实验报告

成绩：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 班级学号 | 202015116 | 姓名 | 梁祺若 | 实验日期 | 2022-10-20 | 仪器编号 |  |
| 实验名称 | 实验6-8 R的支持向量机，一般聚类和特色聚类 | | | | | | |

**1．实验题目**

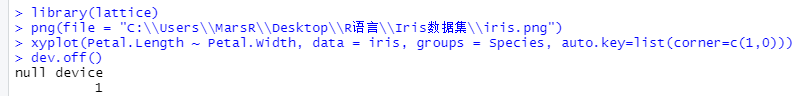
R的支持向量机，一般聚类和特色聚类

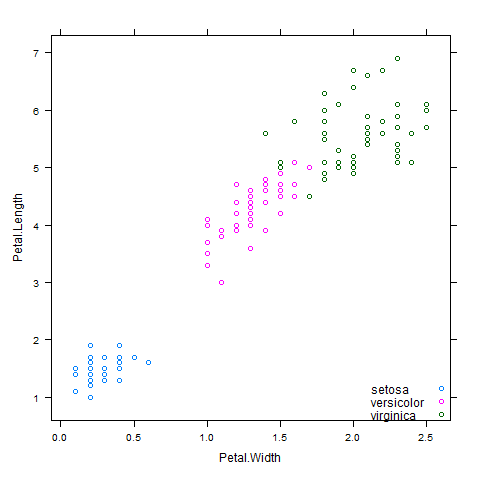
**2．实验步骤**

本题目由三部分组成：（1）利用R软件对iris数据做支持向量机；（2）利用R软件对iris数据做一般聚类；（3）利用R软件对\*\*数据做特色聚类。

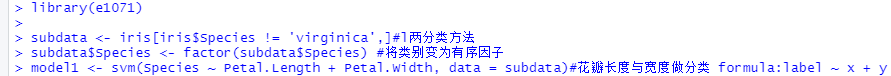
1. 利用R软件对iris数据做支持向量机

* 数据来源：https://cloud.tencent.com/developer/article/1869024
* 初步通过图像判断数据的分布情况

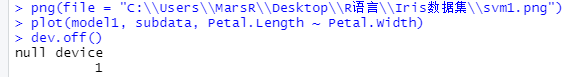


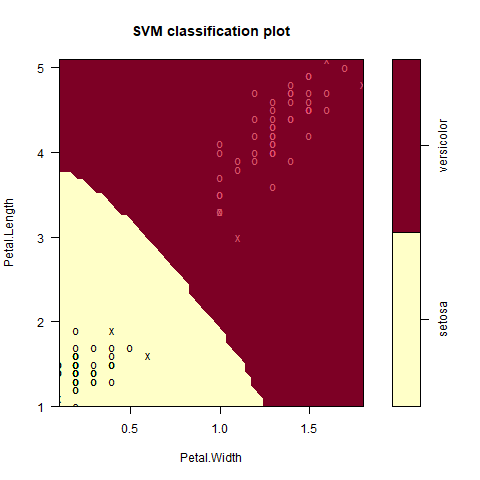


* 由图可以看出setosa很容易与versicolor和virginica分隔开，但后两者分隔不明显，这次试验先选择setosa和versicolor来进行实验

