**毕业设计文件说明**

1. **数据集energydata文件夹**

存放需要用到energydata数据集和处理后数据集以及数据集描述，还有涉及到该数据集的论文。

**该文件夹里有五个子文件，分别是原版数据、删除第一天和倒数第一天的残缺数据、无时间无随机变量特征的数据、数据集详细描述文件和数据集相关论文文件。**

0原版.csv——无任何删减原版数据集

1-energydata-del1day.csv——删除第一天和倒数第一天的残缺数据

energydata-无时间、随机变量特征.csv——无时间、随机变量特征数据

数据集相关描述.docx——数据集描述

相关论文.pdf——数据集相关论文

1. **DBSCAN文件夹**

dbscan\_run,是基于密度算法

Normalized：是数据预处理，归一化算法

1. **Energy文件夹**

里面都是和RECS相关的建筑能耗的数据信息

RECS18子文件夹——包括建筑能耗13个特征和人相关的五个特征

BTU——totalBtu 09年能耗总量

00/11/22分别代表与RECS人的相关特征的3个类别。

Energy——建筑能耗的13个相关特征

00/11/22/分别代表与RECS人的相关特征的3个类别。

1. **GetCsvColum文件夹**

Demo——调用GetCsvColum算法，可以有实现以下几种功能：

1. 得到CSV文件中的某一列
2. 可以根据某一列得到另一列
3. 可以根据其他列得到某一列
4. 可以根据其他列的条件，得到某一列对应的数据

GetCsvColum——得到CSV文件中的某一列

Normalized：是归一化处理数据算法

1. **kmeans文件夹**

Elbow——肘部法则算法

Silhouette——轮廓系数算法

kmeans\_handwirrten——kmeans聚类算法

kmeansOutData——聚类分类类别算法

PlotFigure——画频率直方图算法

1. **PlotFigure文件夹**

里面分别是RECS和ATUS画频率直方图的数据。

前缀带有RE的代表只有人相关的五个特征。

反之，则是ATUS的九个特征和RECS的十个特征。

1. **spectral\_cluster文件夹**

spectral\_clustering——谱聚类算法

1. **statisticalAnalysis文件夹**

统计RECS和ATUS聚类分类别的相关特征的数据信息

两个子文件里是每个特征的统计信息。

RECS\_total是五个特征聚类三个类别的RECS的统计信息

ATUS\_total是五个特征聚类三个类别的ATUS的统计信息

1. **behavior文件夹**

里面是和ATUS相关的人的行为活动的数据信息

ReNewbehavior 是按照论文把人的一天二十四小时的活动分为九类后，根据ATUS聚类类别（3类），分别统计的三类活动表格