实验四 聚类

**学号：**

**姓名：**

1. 实验目的

掌握三种聚类算法：K-means，AGNES，DBSCAN。

2. 实验要求

通过本实验应达到如下要求：

（1）掌握三种聚类算法的基本思想

（2）熟练使用Python或其他工具实现聚类算法

3. 实验器材

（1）计算机一台

（2）Python或其他编程工具

4. 实验内容

（1）数据集：可以使用sklearn库中的iris数据集，其中每个样本有4个特征参数，分别为花萼长度，花萼宽度，花瓣长度，花瓣宽度4个属性。聚类问题选取特征空间的4个维度，不选取数据集中的标签，可视化时可选取前2个维度绘图。（数据集不限）。

**选取数据集为：**

（2）分散性聚类（Kmeans）

（a）**阐述分散性聚类（Kmeans）的基本思想：**

（b）构造KMeans聚类器，**聚类的个数K=** ，**代码：**

（c）聚类结果可视化。使用iris数据集时，横纵坐标分别为两个特征Sepal Length和Sepal Width。使用python实现时，可以使用plt.scatter绘制聚类结果。**代码：**

**可视化图：**

（2）层次聚类（AGNES）

**（a）阐述层次聚类（AGNES）的基本思想：**

（b）构造聚类器，**代码：**

（c）聚类结果可视化。使用iris数据集时，横纵坐标分别为两个特征Sepal Length和Sepal Width。使用python实现时，可以使用plt.scatter绘制聚类结果。**代码：**

**可视化图：**

（3）密度聚类（DBSCAN）

**（a）阐述密度聚类（DBSCAN）的基本思想：**

（b）构造聚类器，**代码：**

（c）聚类结果可视化。使用iris数据集时，横纵坐标分别为两个特征Sepal Length和Sepal Width。使用python实现时，可以使用plt.scatter绘制聚类结果。**代码：**

**可视化图：**

（4）**从上面三个可视化图，对比K-means聚类、AGNES聚类和DBSCAN聚类结果的差异，分析三种聚类算法的优缺点：**

**5. 实验心得**