

The background is a dark blue field filled with digital motifs. On the left, a white wind turbine stands on a circular base that resembles a technical diagram or a radar screen. The base has concentric circles and various small icons. Scattered throughout the background are vertical strings of binary code (0s and 1s) and a network of thin, light-blue lines connecting small white dots, creating a sense of data flow and connectivity.

风机叶片开裂故障预警

老板来碗酸辣粉



目录

DIRECTORY

1

团队介绍

2

比赛思路

3

算法模型

4

方案潜力

第一节

1

团队介绍

2

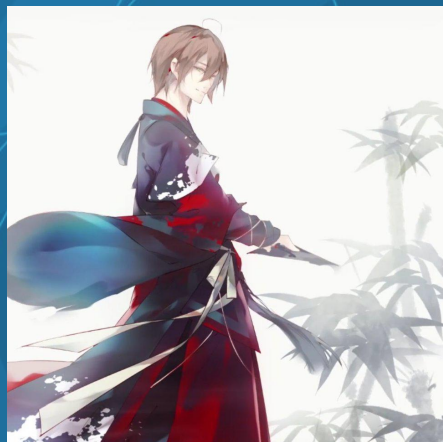
比赛思路

3

算法模型

4

方案潜力



宋颖

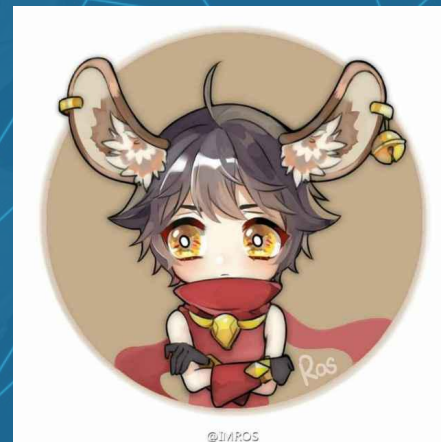
华南理工大学
电气工程及其自动化



罗宾理

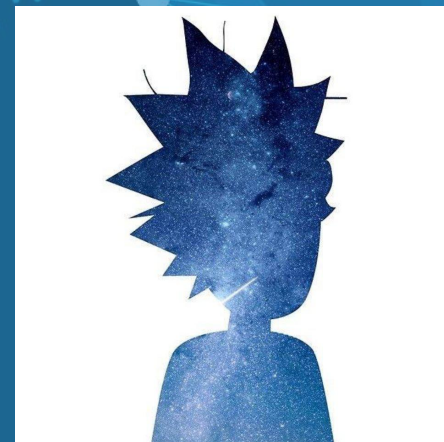
中南大学
软件工程

2018 KDD CUP 冠军
2018 工业AI 数据竞
赛 亚军
2018 IJCAI 广告转化
率预估 季军



彭涛

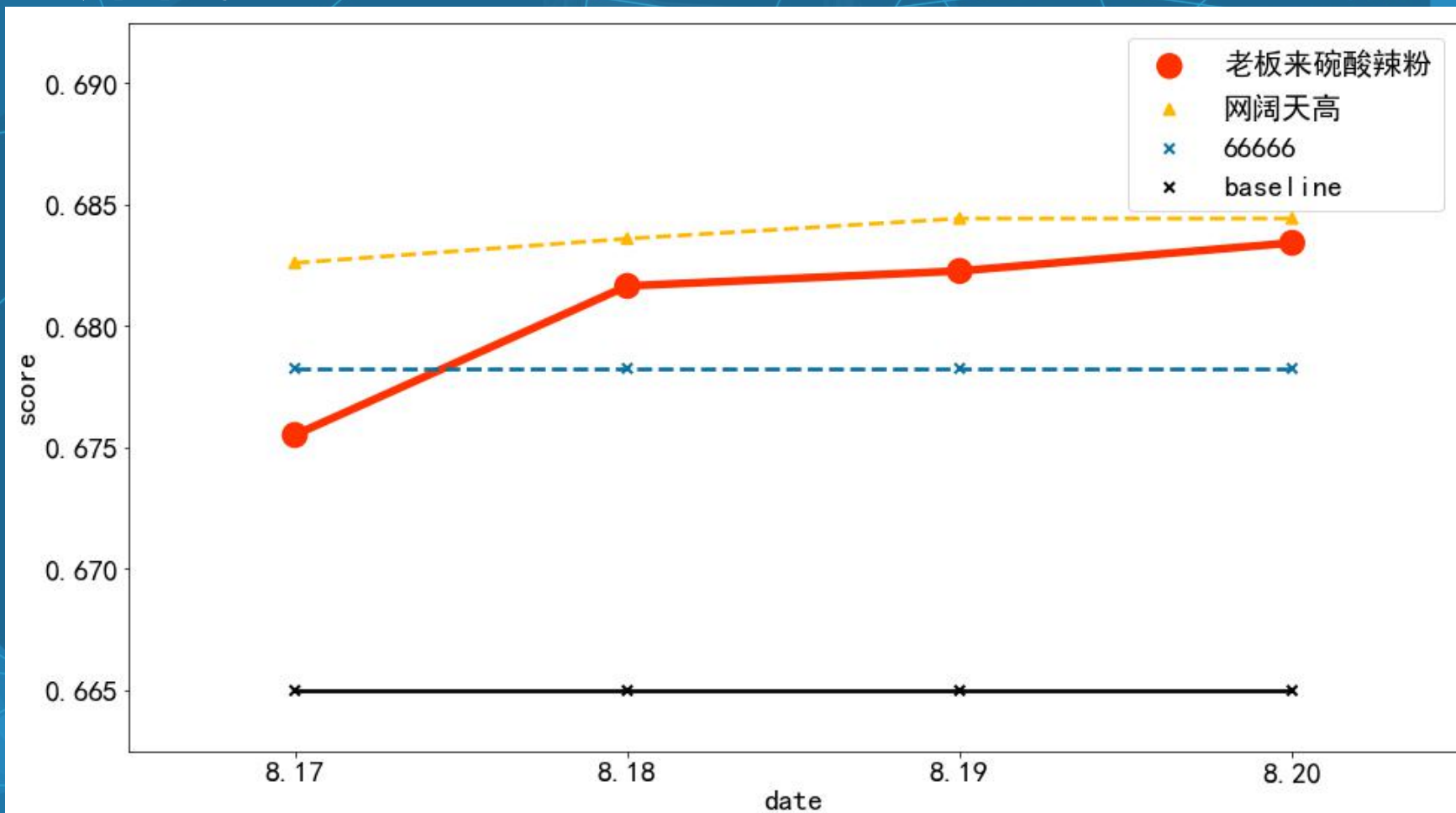
