、‘画布右键菜单’、‘节点右键菜单’、‘画布工具栏’、‘实验工具栏’、‘参数面板’等，如图：




接下来为大家逐一简单介绍：

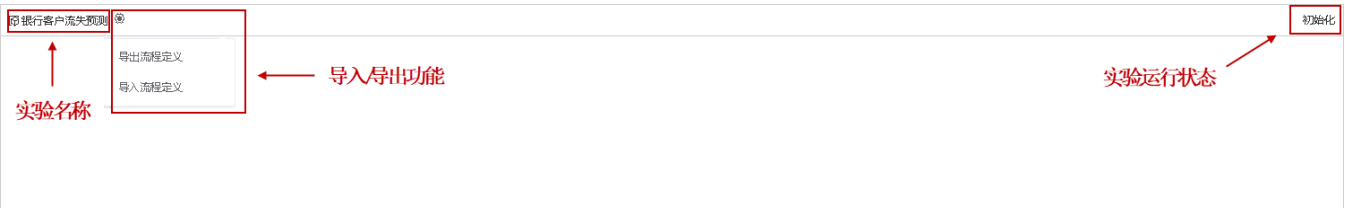
**右侧资源树**

| 名称 | 说明 | 详细介绍 |
| --- | --- | --- |
| 数据源 | 数据源中的节点是数据挖掘实验数据输入节点，目前有文本数据源，关系数据源，数据集，以及产品内置示例数据源。 | 数据源节点 |
| 目标源 | 目标源中的节点是数据挖掘实验结果输出节点，目前支持将结果输出至关系目标源，或导出数据到HDFS。 | 目标源节点 |
| 已训练模型 | 已训练的模型是用户将训练并运行通过后的模型封装保存为节点对象，支持复用，可以简化用户定制工作流的操作。 | |
| 数据预处理 | 数据预处理中是提供对’脏’数据清洗，初步加工等一系列处理功能的节点，具体内容参考详细介绍。 | 数据预处理节点 |
| 特征工程 | 特征工程中包括常见特征工程方法节点。通过特征工程，能够最大限度地从原始数据中提取特征，以供算法和模型使用。 | 特征工程节点 |
| 统计分析 | 统计分析中包括基础数理统计方法，用以数据的分析统计。 | 统计分析节点 |
| 评分卡分析 | 评分卡是一种通用的建模框架，将原始数据通过分箱后进行特征工程变换，继而应用于线性模型进行建模的一种方法。 | 评分卡分析 |
| 文本分析 | 文本分析中包括进行文本分析前，对文本数据加工处理的一系列节点。 | 文本分析节点 |
| 机器学习 | 产品提供常见的机器学习算法节点按照算法类别，分别放在分类算法，回归算法，聚类算法，关联规则文件夹中。而训练，预测，评估均为模型搭建中必须的节点，需搭配算法节点使用。 | 机器学习算法节点 |
| 脚本模块 | 目前产品提供的算法节点不能够涵盖所有算法，故提供PYTHON脚本、SQL脚本扩展产品的算法能力。 | 脚本节点 |
| 自定义模块 | 自定义模块存放用户保存自行编写的PYTHON脚本、SQL脚本节点。 | 自定义模块 |

| | | |
|---|---|---|
| 服务 | 组合使用服务中的两个节点，可以将数据挖掘实验发布为web服务，提供接口，应用程序可以实时调用数据挖掘实验，并能同步获取执行结果。 | 服务节点 |

**实验信息栏**

实验信息栏位显示的内容依次为实验名称，导入导出功能按钮，和实验运行情况，如下图：




其中，导入/导出功能介绍如下：

| 名称 | 说明 |
|---|---|
| 导出流程定义 | 将画布上的实验流程以文件的形式导出。 |
| 导入流程定义 | 新建空白实验，可以将保存的实验流程文件导入该空白实验，能够实现数据挖掘实验的共享，以及文件形式的备份。 |

