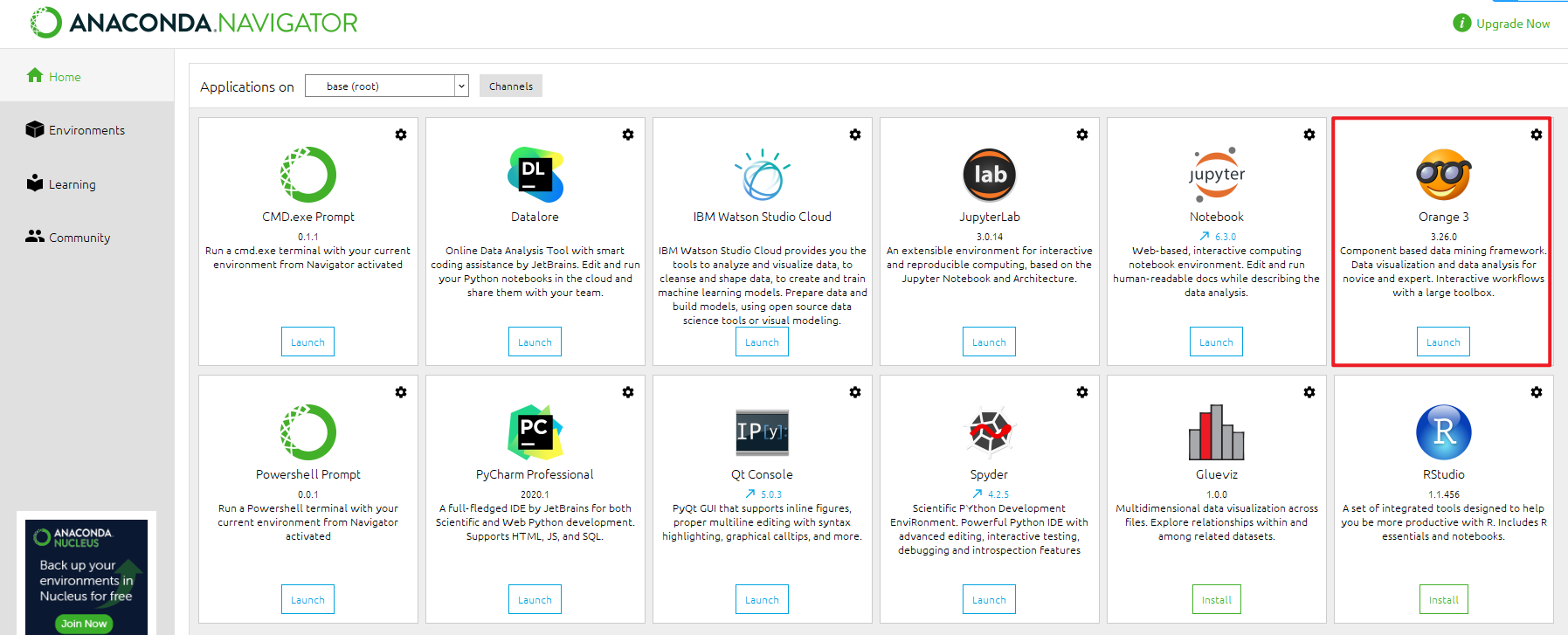
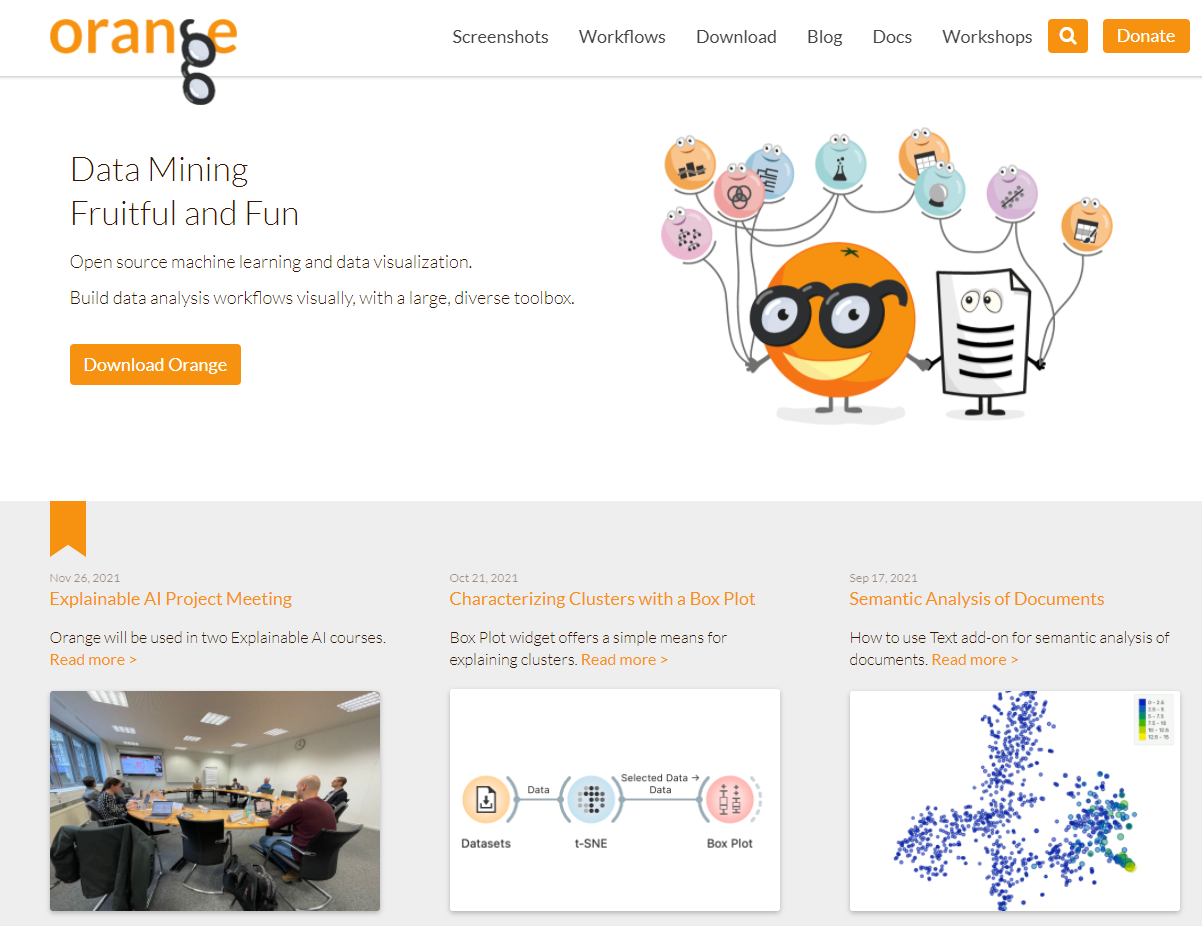
1. **orange3的基本用法**