南京大学
应用统计专业



董文辉

电子科技大学
电子通信与工程

B榜分数曲线



第二节

1

团队介绍

2

比赛思路

3

算法模型

4

方案潜力

赛题回顾

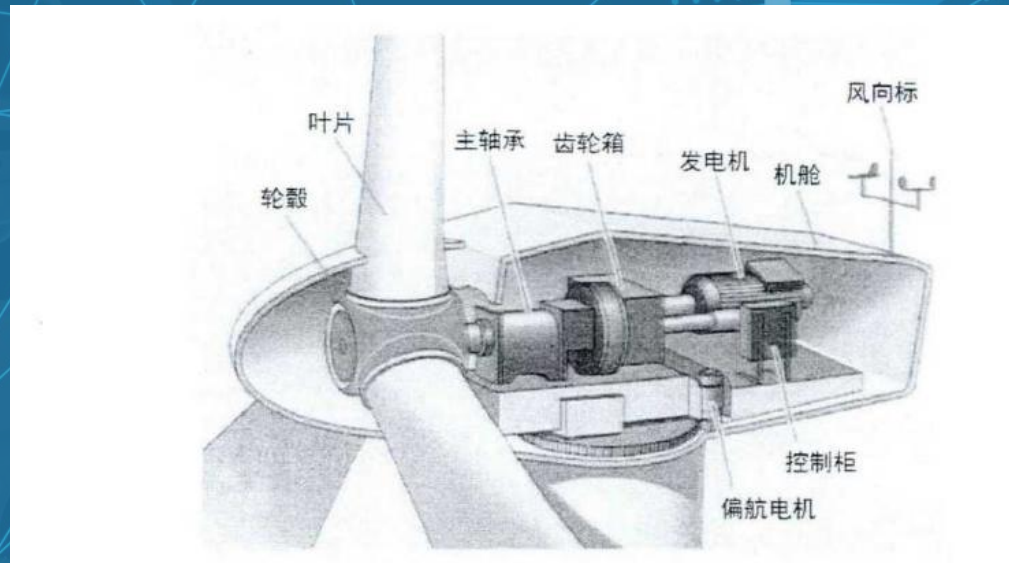
任务描述：

利用SCADA采集的风机工况数据，对于风机未来一周内是否会发生故障进行预测

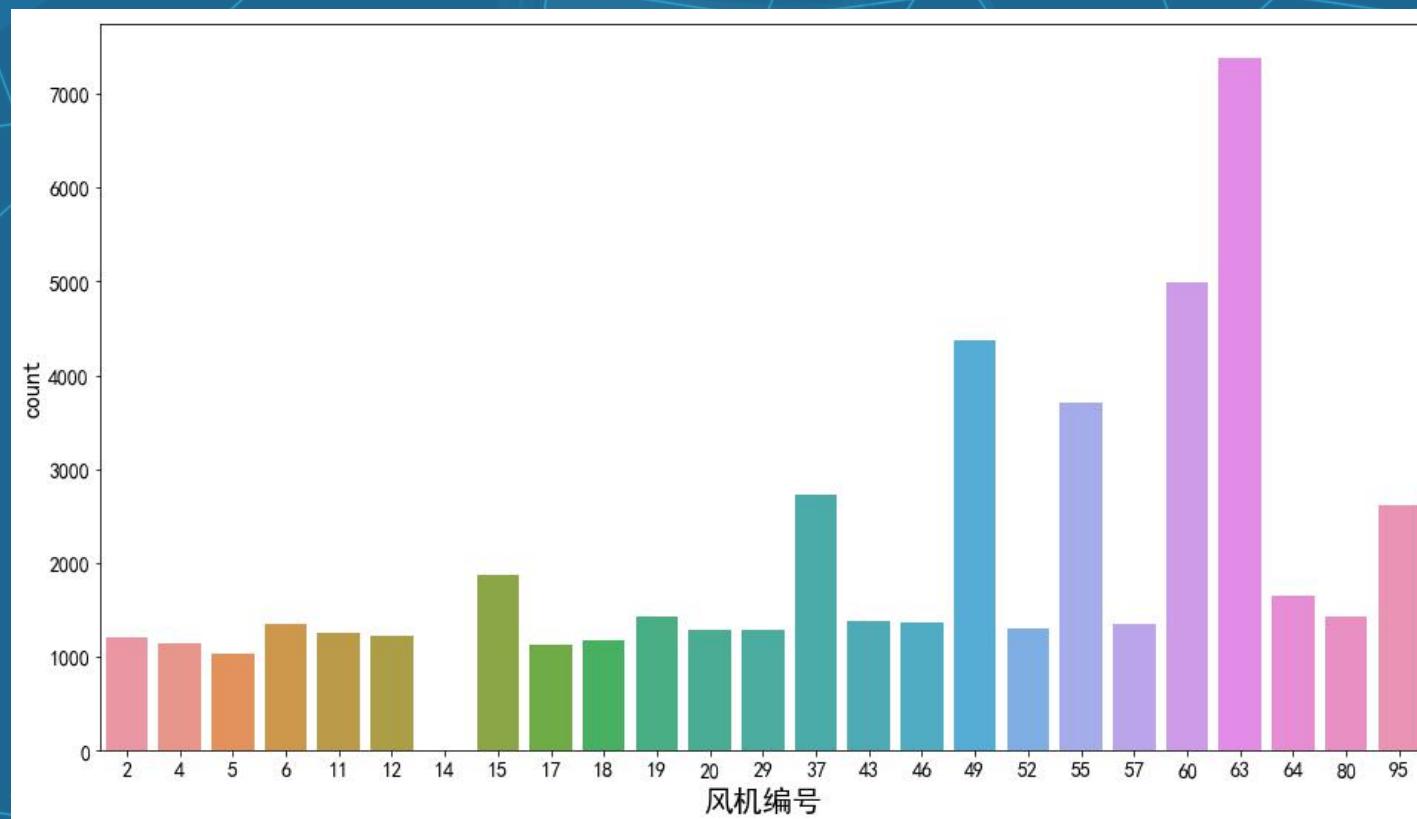
赛题理解：

由于风机开裂故障多发在盛丰期，由此推断风机故障的主要原因是谐振。谐振与风机的机械特性密切相关，每个风机的共振频率都不大一致。

除此之外，谐振的特点是，能够在短时间内造成巨大的破坏力。因此给出的data与label并不是完全准确的对应关系（脏数据）。很有可能的是，事故前6天的数据都是正常的，只有最后那一段时间存在异常

