

# README

## 一、运行方式

hadoop jar Kmeans.jar Kmeans <初始中心路径> <训练集路径> <输出文件夹> <最大迭代轮数>

e.g. `hadoop jar Kmeans.jar Kmeans /input/initial_centers /input/dataset_mini.data /output 10`

## 二、原理

- Kmeans的思路同第九章PPT中一样
- 本次的主要难点是文件系统（代码注释应该写的比较清楚，不懂随时问我啦o( ￣▽￣)ブ）
- `Helper.java` 中是本次实验中常用的辅助函数