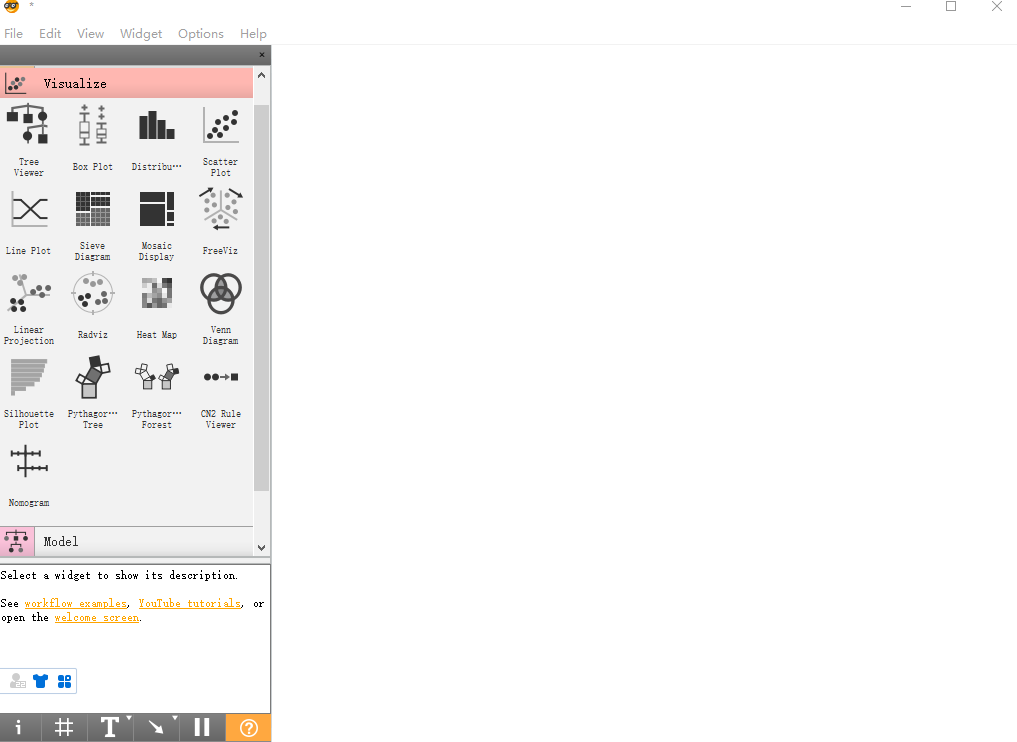
Anaconda 中集成了Orange3



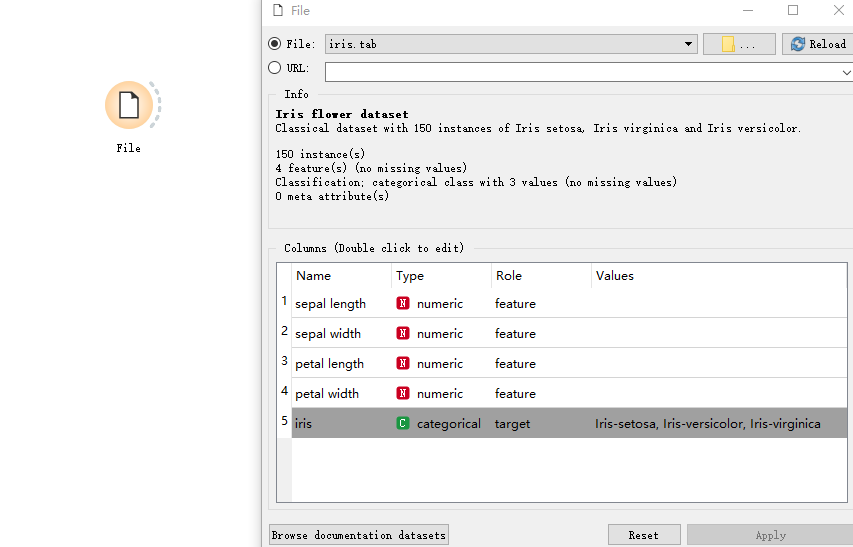
也可以通过官网下载安装



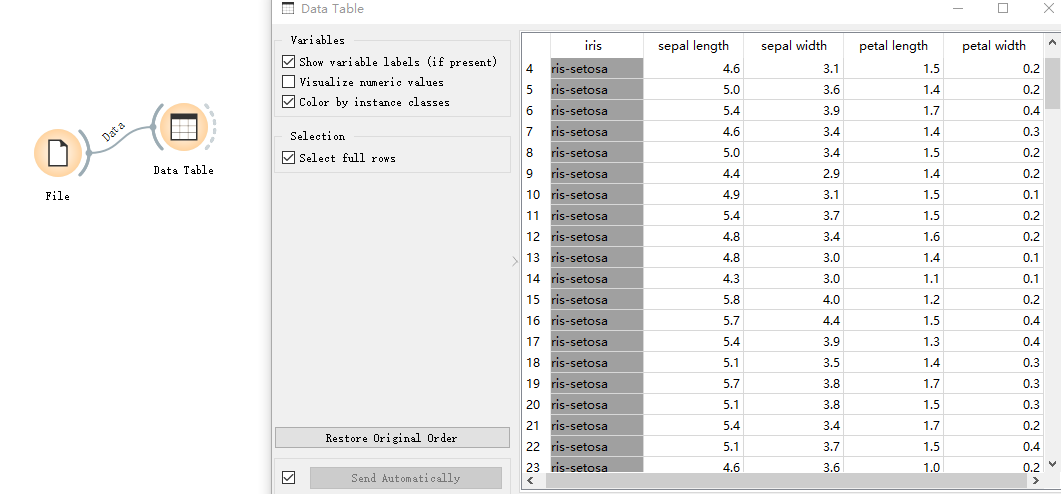
初始页面



导入数据



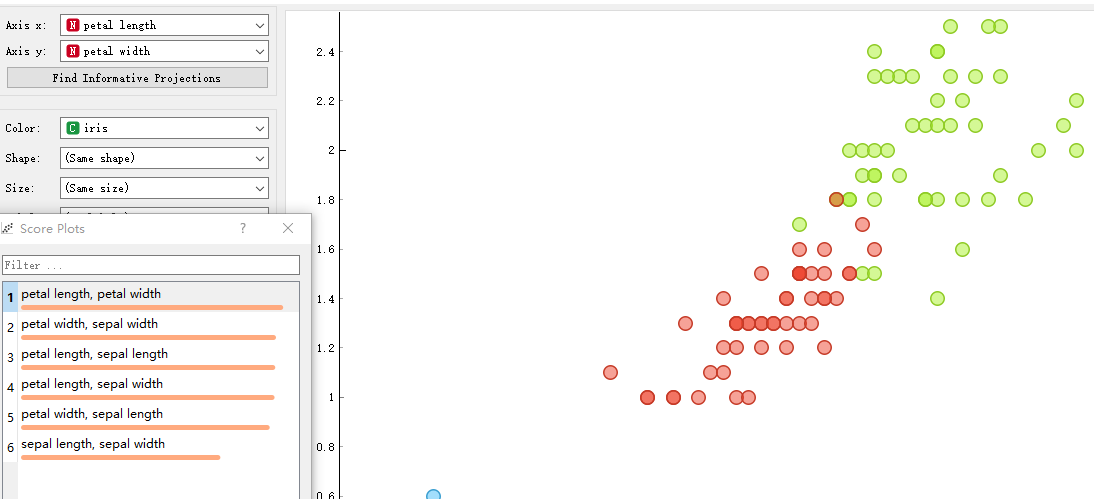
数据显示