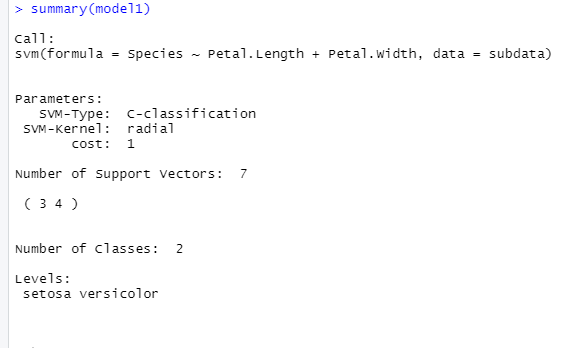


* SVM结果可视化



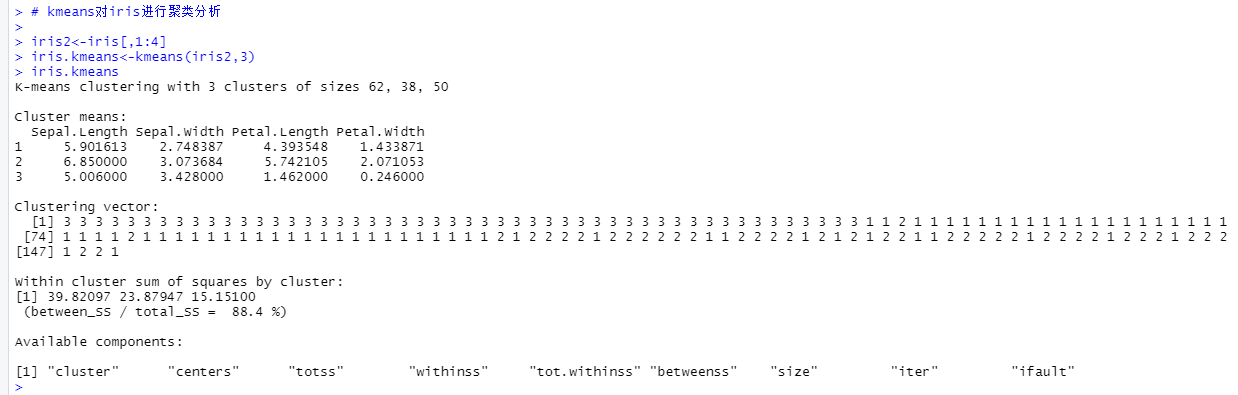


* 使用summary来查看SVM构建情况

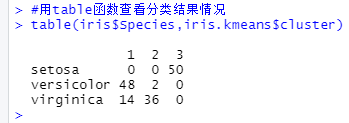


1. 利用R软件对iris数据做一般聚类。

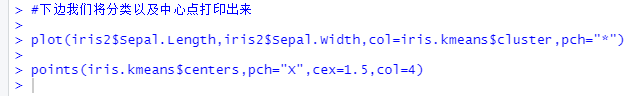
* 数据来源：<https://cloud.tencent.com/developer/article/1869024>
* Kmeans
* Kmeans对iris进行聚类分析