**画布区域**

**画布工具栏**

| 样式 | 名称 | 说明 |
|---|---|---|
| | 放大 | 用于放大画布区域。 |
| | 缩小 | 用于缩小画布区域。 |
| | 原始大小 | 复原画布区域至默认大小。 |
| | 撤销 | 在画布区域进行实验搭建过程中，用于撤销拖拽节点、节点连线等操作。 |
| | 还原 | 还原相对于撤销而言，对撤销的部分进行还原。 |

注：撤销和还原使用的约束。

**画布右键菜单**

| 名称 | 说明 |
|---|---|
| 粘贴 | 与节点右键菜单的“复制”结合使用，粘贴复制的节点。 |

| | |
|---|---|
| 添加备注 | 对实验或节点添加备注信息进行记录。备注示例图如下：<br> <br><br>备注框的操作如下：<br>• 移动：鼠标点击备注框可以进行移动。<br>• 大小更改：鼠标移到备注框边缘的圆圈处，点击拖动即可改变大小。<br>• 边框设置：选择 **边框设置**，弹出“组件设置”窗口，用于设置边框的样式、宽度和颜色。<br> <br><br>• 编辑：编辑备注的内容，备注编辑采用的是文本组件，具体设置可参考 文本组件。<br>• 删除：删除该备注。 |

**节点右键菜单**

各节点资源的右键菜单支持对工作流的相关操作。

各节点资源的右键菜单如下：

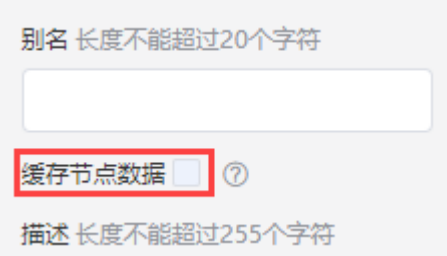
| 节点资源分类 | 右键菜单 |
|---|---|

| | |
|---|---|
| “评估”节点资源 | 删除<br>复制<br>执行到此处<br>执行该节点<br>从当前节点开始执行<br>添加备注<br>查看分析结果<br>查看日志 |
| “训练”节点资源 | 删除<br>复制<br>执行到此处<br>执行该节点<br>从当前节点开始执行<br>添加备注<br>查看分析结果<br>查看日志<br>模型 > |

| | |
|---|---|
| “PYTHON脚本”节点资源 | 删除<br>复制<br>执行到此处<br>执行该节点<br>从当前节点开始执行<br>添加备注<br>查看日志<br>查看输出<br>保存脚本 |
| 其它节点资源 | 删除<br>复制<br>执行到此处<br>执行该节点<br>从当前节点开始执行<br>添加备注<br>查看日志<br>查看输出 |
| 选择多个节点 | 删除<br>复制<br>分组 |

这些右键菜单各项的说明如下：

| 右键菜单 | 说明 |
|---|---|
| 删除 | 表示删除当前节点资源。 |
| 复制 | 复制选中的节点，可以是一个或者多个,与画布右键菜单的“粘贴”节点相结合使用。 |
| 执行到此处 | 表示运行工作流时到当前节点资源结束。 |

| | |
|---|---|
| 执行该节点 | 表示单独执行该节点。<br>需要在挖掘配置中配置HDFS，且挖掘、ETL右侧面板勾选缓存节点数据时，节点的右键才能显示**执行该节点**和**从当前节点开始执行**。<br>  <br> |
| 从当前节点开始执行 | 表示运行工作流时从当前节点资源开始执行。 |
| 添加备注 | 同画布右键菜单的 添加备注 。 |
| 查看日志 | 用于查看当前节点资源的运行日志。 |
| 查看输出 | 用于查看当前节点资源的输出列表。<br>• 下载预览数据：在“查看输出”窗口可以下载预览数据。<br> <br>此处会把预览的数据以csv文件的方式下载到本地，不会下载全量数据，数据量最多100条。 |
| 查看分析结果 | 用于查看当前工作流的分析结果。 |
| 模型 | 用于将当前模型保存为“已训练模型”，便于搭建其它工作流时可以直接引用。 |
| 保存脚本 | 用于将当前PYTHON脚本保存到自定义模块下，便于复用。详情请参考 自定义模块。 |
| 分组 | 将选择的多个节点合并为一组，可收缩和展开合并的节点，合并后便于归类和移动节点。<br>注意：分组节点不允许复制和再次分组。 |