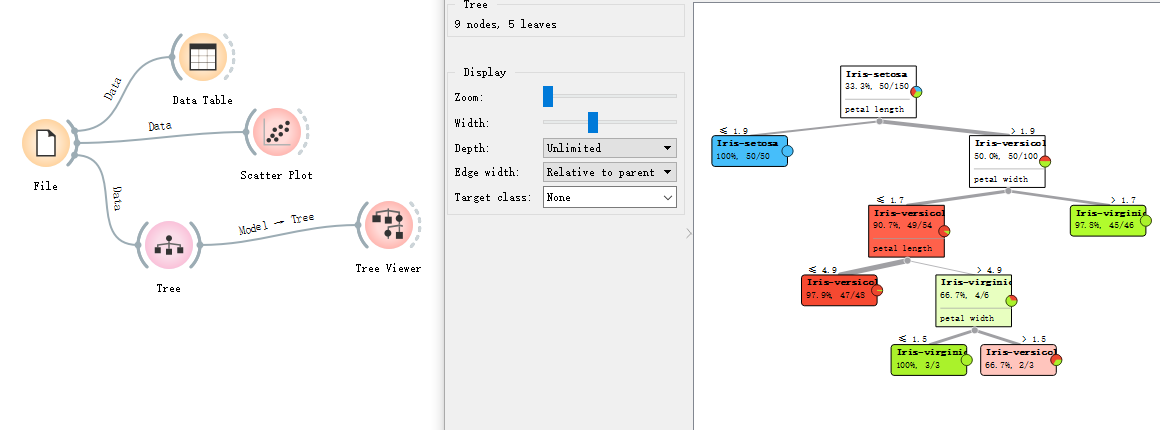
绘制散点图



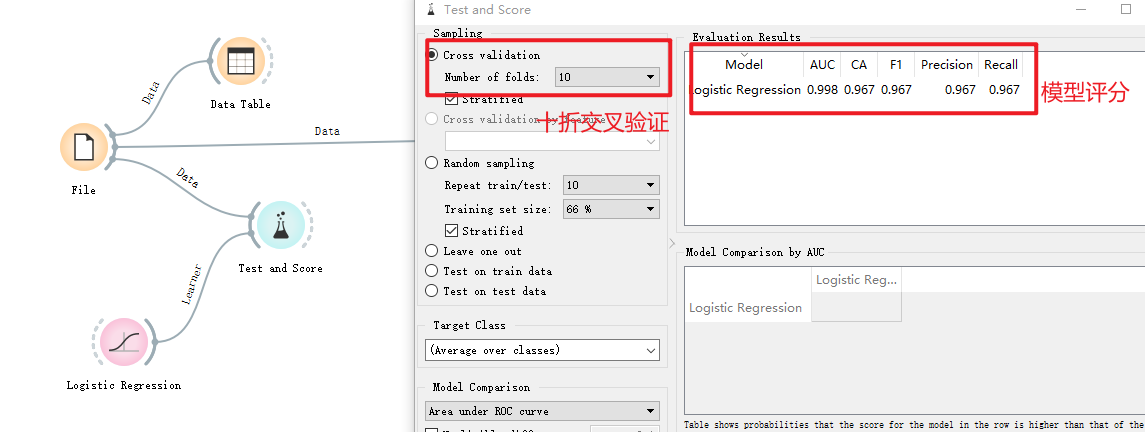
特征评分，并选择最高评分绘制散点图



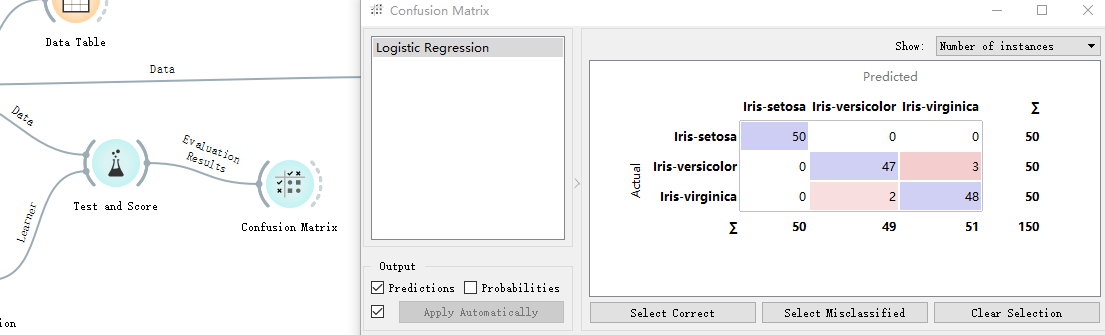
使用决策树算法



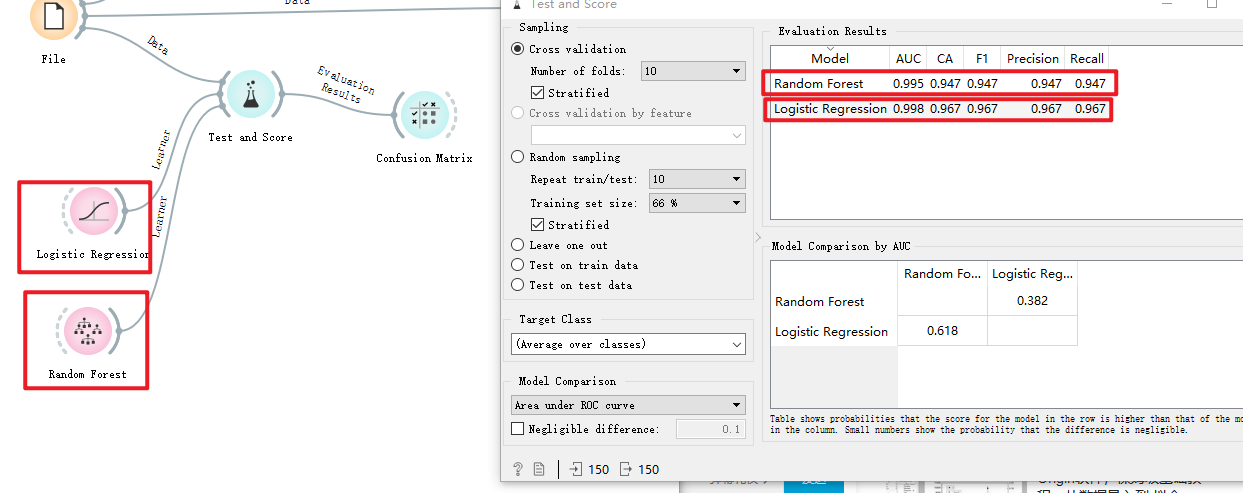
利用逻辑回归算法进行分类并评分（十折交叉验证）



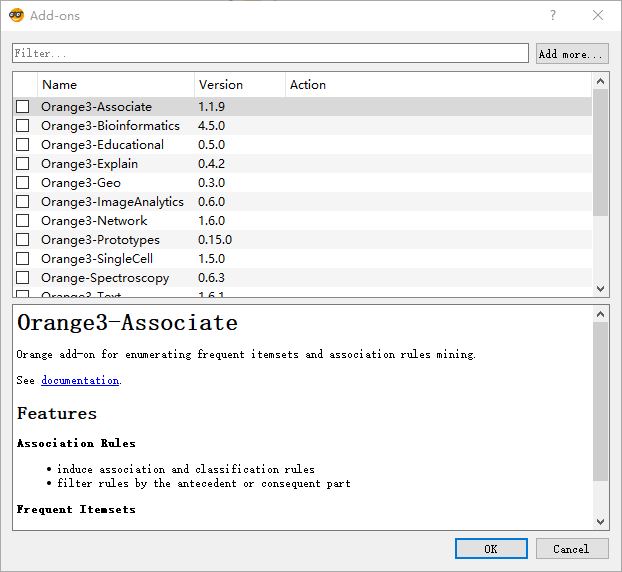