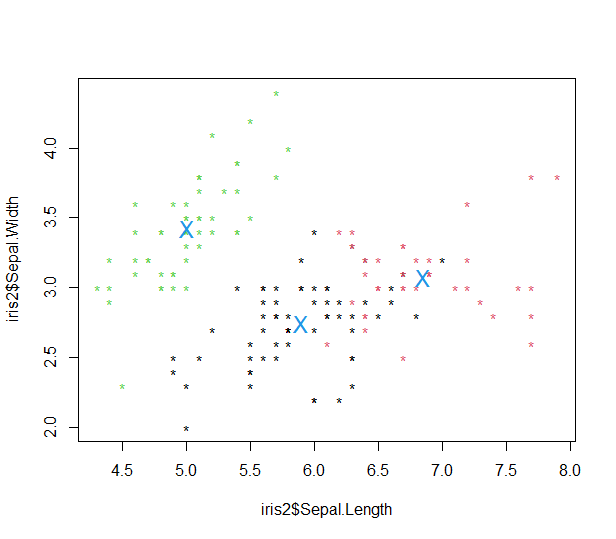


* 用table函数查看分类结果情况

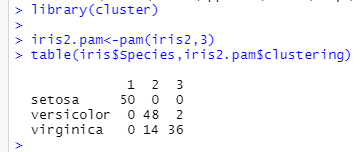


* 下边我们将分类以及中心点打印出来

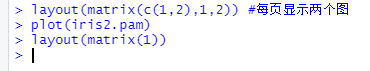


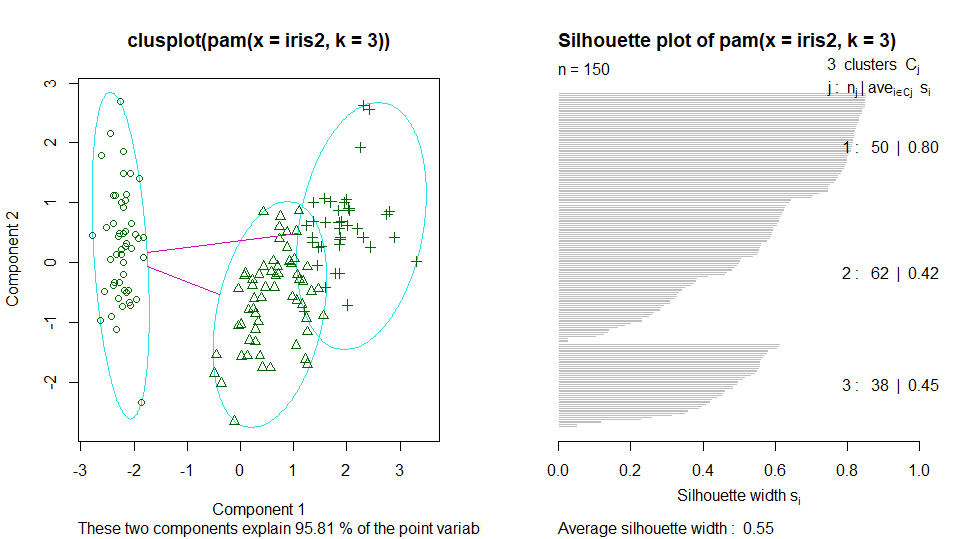


* K-mediods
* 使用K-mediods方法中的pam算法来进行聚类分析