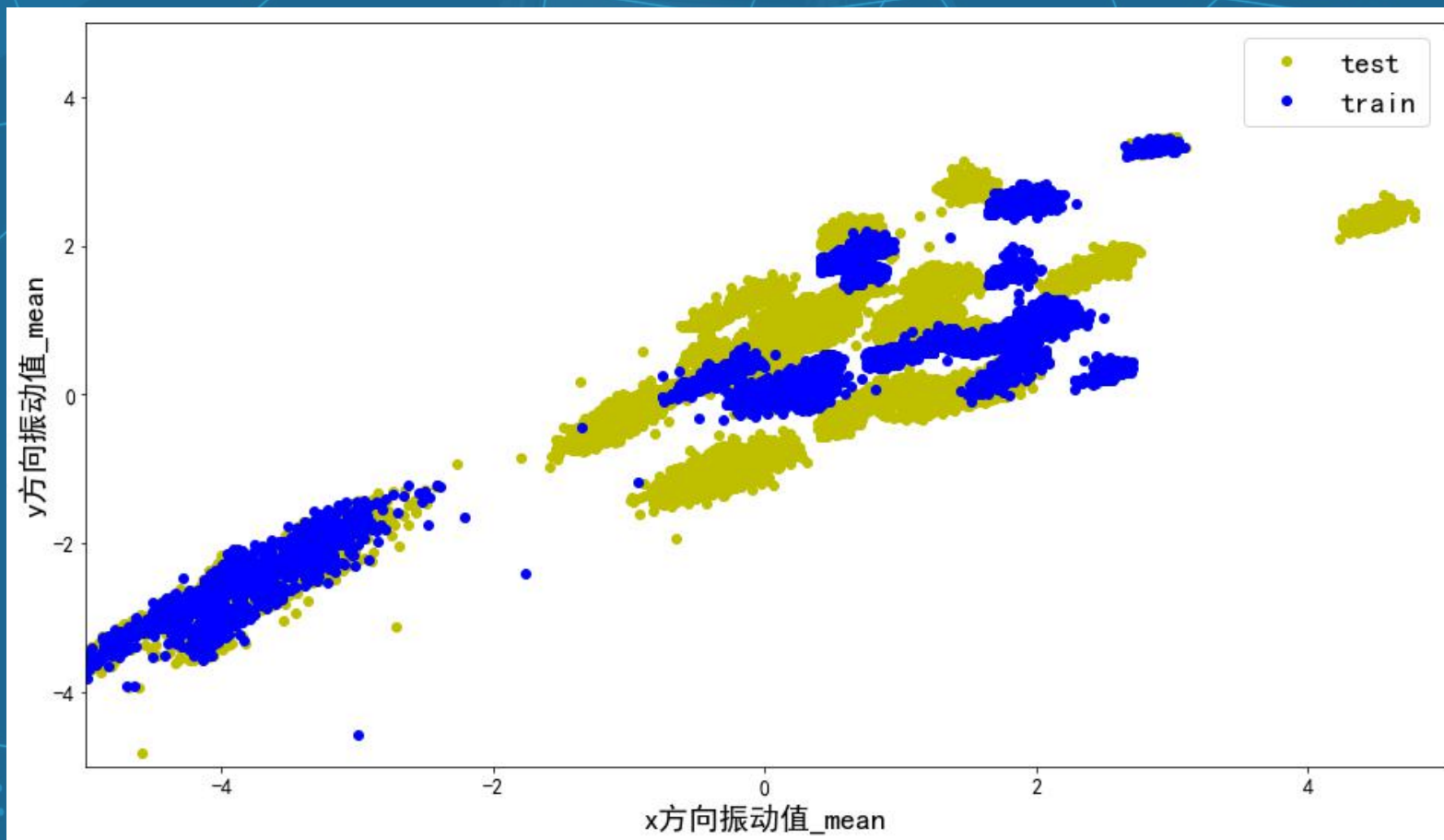


train数据分布

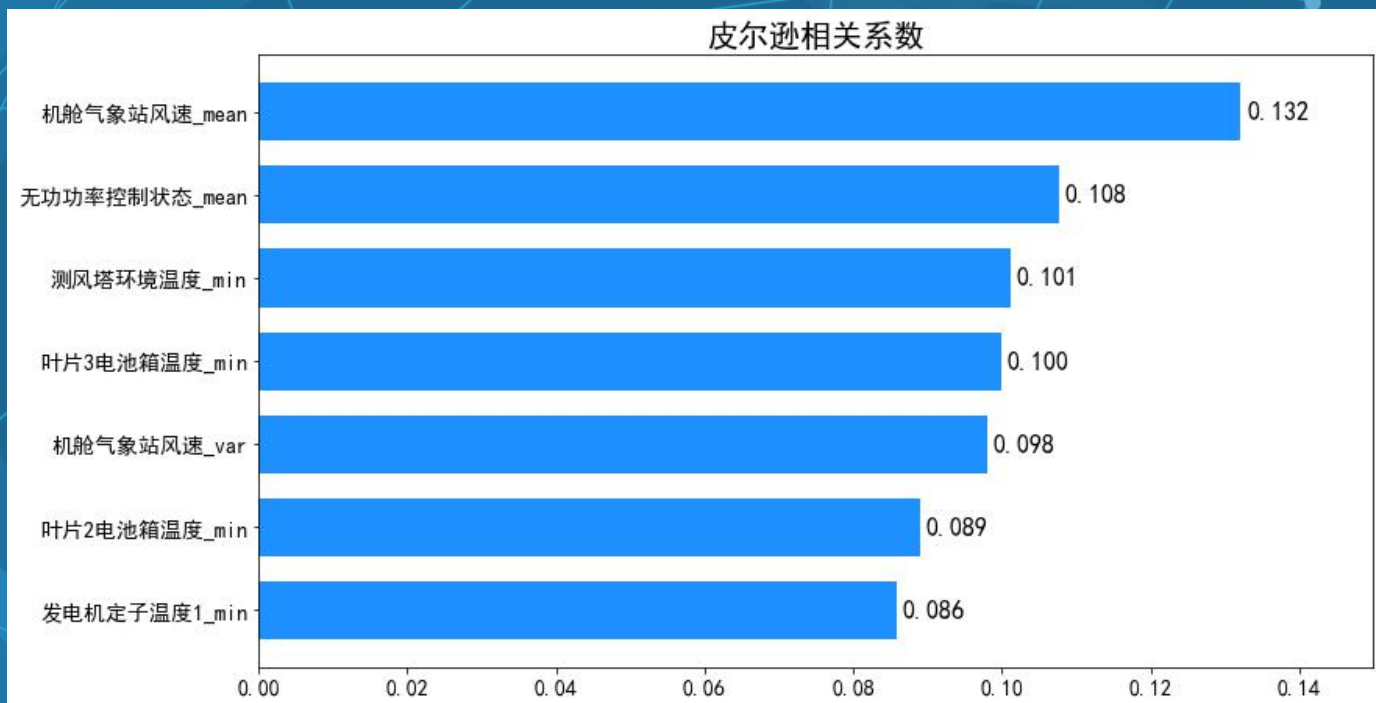


test: 无风机编号

特征聚类



皮尔逊相关系数分析



- 机舱气象站风速_mean
- 无功功率控制状态_mean
- 测风塔环境温度_min
- 叶片3电池箱温度_min
- 机舱气象站风速_var
- 发电机定子温度1_min