查看模型混淆矩阵

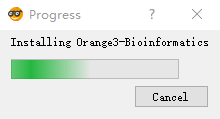


多模型对比



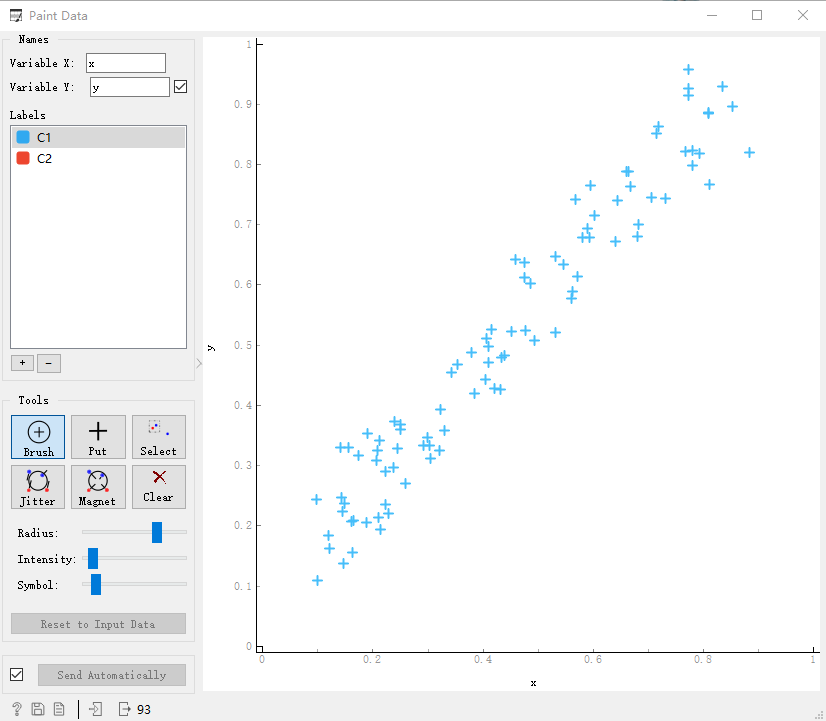
安装其他工具包 Options->Add-one



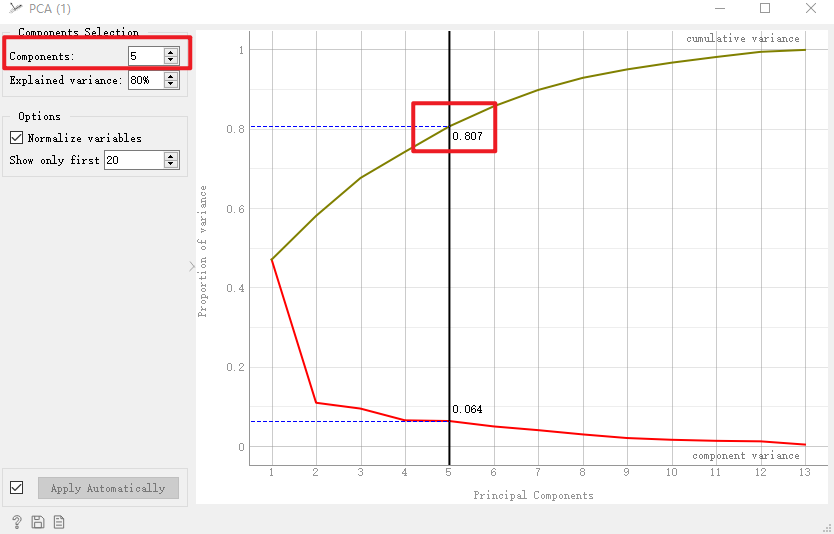


主成分分析 PCA

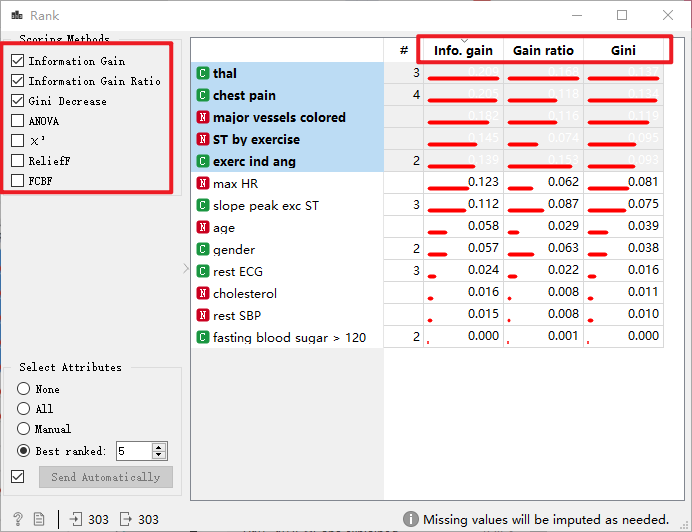
Paint data



