

作业 2

内容

阅读 science 论文 Clustering by fast search and find of density peaks , 实现文中的聚类算法, 在论文提供的一个数据集上进行实验, 给出必要的实验结果。

鼓励大家对论文中的方法做一些改进, 提出质疑。或者对这个算法进行一定的分析 (比如 d_c 对结果的影响, 噪声点的判断是否合理等等), 大家也可以在其它数据集上进行实验, 给出一个比较深入的分析结果, 写在报告中。

要求

编程语言不限, 推荐 Python, R 或者 matlab。

有很多已经实现的代码, 比如 Python Package Index 有 Python 的实现, CRAN 中也有 R 语言的实现, 大家可以适当参照, 但不要直接拿过来用。

提交

源代码+报告

说明

Dcluster-0.5.2.tar.gz 为 Python 安装包

DataS1.zip 为论文中的 S1 数据集

Laio-talk.pdf 是作者的一个 slides