**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。**

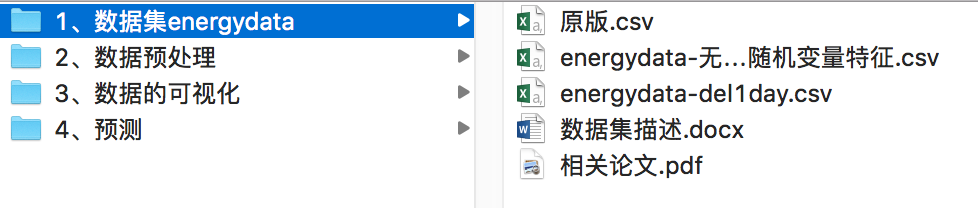
原版.csv——无任何删减原版数据集

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. **数据预处理文件夹**

**数据集子文件夹**

原版.csv——无任何删减原版数据集

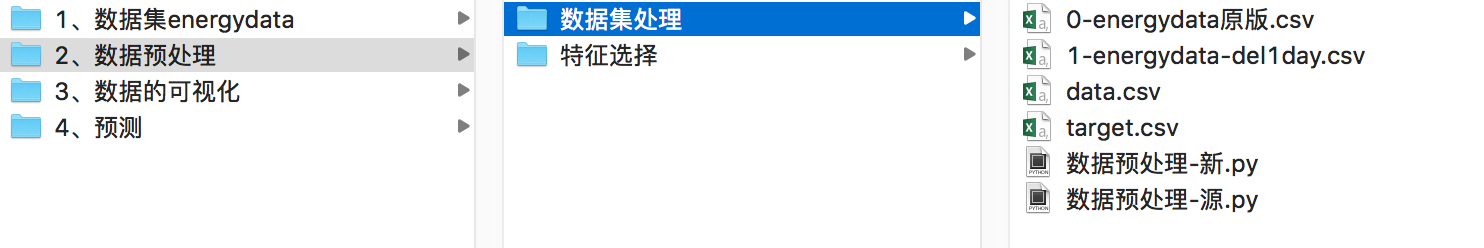
energydata-del1day.csv——删除第一天和倒数第一天的残缺数据