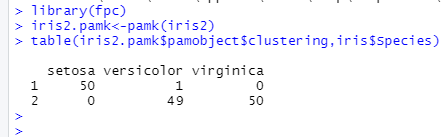


* Pam算法画图

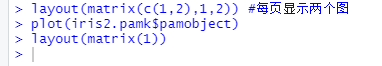


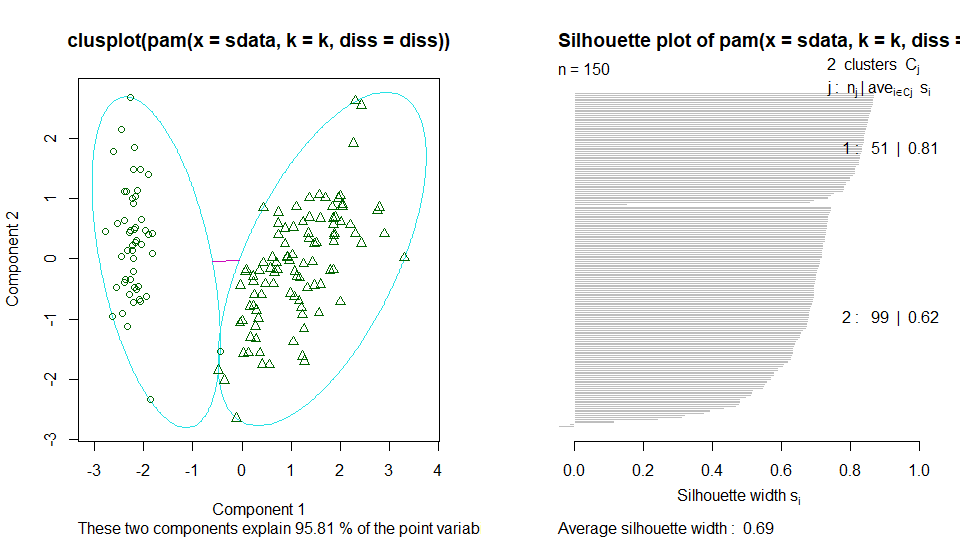


* 使用K-mediods方法中的pamk算法来进行聚类分析

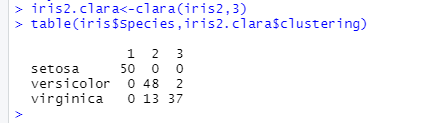


* pamk算法画图

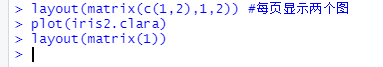