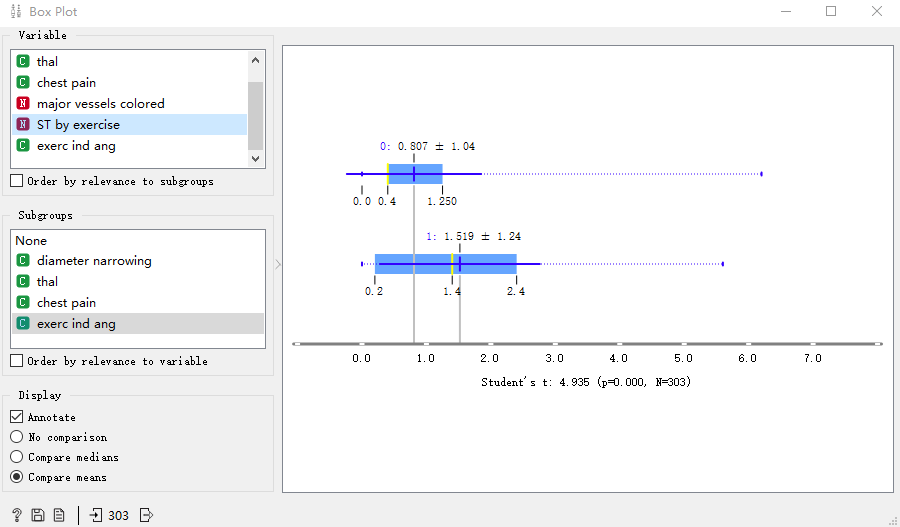
5个主成分可以解释80.7%的数据



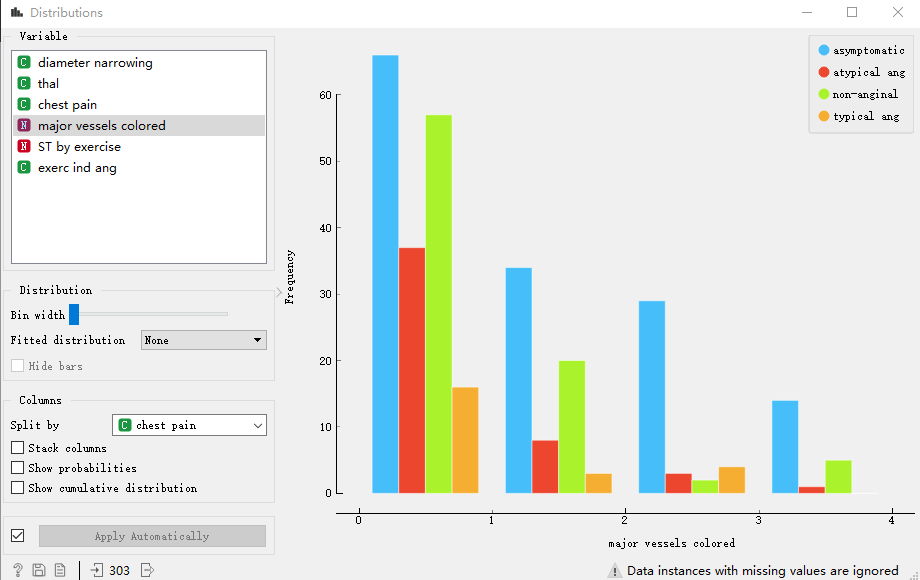
Rank 成分评分



属性与类的关联性



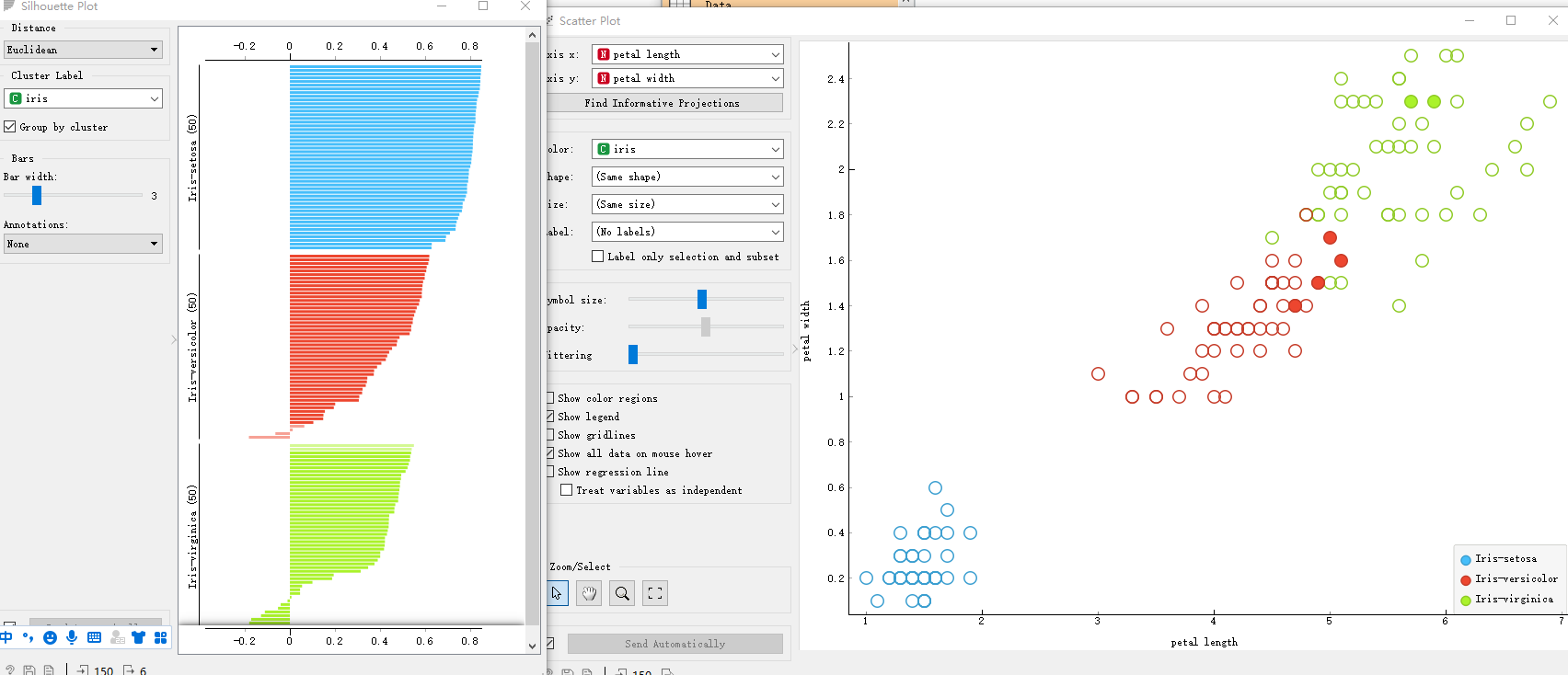
属性密度分布图



K-mean

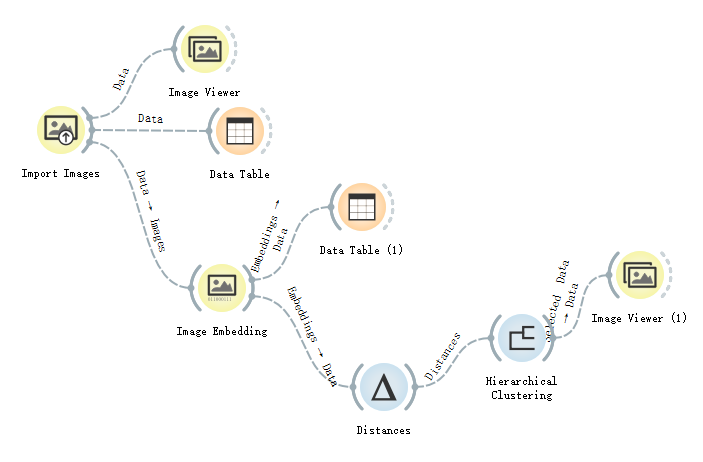
Educational插件

