**国能日新功率预测题目说明**

## 公司简介

国能日新科技股份有限公司（简称：国能日新）是一家专业提供新能源软件产品及服务的高新技术企业，其拥有多项专业资质，具备自主研发、生产制造和全方位营销等综合能力，专注于为新能源企业在并网自动化、生产运营、企业信息化等方面提供优质的解决方案。通过专业的技术产品及全面的服务体系为新能源企业的更优发展提供技术保障。

## 赛题任务

通过学习历史一段时间内的wrf预测数据和对应的发电功率数据进行模型训练，结合未来时刻的wrf预测数据，利用训练好的模型预测未来相应时刻的发电功率。

看数据应该是太阳能

## 数据描述

我司将提供若1个电站的历史功率数据，历史气象数据（见实测数据.csv），其具体数据类型如下：

（1）电站历史运行数据，如表1：

表1 电站历史运行数据

|  |  |  |
| --- | --- | --- |
| 电站历史运行数据 | | |
| 时间 | 历史电站实际辐照度 | 历史电站实际功率 |

（2）气象信息数据(分气象A、B、C、D四个文件)，具体如表2：

表2 气象信息数据

|  |  |  |  |
| --- | --- | --- | --- |
| **光** | | | |
| **气象数据包A** | **气象数据包B** | **气象数据包C** | **气象数据包D** |
| 特征量1 | 特征量1 | 特征量1 | 特征量1 |
| 特征量2 | 特征量2 | 特征量2 | 特征量2 |
| 特征量3 | 特征量3 | 特征量3 | 特征量3 |
| 特征量4 | 特征量4 | 特征量4 | 特征量4 |
| 特征量5 | 特征量5 | 特征量5 | 特征量5 |
| 特征量6 | 特征量6 | 特征量6 | 特征量6 |
| 特征量7 | 特征量7 |  |  |
| 特征量8 | 特征量8 |  |  |
| 特征量9 | 特征量9 |  |  |
| 特征量10 | 特征量10 |  |  |
| 特征量11 | 特征量11 |  |  |
| 特征量12 | 特征量12 |  |  |
| 特征量13 | 特征量13 |  |  |
| 特征量14 | 特征量14 |  |  |
| 特征量15 | 特征量15 |  |  |
| 特征量16 | 特征量16 |  |  |
| 。。。 |  |  |  |
| 特征量79 |  |  |  |

这些数据都是光。可能是4个模型分别给的。有随机云层影响。可能有交互作用。不知道有没有温、压、湿、风作用。

相关性大表

one hot coding

额外维度还有哪些

“一对多”再清洗

说明：

1. 气象信息数据提供了四套完全不同的数据，可以分别单独作为输入进行训练，也可自由组合作为模型的输入；
2. 实际功率中的负值是因为机组在发电不足时自身会消耗电能。
3. 实际数据由于通信等问题存在异常值已经进行剔除。

