模式识别第四次作业报告

1. **实验内容**
2. 用PCA对dataset3.txt样本的特征进行降维，画出各个主成分的方差，并据此确定若干主成分构成新的特征；
3. 在新特征表示的基础上，分别试验将训练样本聚成1、2、3、4、5、6个聚类，绘制误差平方和随聚类数变化的折线。并据此确定最合适的聚类数进行聚类，用合适的方式表示出聚类结果并对其进行分析和讨论。
4. **实验方法**
5. **实验设计**

特征降维：主成分分析(PCA)方法

聚类分析：Cmeans聚类

聚类效果评估：轮廓系数（Silhouette Coefficient）分析

1. **实验原理**
   * 1. 主成分分析(PCA)方法
     2. Cmeans聚类
     3. 轮廓系数（Silhouette Coefficient）分析
2. **实验过程**
3. **程序来源说明**
4. **结果分析**