Silhouette Plot轮廓图



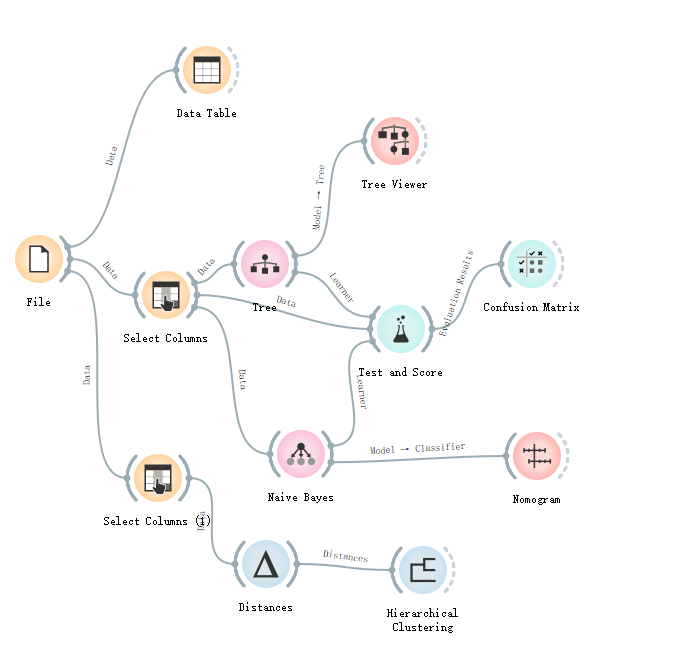
图像转数字

Add-one ImageAnalytics

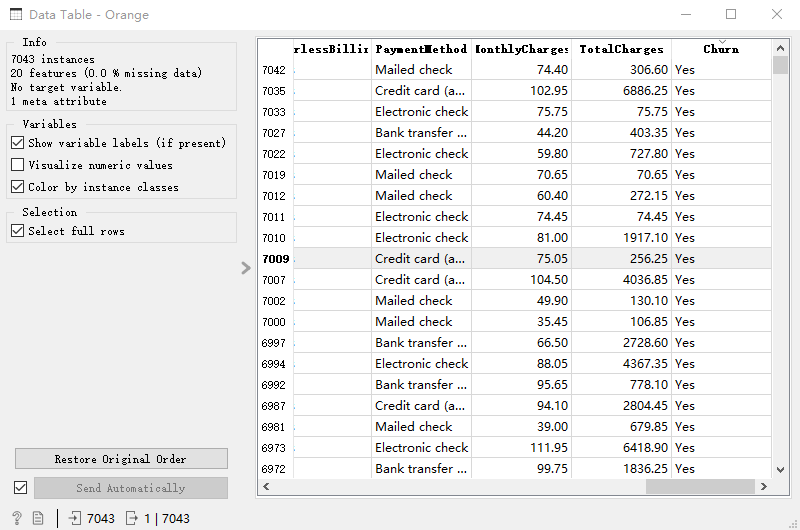


1. **kaggle数据集实战**

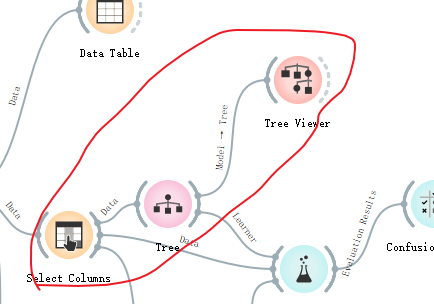
使用数据集：WA\_Fn-UseC\_-Telco-Customer-Churn.csv 电信客户流失数据