第三节

1

团队介绍

2

比赛思路

3

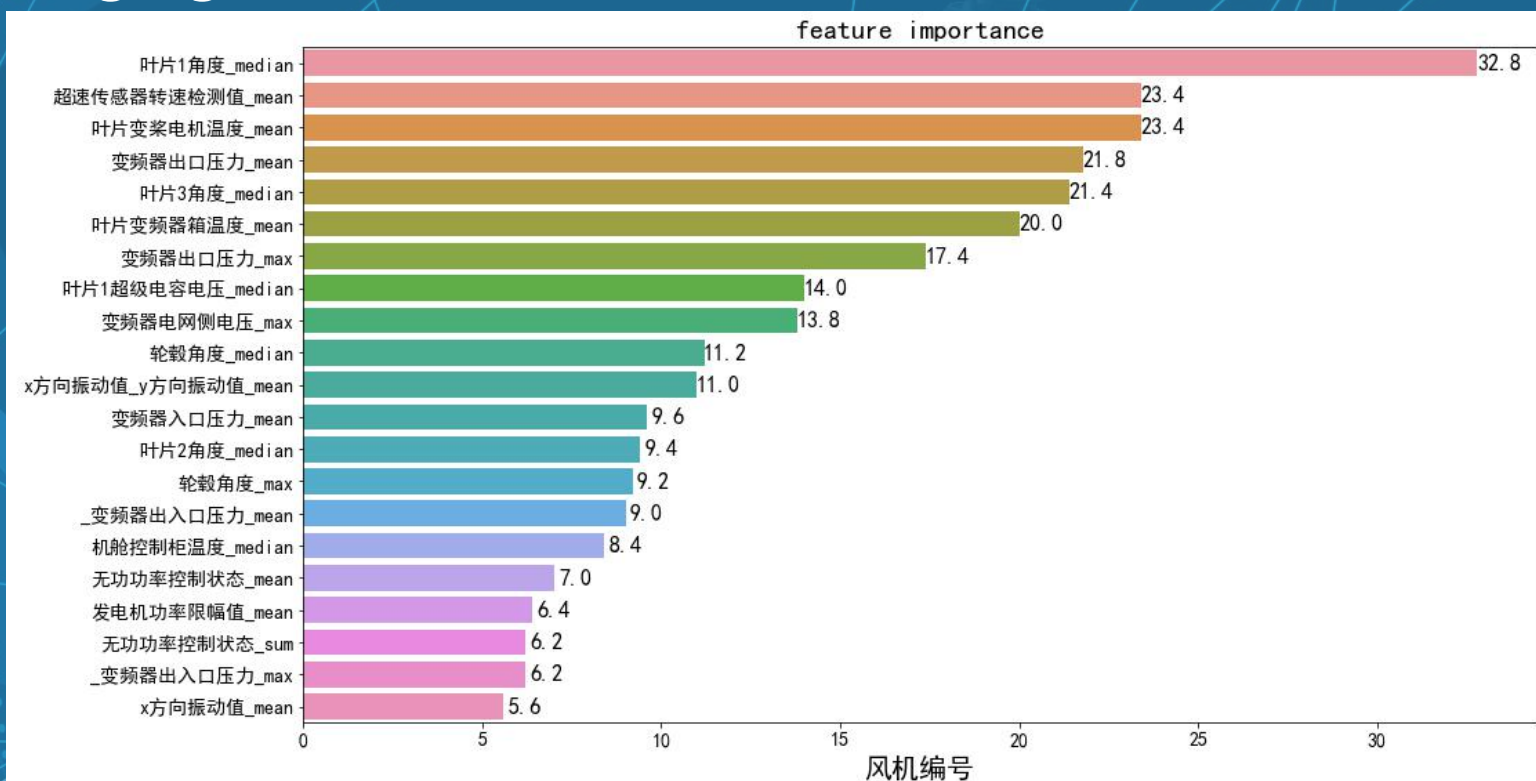
算法模型

4

方案潜力

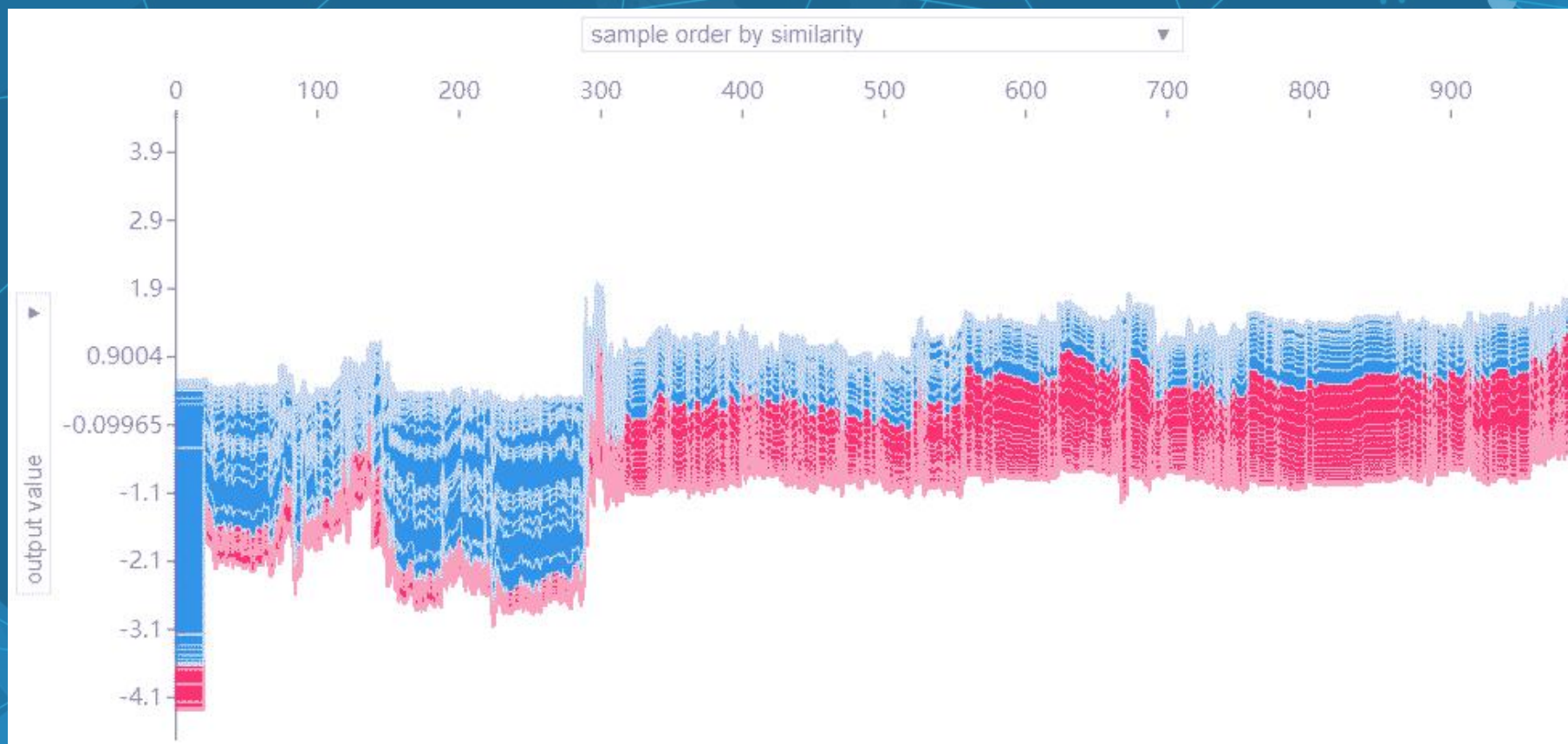
特征选取&lightgbm模型训练

- 选取相关系数前5的特征