## 评价目标

**考核目标为：日单装机偏差考核分数的平均值，即先计算每天所有考核点的总和，然后对所有日的结果求平均。**

偏差计算方式：



其中： 是点数；

是96点；

是第i点可用功率预测值；

是第i点的实际功率。

当为零时，当在装机容量的3%以内，则不予考核；当超出装机容量的3%时，误差值按100%计算。

当为零时，当在装机容量的3%以内，则不予考核；当超出装机容量的3%时，误差值按100%计算。

考核分数计算：



**日单装机考核分数=*sum（Scorei）/cap*, *Scorei*为每个点的考核分数,cap为每个电站的装机容量。**

举例说明，某装机容量为300MW的电站某日的预测功率和实际功率如下，则该日的考核分数为0.545774分，该日的单装机考核分数为0.545774/300=0.001819：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 时间 | 预测功率 | 实际功率 | 绝对偏差 | 考核分 |
| 2018/1/1 0:00 | 0 | 0 | 0 | 0 |
| 2018/1/1 0:15 | 0 | 0 | 0 | 0 |
| 2018/1/1 0:30 | 0 | 0 | 0 | 0 |
| 2018/1/1 0:45 | 0 | 0 | 0 | 0 |
| 2018/1/1 1:00 | 0 | 0 | 0 | 0 |
| 2018/1/1 1:15 | 0 | 0 | 0 | 0 |
| 2018/1/1 1:30 | 0 | 0 | 0 | 0 |
| 2018/1/1 1:45 | 0 | 0 | 0 | 0 |
| 2018/1/1 2:00 | 0 | 0 | 0 | 0 |
| 2018/1/1 2:15 | 0 | 0 | 0 | 0 |
| 2018/1/1 2:30 | 0 | 0 | 0 | 0 |
| 2018/1/1 2:45 | 0 | 0 | 0 | 0 |
| 2018/1/1 3:00 | 0 | 0 | 0 | 0 |
| 2018/1/1 3:15 | 0 | 0 | 0 | 0 |
| 2018/1/1 3:30 | 0 | 0 | 0 | 0 |
| 2018/1/1 3:45 | 0 | 0 | 0 | 0 |
| 2018/1/1 4:00 | 0 | 0 | 0 | 0 |
| 2018/1/1 4:15 | 0 | 0 | 0 | 0 |
| 2018/1/1 4:30 | 0 | 0 | 0 | 0 |
| 2018/1/1 4:45 | 0 | 0 | 0 | 0 |
| 2018/1/1 5:00 | 0 | 0 | 0 | 0 |
| 2018/1/1 5:15 | 0 | 0 | 0 | 0 |
| 2018/1/1 5:30 | 0 | 0 | 0 | 0 |
| 2018/1/1 5:45 | 0 | 0 | 0 | 0 |
| 2018/1/1 6:00 | 0 | 0 | 0 | 0 |
| 2018/1/1 6:15 | 0 | 0 | 0 | 0 |
| 2018/1/1 6:30 | 0 | 0 | 0 | 0 |
| 2018/1/1 6:45 | 0 | 0 | 0 | 0 |
| 2018/1/1 7:00 | 0 | 0 | 0 | 0 |
| 2018/1/1 7:15 | 0 | 0 | 0 | 0 |
| 2018/1/1 7:30 | 0 | 0 | 0 | 0 |
| 2018/1/1 7:45 | 0 | 0 | 0 | 0 |
| 2018/1/1 8:00 | 0 | 0 | 0 | 0 |
| 2018/1/1 8:15 | 0 | 0.052 | 0 | 0 |
| 2018/1/1 8:30 | 0 | 12 | 1 | 0.06 |
| 2018/1/1 8:45 | 23.2471 | 3.684 | 0.841529 | 0.074568 |
| 2018/1/1 9:00 | 54.063 | 26.4347 | 0.511039 | 0.084078 |
| 2018/1/1 9:15 | 70.2819 | 46.5843 | 0.337179 | 0.048206 |
| 2018/1/1 9:30 | 87.5821 | 76.2037 | 0.129917 | 0 |
| 2018/1/1 9:45 | 105.423 | 105.306 | 0.00111 | 0 |
| 2018/1/1 10:00 | 122.723 | 127.22 | 0.036643 | 0 |
| 2018/1/1 10:15 | 140.564 | 143.76 | 0.022737 | 0 |
| 2018/1/1 10:30 | 157.864 | 157.52 | 0.002179 | 0 |
| 2018/1/1 10:45 | 175.705 | 168.732 | 0.039686 | 0 |
| 2018/1/1 11:00 | 193.546 | 178.849 | 0.075935 | 0 |
| 2018/1/1 11:15 | 209.224 | 187.936 | 0.101747 | 0 |
| 2018/1/1 11:30 | 223.28 | 195.832 | 0.122931 | 0 |
| 2018/1/1 11:45 | 237.337 | 201.957 | 0.149071 | 0 |
| 2018/1/1 12:00 | 250.852 | 206.025 | 0.178699 | 0 |
| 2018/1/1 12:15 | 259.502 | 209.254 | 0.193632 | 0 |
| 2018/1/1 12:30 | 262.206 | 211.583 | 0.193066 | 0 |
