光伏功率预测大赛

https://www.dcjingsai.com/common/cmpt/国能日新光伏功率预测大赛\_竞赛信息.html

## 竞赛背景

光伏发电具有波动性和间歇性，大规模光伏电站并网运行可能对电力系统的安全稳定经济运行造成影响。对光伏电站的输出功率进行准确率预测，有助于调度部门统筹安排常规能源和光伏发电的协调配合，及时调整调度计划，合理安排电网运行方式。因此，本题旨在通过利用气象信息、历史数据、组件信息等，通过机器学习、人工智能方法，预测未来发电功率，为进一步为光伏发电功率提供准确的预测结果。

## 任务

通过学习一段时间内的环境数据和对应的光伏发电功率，训练模型，通过给定某时间点预测的环境数据，预测该时间点的光伏发电功率。

## 数据

\*注 :报名参赛或加入队伍后，可获取数据下载权限。

训练集数据提供了4个电场的脱敏后的环境数据和电场实际辐照度和电场发电功率。测试集数据提供了4个电场的脱敏后的环境数据，需要利用这些数据预测每个时间点的光伏发电功率。

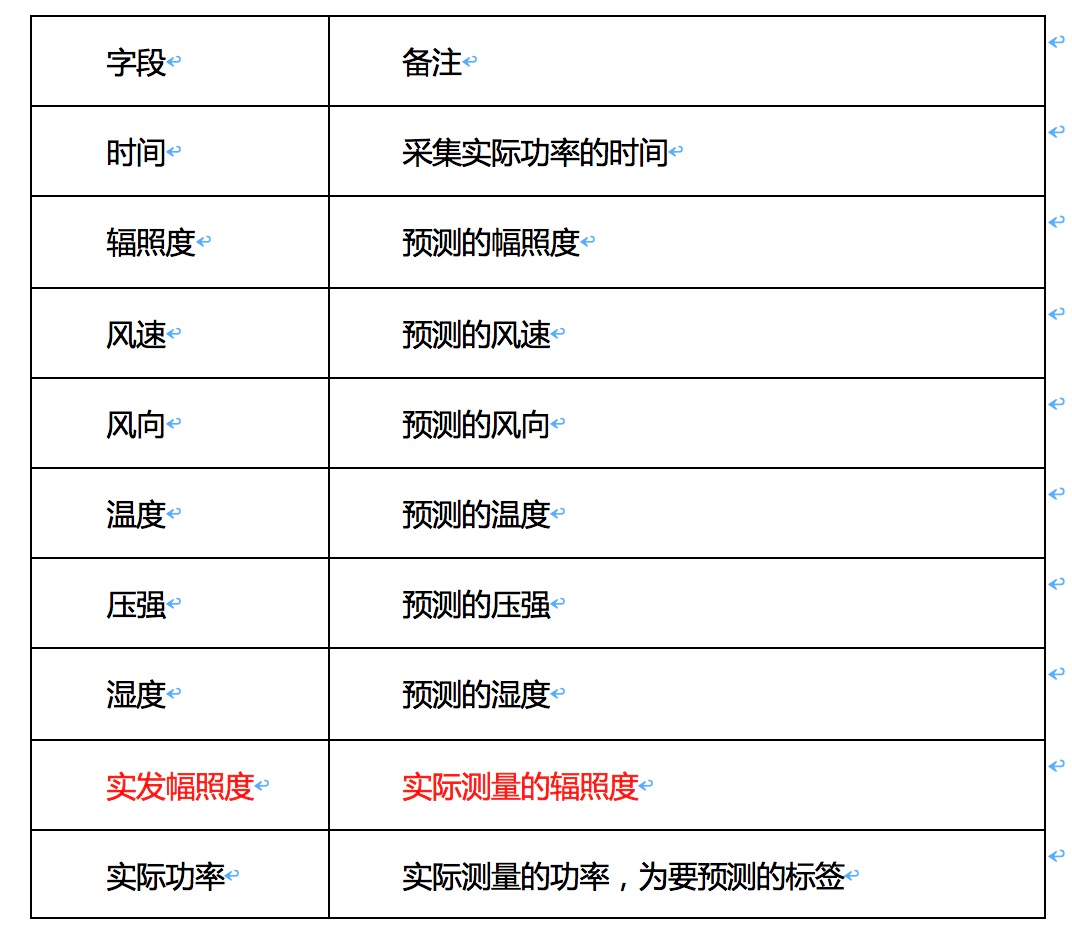
**注意：**

为了贴近实际应用，环境数据提供的是预测值，不是实测值，训练集中的电场实际辐照度和电场实际发电功率为脱敏后的实测值。

训练集和测试集的描述如下：

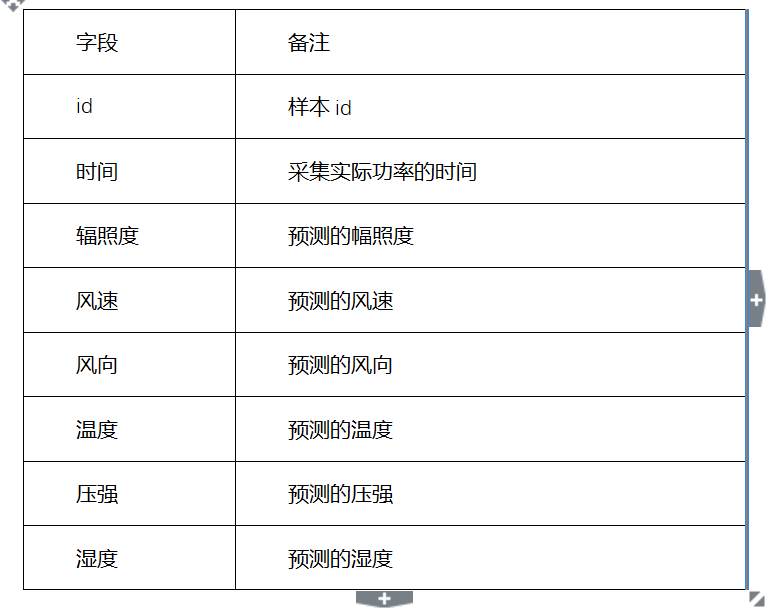
训练集有train\_1.csv, train\_2.csv, train\_3.csv, train\_4.csv共4个文件，测试集有test\_1.csv,test\_2.csv,test\_3.csv,test\_4.csv共4个文件，分别为电场1，电场2，电场3，电场4的训练集数据和测试集数据。