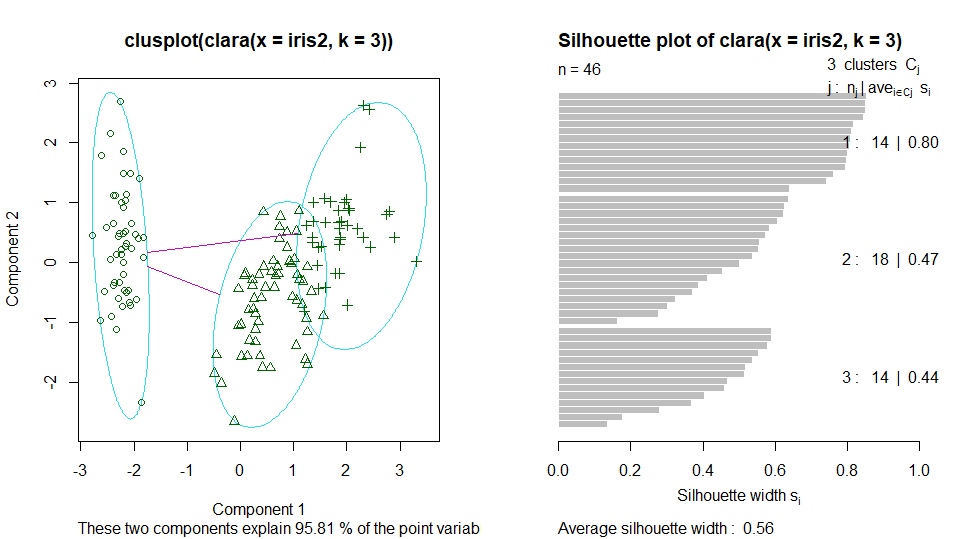


* 使用K-mediods方法中的clara算法来进行聚类分析

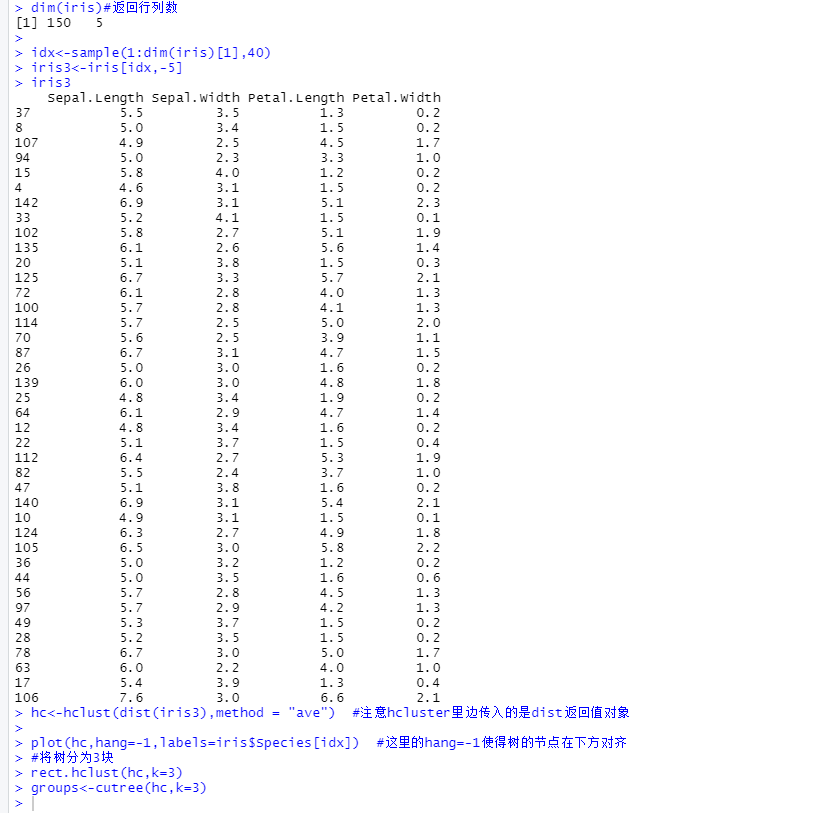


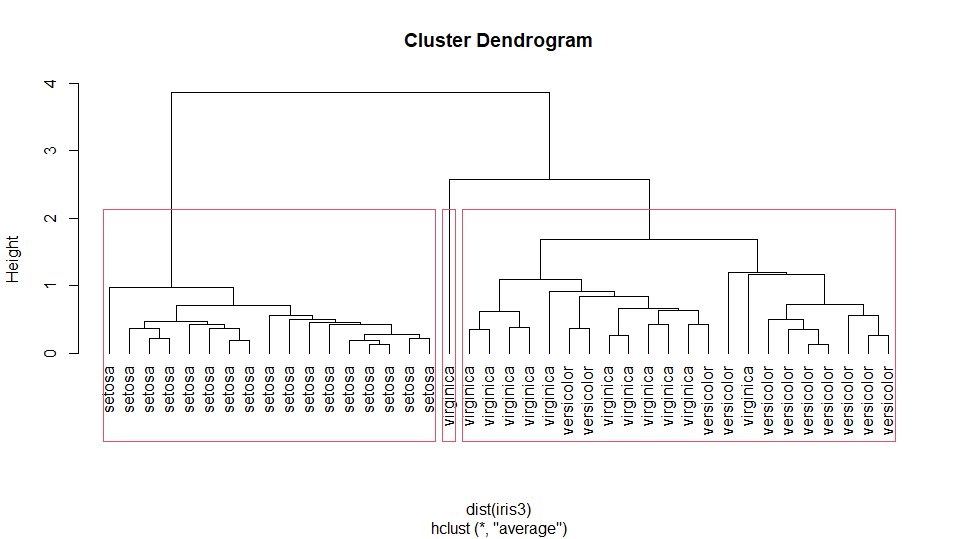
* Clara算法画图





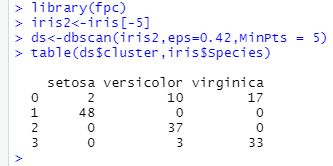
* 通过上述分类结果可以看到，pam和calra算法分类结果基本类似，但是pamk将三类分为了两类。
* 层次聚类HCluster