- lightgbm模型特征重要性前40的特征



- 叶片角度
- 超速传感器转速检测值
- 叶片变桨电机温度
- 变频器出口压力
- 叶片变频器箱温度
- 叶片超级电容电压

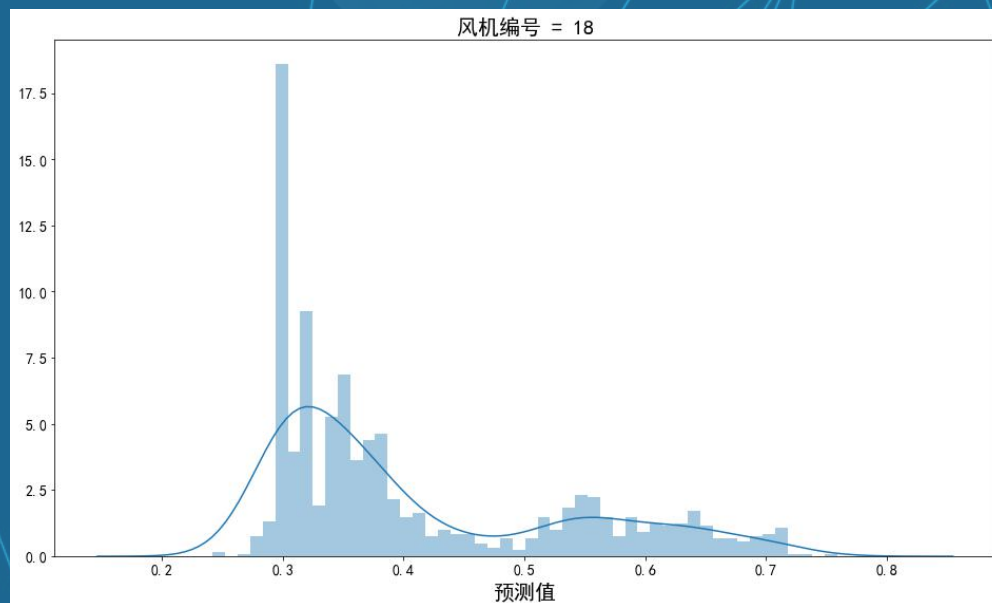
lightgbm模型结果





KNN&DeepFFM带来的提升

KNN在某些风机编号的预测上表现得非常出色



DeepFFM由于与lgb模型的差异性，融合之后能够带来一定程度上的提升

三者融合，能在最大限度上保证泛化能力

第四节

1

团队介绍

2

比赛思路

3

算法模型

4

方案潜力

端到端模型
响应时间测试
能够做到与SCADA同步展现风机状态

SCADA
原始数据

0.026s



最终结果