**实验10 Kmeans算法实现**

**一、实验名称**：

Kmeans算法实现

**二、实验目的：**

掌握KNN算法基本原理

**三、实验要求**

1、使用python实现Kmeans；

2、测试数据标准化对kmeans算法的影响

**四、实验平台**

计算机、Python、Anaconda

**五、实验数据**

processed.cleveland.data

UCI公开数据集-heart disease，属性信息如下:

1. #3 (age)

2. #4 (sex)

3. #9 (cp)

4. #10 (trestbps)

5. #12 (chol)

6. #16 (fbs)

7. #19 (restecg)

8. #32 (thalach)

9. #38 (exang)

10. #40 (oldpeak)

11. #41 (slope)

12. #44 (ca)

13. #51 (thal)

14. #58 (num) (the predicted attribute)

数据集参考网址：https://archive.ics.uci.edu/ml/datasets/Heart+Disease

**六、通过实验,要求同学在实验中解决以下问题，并递交完整的实验报告**

1. 使用Kmeans算法对上述数据进行聚类，然后使用标签数据对聚类效果进行评价

2. 阐述如何对kmeans算法进行并行化执行以提高算法效率？

3. 总结实验中遇到的问题及解决办法、心得、体会。