# 用电量预测模型

罗磊 2018-07-04

## GBDT模型

将数据样本按照7：3的比例分割为训练集和测试集，使用上一时刻的数据来预测下一时刻的数据，使用GBDT模型对样本训练集进行训练，然后在测试集上进行测试。GBDT模型的参数设置为：

表格 1. GBDT模型参数设置

|  |  |
| --- | --- |
| 参数名称 | 参数值 |
| n\_estimators | 300 |
| learning\_rate | 0.01 |
| min\_samples\_split | 100 |
| min\_samples\_leaf | 200 |
| max\_depth | 4 |
| subsample | 1 |
| random\_state | 10 |
| loss | ‘lad’ |

结果如下：



求得的平均绝对百分偏差为MAPE = 0.1611。

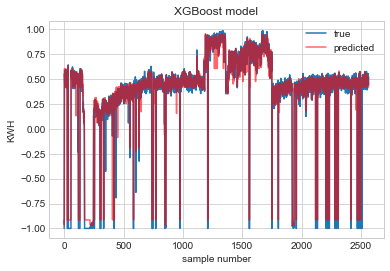
## XGBoost模型

同GBDT模型一样，XGBoost模型也使用上一时刻的数据来预测下一时刻的数据，参数设置为：

表格 2. XGBoost模型参数设置

|  |  |
| --- | --- |
| 参数名称 | 参数值 |
| n\_estimators | 100 |
| learning\_rate | 0.1 |
| seed | 0 |
| loss | ‘mse’ |

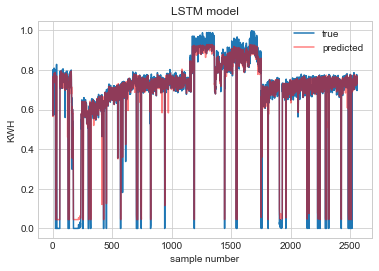
结果如下：



求得的平均绝对百分偏差为MAPE = 0.1609。

## LSTM模型

使用Keras框架搭建该LSTM模型，首先进行数据归一化，然后按照7：3的比例，构造训练集和测试集。设置预测序列长度seq\_len = 10，即使用前10位用电量数据去预测接下来时刻的用电量数据。模型迭代次数epochs = 1000，使用批式随机梯度下降法，设置batch\_size = 3000。对测试集预测结果如下：



求得的平均绝对百分偏差为MAPE = 4.8434。

## K折检验

下面均使用K折检验判断XGBoost模型和LSTM的准确度情况，其中参数*k*-fold = 10。得到下表：

表格 3. XGBoost和LSTM两模型进行K折检验得到的MAPE结果比较

|  |  |  |  |
| --- | --- | --- | --- |
| XGBoost | | LSTM | |
| 0 | 0.007934 | 0 | 0.2237 |
| 1 | 0.005522 | 1 | 0.3739 |
| 2 | 0.005753 | 2 | 0.1490 |
| 3 | 0.006681 | 3 | 0.1564 |
| 4 | 0.006625 | 4 | 0.1755 |
| 5 | 0.007439 | 5 | 0.1742 |
| 6 | 0.006131 | 6 | 0.4220 |
| 7 | 0.013443 | 7 | 0.1591 |
| 8 | 0.008696 | 8 | 0.1532 |
| 9 | 0.004889 | 9 | 0.1750 |
| mean | 0.007312 | mean | 0.2162 |

从两模型的MAPE来看， **XGBoost略胜一筹，更适合解决这个问题**，不过还需要后续调参来进一步确认。

## 考察参数seq\_len对模型MAPE的影响

从上图可以看出，随着seq\_len的增大，LSTM的预测准确度呈现先增加后降低的趋势，但在该程序中的表现不如XGBoost（未调参）。