

实战案例3：充电桩故障分类与检测

作者：Robin

日期：2018/05

数据集来源：[点石](#)

声明：[小象学院](#)拥有完全知识产权的权利；只限于善意学习者在本课程使用，不得在课程范围外向任何第三方散布。任何其他人或机构不得盗版、复制、仿造其中的创意，我们将保留一切通过法律手段追究违反者的权利

1. 案例描述

该项目的目的主要是基于充电桩的各项信号预测分析充电桩的工作状态及影响因素。

2. 数据集描述

- 点石[提供的数据集](#)包含85500条数据记录，每条记录包含8项信息。
- 数据字典
 - S1**: K1K2驱动信号，浮点型
 - S2**: 电子锁驱动信号，浮点型
 - S3**: 急停信号，浮点型
 - S4**: 门禁信号，浮点型
 - S5**: 电压的总谐波失真，浮点型
 - S6**: 电流的总谐波失真，浮点型
 - label**: 正常情况的标签（0代表正常，1代表异常），整型

3. 任务描述

- 使用scikit-learn建立不同的机器学习模型进行充电桩的工作状态预测

4. 主要代码解释

- 代码结构

```
lect03_proj
├── data
│   ├── charging_pile.csv    # CSV数据文件
├── output
│   ├── model_comparison.csv  # 模型比较结果的CSV文件(程序的输出)
│   ├── pred_results.png    # 模型比较结果的png文件(程序的输出)
├── config.py                # 配置文件
├── utils.py                 # 工具类文件
├── main.py                  # 主程序
└── lect03_proj_readme.pdf   # 案例讲解文档
```

- main.py**

使用的模型及相关参数配置。该项目中使用了2个机器学习模型，并为不同的学习模型指定了参数空间。如：kNN，指定了3个k值用于比较对结果的影响：5, 11, 15。

```
def main():
    ...
    model_name_param_dict = {'kNN': [5, 11, 15],
                             'LR': [0.1, 1, 10]}
    ...
```

- **main.py**

选择最优模型，`np.argmax(list)`返回list中最大值的索引号，如`np.argmax([5, 10, 2])`的返回值是1，因为第1个位置上的值最大，为10。之后根据该索引号获取最优的准确率及模型。

```
def train_test_model(X_train, y_train, X_test, y_test, param_range, model_name='SVM'):
    ...
    # 记录最优模型
    best_idx = np.argmax(scores)
    best_acc = scores[best_idx]
    best_model = models[best_idx]
    ...
```

5. 案例总结

- 该项目通过预测充电桩的工作状态，巩固并实践了使用scikit-learn搭建简单的预测模型：
 - 机器学习流程
 - scikit-learn中常用预测模型的使用，如：kNN, logistic regression等
 - 参数的选择

6. 课后练习

- 参考随堂代码，试着使用不同的参数空间观察对结果的影响。

参考资料

1. [scikit-learn官方教程](#)
2. [通过scikit-learn理解机器学习](#)