训练集字段如下：



测试集数据字段，除无字段实发幅照度和实际功率外，增加了id列为样本id，每条记录的样本id是唯一值，其它字段同训练集数据。

测试集字段如下：



补充说明：

1.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。

2.实发幅照度中的负值视为噪声数据。

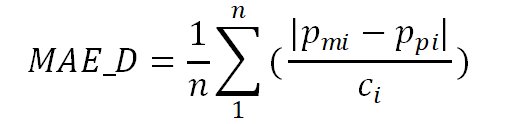
## 评分标准

评分算法

other

评分算法分三步计算：

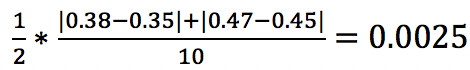
i.日平均绝对偏差的计算公式如下：



其中， MAE\_D为每日的平均绝对偏差，n为当天实际功率的值大于等于装机功率3%（即实际功率值大于等于Ci\*0.03）的测量样本的数量；为实际功率；为预测功率；Ci为装机功率，电站1，2，3，4的装机功率的值分别为10，10，40，50。

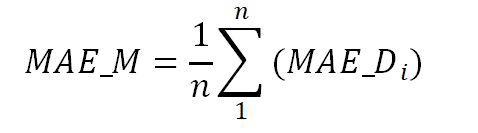
举例说明，某一天电站1预测的功率如下：



则当日平均绝对偏差为

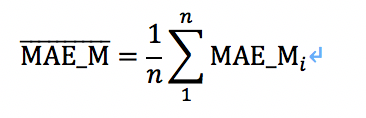
即实际功率为0.16和0.23的两行未参与计算最终分数。

ii.每月的绝对平均偏差的平均值计算方法如下：



MAE\_M为每月的绝对平均偏差，n为当月天数，为第i天的日平均绝对偏差。

iii.计算单个电场月MAE的平均值



其中， 为第 月的月平均偏差，n 为总月数。

最终分数为4个电场 的的平均值：

其中，为第个电场的值。

本次比赛未划分A/B榜

## 结果提交

