



链滴

# 单指标时序异常检测学习笔记

作者: [Dadong-Zhang](#)

原文链接: <https://ld246.com/article/1600357211333>

来源网站: [链滴](#)

许可协议: [署名-相同方式共享 4.0 国际 \(CC BY-SA 4.0\)](#)

# 单指标时序异常检测学习笔记

## Metis开源项目

Metis采用Tsfresh和XGBoost组件进行时序异常检测，采用了三个时间窗口去进行预测

## 特征数据格式

时间窗默认180分钟，数据粒度1分钟

数据 一时间窗	上一时间窗	当前时间	
当前时刻A	√	√	
昨日同期时刻	√	√	√
上周同期时刻	√	√	√

## 算法思路

先使用三种非监督方法进行检测，当某一方法输出异常时使用全监督的XGBOOST进行检测

## 非监督算法

### Statistic方法

3sigma法，检测当前时刻点数值与历史数值均值相差是否大于三倍标准差

### EWMA方法

指数加权移动平均法，相比于3sigma法给予历史数据一个指数级递减的权重，距离当前数据点越近权重越大

阈值为一下公式所示，其中 $\alpha$ 为学习率

### Polynomial Interpolation方法

多项式回归

### IForest方法（未使用）

孤立森林方法

## 全监督训练方法

### XGBOOST

XGBOOST是GBDT (Gradient Boosting Decision Tree) 的高效实现方法, 并在Kaggle竞赛上获奖多。

## 参考

[腾讯织云Metis异常检测原理剖析](#)

[通俗理解kaggle比赛大杀器xgboost](#)

[终于有人说清楚了--XGBoost算法](#)

[基于时间序列的异常检测算法小结](#)

## ARIMA (差分整合移动平均自回归模型)

ARIMA更是一种类似于EWMA的时间序列预测模型

### 进行预测的一般步骤

1. 数据获取与清洗: 包括对数据进行预处理 (比如利用数据平滑消除数据抖动)
2. 数据预处理: 时序图 (原始图), 自相关、偏自相关 (消除序列的周期性与时间相关性), 差分 (除自增减趋势, 会损失数据)
3. 模型识别: 通过算法分析数据特征来确定使用何种检测算法或预测算法
4. 模型定阶
5. 参数估计
6. 模型验证