| 2018/1/1 12:45 | 265.449 | 213.064 | 0.197345 | 0 |
| 2018/1/1 13:00 | 268.693 | 213.783 | 0.20436 | 0.005857 |
| 2018/1/1 13:15 | 267.071 | 212.246 | 0.205282 | 0.007054 |
| 2018/1/1 13:30 | 261.124 | 209.946 | 0.195991 | 0 |
| 2018/1/1 13:45 | 255.718 | 206.396 | 0.192877 | 0 |
| 2018/1/1 14:00 | 250.312 | 202.645 | 0.19043 | 0 |
| 2018/1/1 14:15 | 241.121 | 198.359 | 0.177347 | 0 |
| 2018/1/1 14:30 | 228.686 | 187.5 | 0.180098 | 0 |
| 2018/1/1 14:45 | 216.252 | 186.328 | 0.138376 | 0 |
| 2018/1/1 15:00 | 203.818 | 182.057 | 0.106767 | 0 |
| 2018/1/1 15:15 | 189.221 | 174.615 | 0.07719 | 0 |
| 2018/1/1 15:30 | 171.92 | 161.937 | 0.058068 | 0 |
| 2018/1/1 15:45 | 154.62 | 147.165 | 0.048215 | 0 |
| 2018/1/1 16:00 | 137.861 | 130.605 | 0.052633 | 0 |
| 2018/1/1 16:15 | 118.939 | 111.777 | 0.060216 | 0 |
| 2018/1/1 16:30 | 97.3134 | 88.699 | 0.088522 | 0 |
| 2018/1/1 16:45 | 75.6882 | 62.004 | 0.180797 | 0 |
| 2018/1/1 17:00 | 54.6036 | 35.588 | 0.348248 | 0.040474 |
| 2018/1/1 17:15 | 38.3847 | 16.76 | 0.563368 | 0.069739 |
| 2018/1/1 17:30 | 27.5721 | 3.87333 | 0.85952 | 0.090922 |
| 2018/1/1 17:45 | 16.2189 | 0 | 1 | 0.064876 |
| 2018/1/1 18:00 | 5.4063 | 0 | 0 | 0 |
| 2018/1/1 18:15 | 0 | 0 | 0 | 0 |
| 2018/1/1 18:30 | 0 | 0 | 0 | 0 |
| 2018/1/1 18:45 | 0 | 0 | 0 | 0 |
| 2018/1/1 19:00 | 0 | 0 | 0 | 0 |
| 2018/1/1 19:15 | 0 | 0 | 0 | 0 |
| 2018/1/1 19:30 | 0 | 0 | 0 | 0 |
| 2018/1/1 19:45 | 0 | 0 | 0 | 0 |
| 2018/1/1 20:00 | 0 | 0 | 0 | 0 |
| 2018/1/1 20:15 | 0 | 0 | 0 | 0 |
| 2018/1/1 20:30 | 0 | 0 | 0 | 0 |
| 2018/1/1 20:45 | 0 | 0 | 0 | 0 |
| 2018/1/1 21:00 | 0 | 0 | 0 | 0 |
| 2018/1/1 21:15 | 0 | 0 | 0 | 0 |
| 2018/1/1 21:30 | 0 | 0 | 0 | 0 |
| 2018/1/1 21:45 | 0 | 0 | 0 | 0 |
| 2018/1/1 22:00 | 0 | 0 | 0 | 0 |
| 2018/1/1 22:15 | 0 | 0 | 0 | 0 |
| 2018/1/1 22:30 | 0 | 0 | 0 | 0 |
| 2018/1/1 22:45 | 0 | 0 | 0 | 0 |
| 2018/1/1 23:00 | 0 | 0 | 0 | 0 |
| 2018/1/1 23:15 | 0 | 0 | 0 | 0 |
| 2018/1/1 23:30 | 0 | 0 | 0 | 0 |
| 2018/1/1 23:45 | 0 | 0 | 0 | 0 |

最终结果对所测电站的测试数据集日单装机偏差考核求平均。

## 数据使用说明：

1. 实测数据中由于实际运行情况存在缺失，属于正常情况，其中实际辐照度数据仅供参考和用于异常数据剔除，不可直接作为模型输入。
2. 提供的四种气象数据时间长度最为2018年1月1日至今；所以统一以2018年1月1日至2019年12月31日为时间范畴（若气象缺失以气象的时间长度为准），该时间段之外的数据为无效数据可进行删除。
3. 气象数据可单独使用，也可自由组合按时间匹配后使用。

（4） 2018年1月1日至2018年12月31日的气象数据与实测数据按时间匹配后作为模型的训练数据，进行模型训练和交叉验证，其中气象数据为模型输入，实测数据为模型训练目标；2019年1月1日至2019年12月31日的气象数据作为模型的测试输入数据，最终预测结果为2019年1月1日至2019年12月31日的预测功率数据（气象数据缺失的以气象数据为准，如气象数据只到12月27日，则预测结果也只到12月27日）。

**（5）最终提交数据为2019年1月1日至2019年12月31日的与气象数据对应的预测数据，及整个建模过程的思路和说明文档。**

**（6）限时一周内提交。**