data.csv——除appliances外其他特征

target.csv——appliances特征列

数据预处理-源.py——

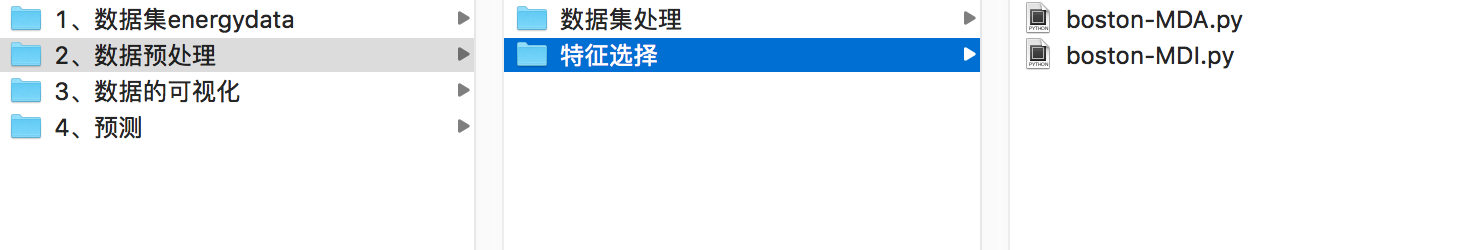
数据预处理-新.py——



**特征选择子文件**

Boston-MDA.py——

Boston-MDI.py——



1. **数据可视化文件夹**

**随机森林决策树可视化子文件夹**

Data-200.csv——除appliances外其他特征的200行数据

feature\_names-26.csv——26个特征

target-200.csv——appliances特征列的200行数据

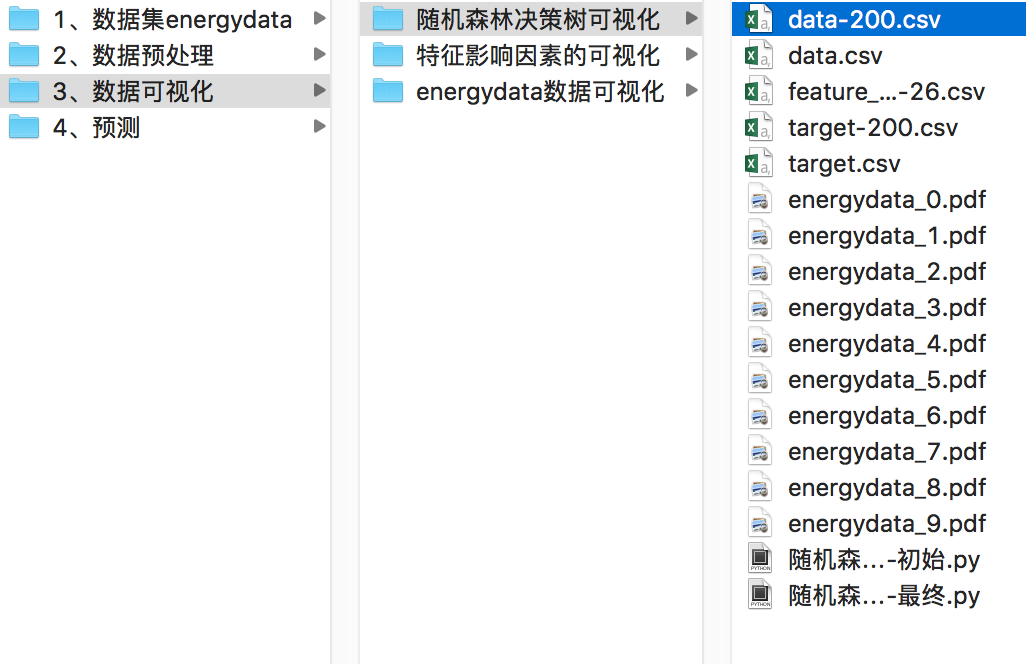
data.csv——除appliances外其他特征

target.csv——appliances特征列

energydata\_0-9.pdf——得到的多个决策树图

随机森林可视化-最终.py——随机森林可视化的最终代码

随机森林可视化-初始.py——随机森林可视化的初始化代码



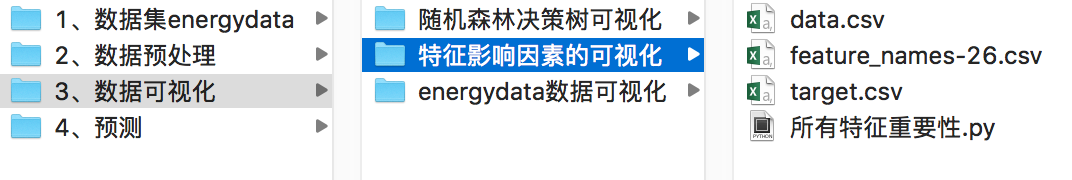
**特征影响因素的可视化子文件夹**

data.csv——除appliances外其他特征

feature\_name-26.csv——26个特征

target.csv——appliances特征列

特征重要性.py——计算特征重要性及可视化代码

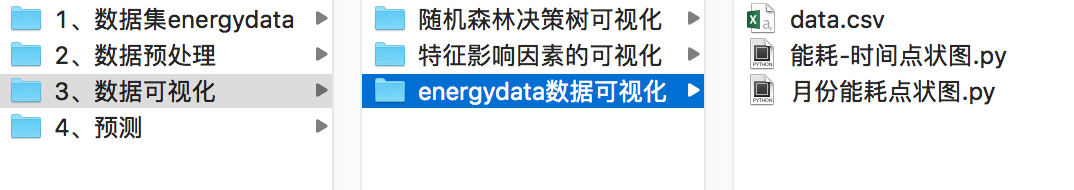


**energydata数据可视化子文件夹**

data.csv——除appliances外其他特征

月份能耗点状图.py——月份能耗点状图代码

能耗-时间点状图.py——能耗-时间点状图代码



1. **模型拟合调参文件夹**

data.csv——除appliances外其他特征

target.csv——appliances特征列

调参.py——调参过程代码

predict.py——预测代码