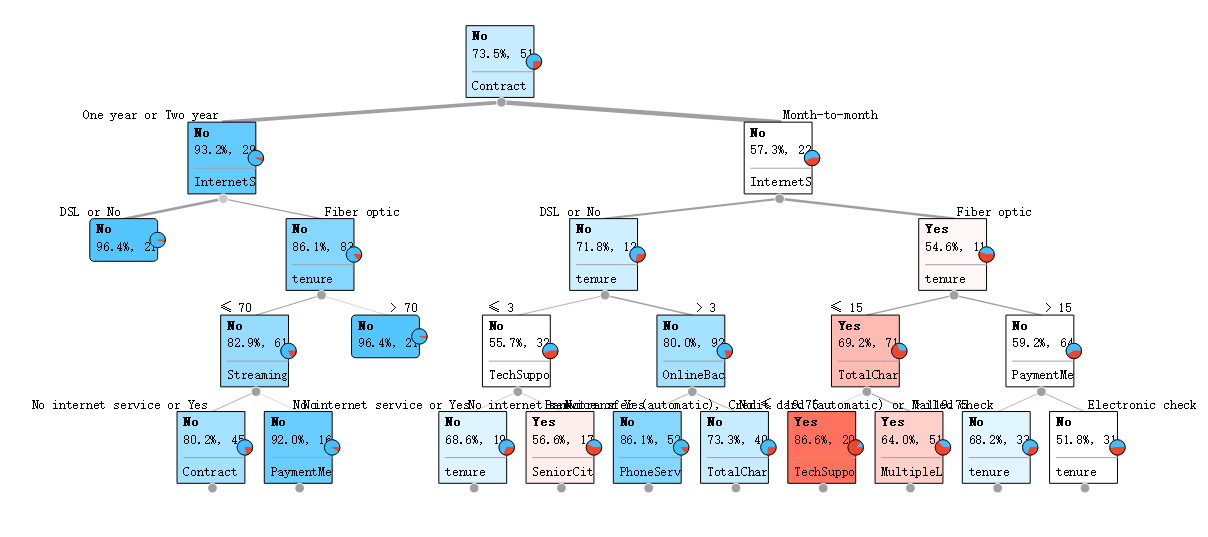


**查看数据：**

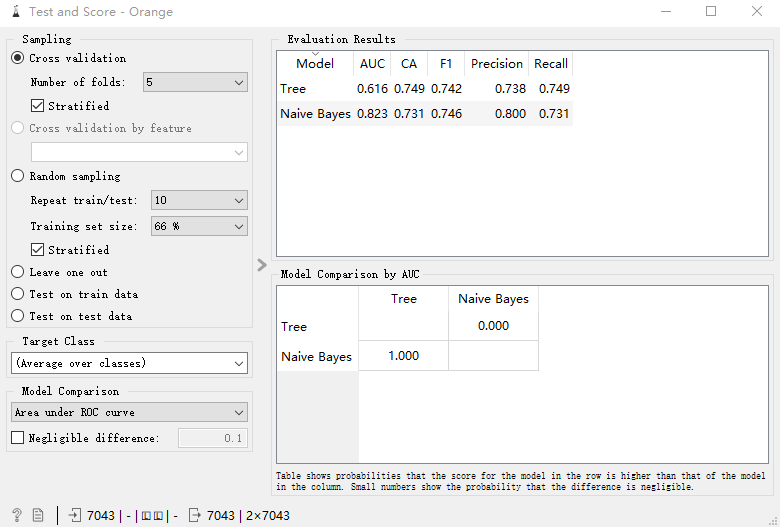


**决策树绘制：**

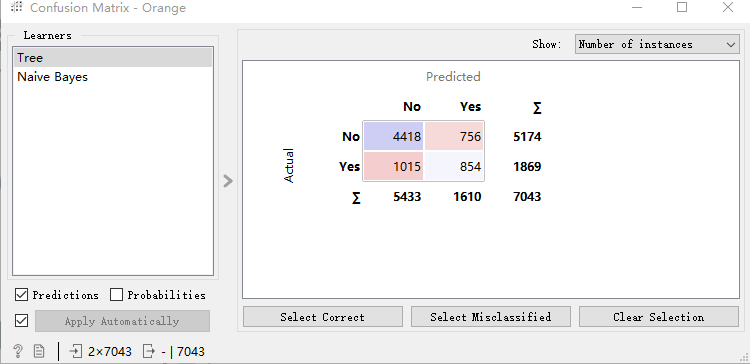




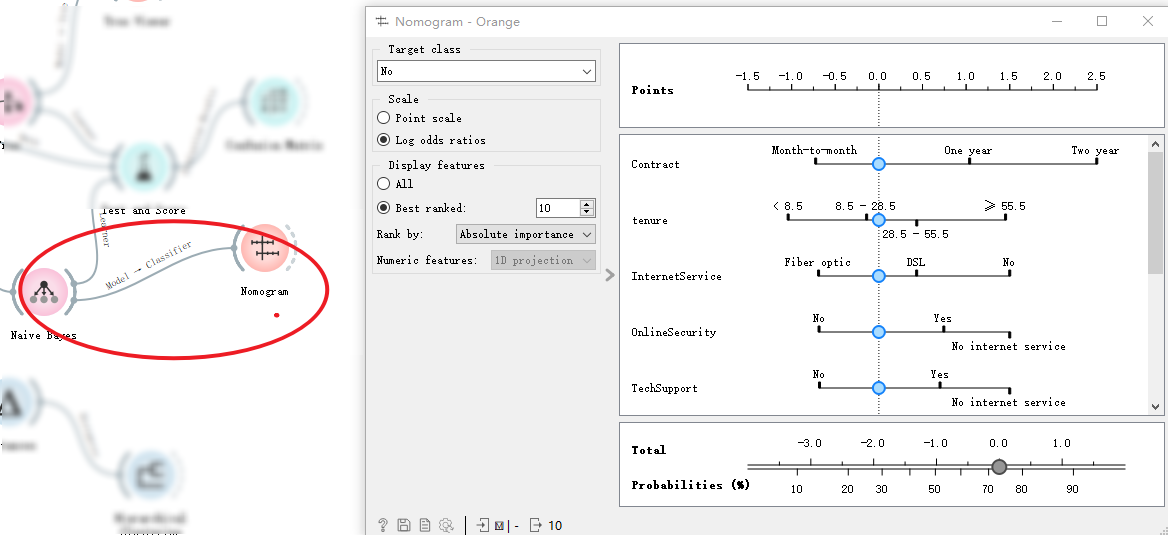
**模型评分：**



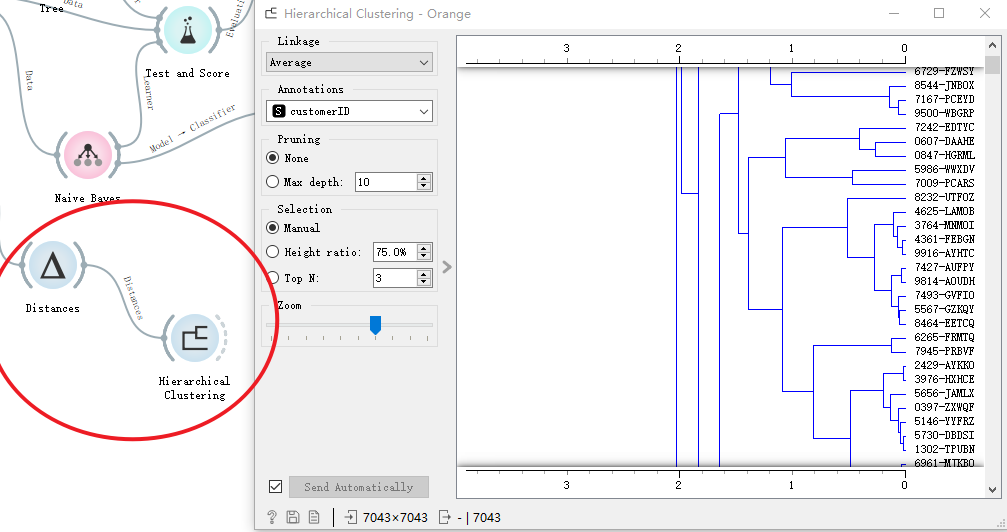
**混淆矩阵：**



**贝叶斯分类器**



**无标签情况下聚类，并用层次聚类法表示**



1. **心得体会**

**使用Orange3做数据挖掘可视化美观，流程清晰明了，操作简单。**

**能够应对基本的数据挖掘与数据分析的工作任务。**

**个人高估了Orange3的使用，以至于此次大作业难度相对简单。**

**在安装其他软件包即Options->Add-ons遇到了一些麻烦，最好是在管理员模式下安装，其中Text、Network工具包的安装需要安装Microsoft Visual C++ Build Tools（Windows环境下）。**