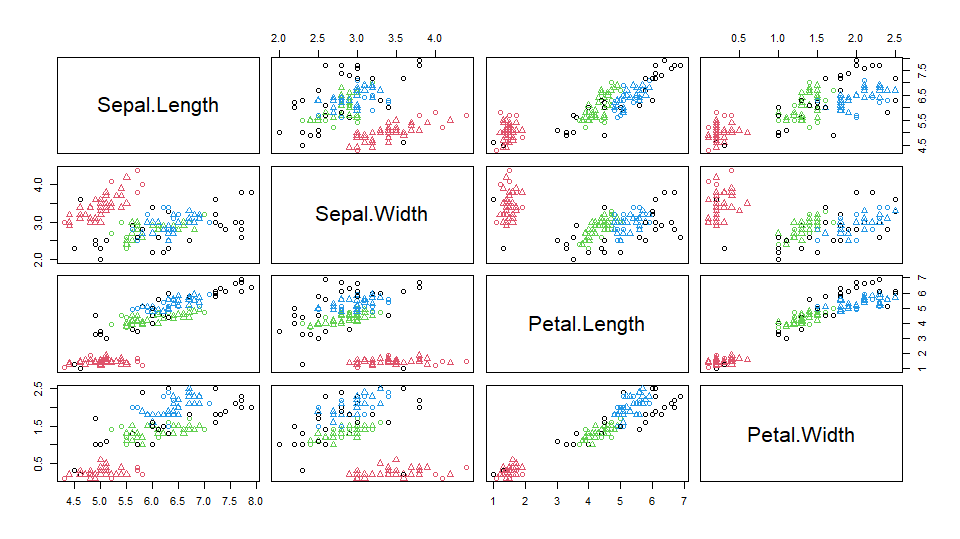
1. 利用R软件对iris数据做特色聚类。

* 基于密度的聚类DBSCAN

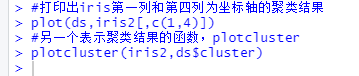


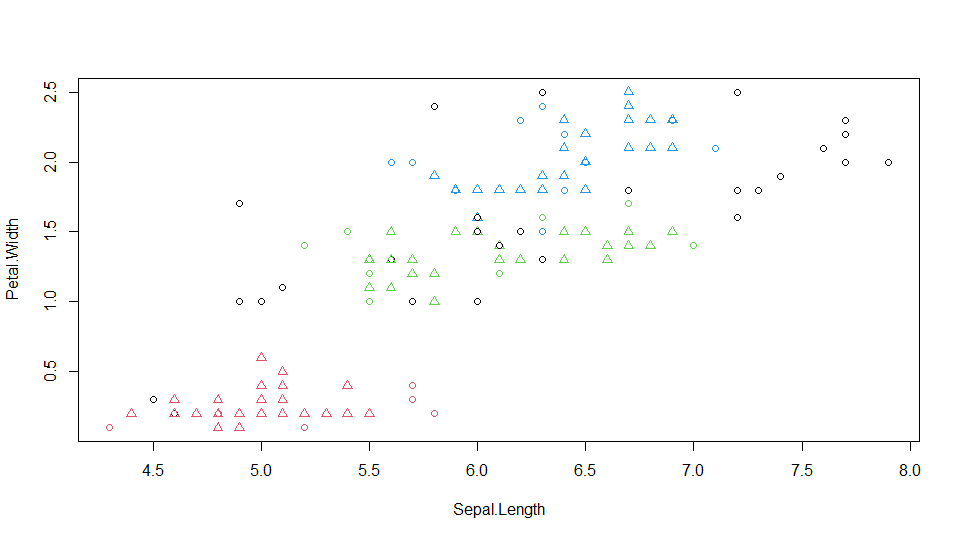
* 打印出ds和irsi2的聚类散点图



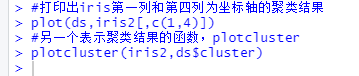


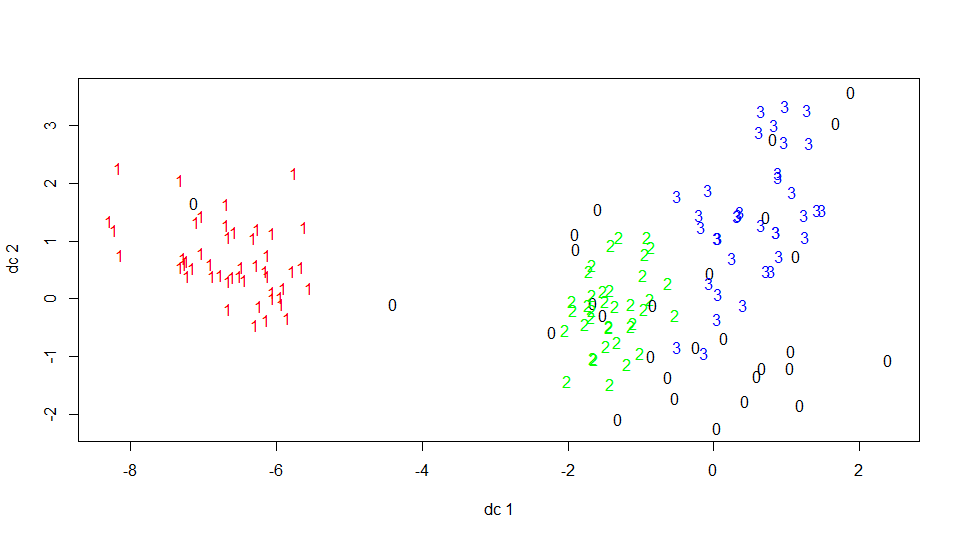
* 打印出iris第一列和第四列为坐标轴的聚类结果





* 另一个表示聚类结果的函数，plotcluster





**3．实验结果**

本次实验对iris数据集做了支持向量机以及一般聚类和特殊聚类分析，特别是对于一般聚类的K-mediods方法，使用了pam、pamk、clara三种不同的算法。通过上述分类结果可以看到，pam和calra算法分类结果基本类似，但是pamk将三类分为了两类。