当前显示 100 条 / 总共有 20640 条数据

| # 房价中位数值 | # 收入中位数 | # 街区房龄 | # 街区平均房间数 | # 街区平均卧室数 | # 街区人口数 | # 平均占地面积 | # 纬度 | # 经度 |
|---|---|---|---|---|---|---|---|---|
| 4.526 | 8.3252 | 41 | 6.984126984 | 1.023809524 | 583 | 2.555555556 | 37.88 | -122.23 |
| 3.585 | 8.3014 | 21 | 6.238137083 | 0.971880492 | 494 | 2.109841828 | 37.86 | -122.22 |
| 3.521 | 7.2574 | 52 | 8.288135593 | 1.073446328 | 506 | 2.802259887 | 37.85 | -122.24 |
| 3.413 | 5.6431 | 52 | 5.817351598 | 1.073059361 | 674 | 2.547945205 | 37.85 | -122.25 |
| 3.422 | 3.8462 | 52 | 6.281853282 | 1.081081081 | 438 | 2.181467181 | 37.85 | -122.25 |
| 2.697 | 4.0368 | 52 | 4.761658031 | 1.103626943 | 567 | 2.139896373 | 37.85 | -122.25 |
| 2.992 | 3.6591 | 52 | 4.931906615 | 0.951361868 | 619 | 2.128404669 | 37.84 | -122.25 |
| 2.414 | 3.12 | 52 | 4.797527048 | 1.061823802 | 578 | 1.788253478 | 37.84 | -122.25 |
| 2.267 | 2.0804 | 42 | 4.294117647 | 1.117647059 | 742 | 2.026890756 | 37.84 | -122.26 |
| 2.611 | 3.6912 | 52 | 4.970588235 | 0.990196078 | 1080 | 2.172268908 | 37.84 | -122.25 |
| 2.815 | 3.2031 | 52 | 5.47761194 | 1.07960199 | 745 | 2.263681592 | 37.85 | -122.26 |

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

**实验工具栏**

| 样式 | 名称 | 说明 |
|---|---|---|
| | 查看历史 | 实验每次执行的评估结果作为历史信息都被记录下来，在历史信息页面展示，方便用户对比多次实验结果，进而选取出最优实验。 |
| | 保存 | 该按钮用于保存新建实验、实验流程变动等。 |

| | | |
|---|---|---|
| | 另存为 | 可将实验另存。 |
| | 运行 | 执行实验流程。 |
| | 部署服务 | 将数据挖掘实验发布为web服务。 |
| | 设置模型自学习 | 发布到生产的服务内所训练的模型，通过模型自学习有可能提高模型准确性。 |
| | 模型批量预测 | 针对批量的数据，采用服务发布的方式将预测结果数据投放到生产。 |
| | 参数设置 | 通过参数筛选用户可查询出需要的数据。目前只有关系数据源支持参数设置功能。 |

**参数面板**

参数面板分为两种状态：

- 未选中画布中任何节点：参数面板区域仅有属性页签。
- 选中画布中节点：参数面板区域依次出现参数、属性、帮助三个页签。

由于两种状态的属性页签的作用一致，统一介绍如下：

| 页签名 | 说明 |
|---|---|
| 参数 | 根据具体节点，显示参数信息。 |
| 属性 | 可以修改实验或者节点的别名，其中，实验还可以增加描述。 |
| 帮助 | 具体节点的说明，及作用介绍。 |