1. 实验任务题目

K-means算法设计与实现

1. 编程语言

C，C++，Java或者学过的其它编程语言

1. 实验数据介绍

**红酒数据集（Wine Data Set）**[**http://archive.ics.uci.edu/ml/datasets/Wine**](http://archive.ics.uci.edu/ml/datasets/Wine)

共178个数据，每个数据特征为13维

13个特征分别为：（13个化学成分，每个成分取值为实数）

1) Alcohol

2) Malic acid

3) Ash

4) Alcalinity of ash

5) Magnesium

6) Total phenols

7) Flavanoids

8) Nonflavanoid phenols

9) Proanthocyanins

10) Color intensity

11) Hue

12) OD280/OD315 of diluted wines

13) Proline

1. 实验完成要求

* 编程实现K-means算法，并在红酒数据集上运行。
* 设置不同K值，不同初始中心，在红酒数据集上进行实验比较。
* 分析k-means的优缺点，并对其中一个或几个缺点进行改进。
* 演示实验并提交代码，统计分析实验结果并上交实验报告；