

Week 2

Azure本身服务器运维所提供的数据，其中包含的数据量更大而广，更有利于模型的泛化。

通过这一组数据建立一个预测器，以预测机器可能出现failures或者errors的概率。

你将接触到不同维度的数据，包括时间序列、属性以及每个ID具备非结构化的数据。其中包含：

- Machine conditions and usage: 机器的运行条件，例如：从传感器收集的数据。
- Failure history: 机器或机器内组件的故障历史。
- Maintenance history: 机器的维修历史，例如错误代码、以前的维护活动或组件更换。
- Machine features: 机器的特性，例如引擎尺寸、品牌和型号、位置。
-

任务要求

使用这一套数据集，第一阶段独立学习并使用XGBoost完成故障的预测器

第二阶段进一步学习Light GBM，并使用LightGBM构建你的最终模型。

参考

XGBoost

<https://www.jianshu.com/p/ac1c12f3fba1>

LightGBM

<https://zhuanlan.zhihu.com/p/99069186>
