LogicMonitor

LogicMonitor-AI
异常检测

AIOps挑战赛答辩

姚睿　　2018/05/19

LogicMonitor

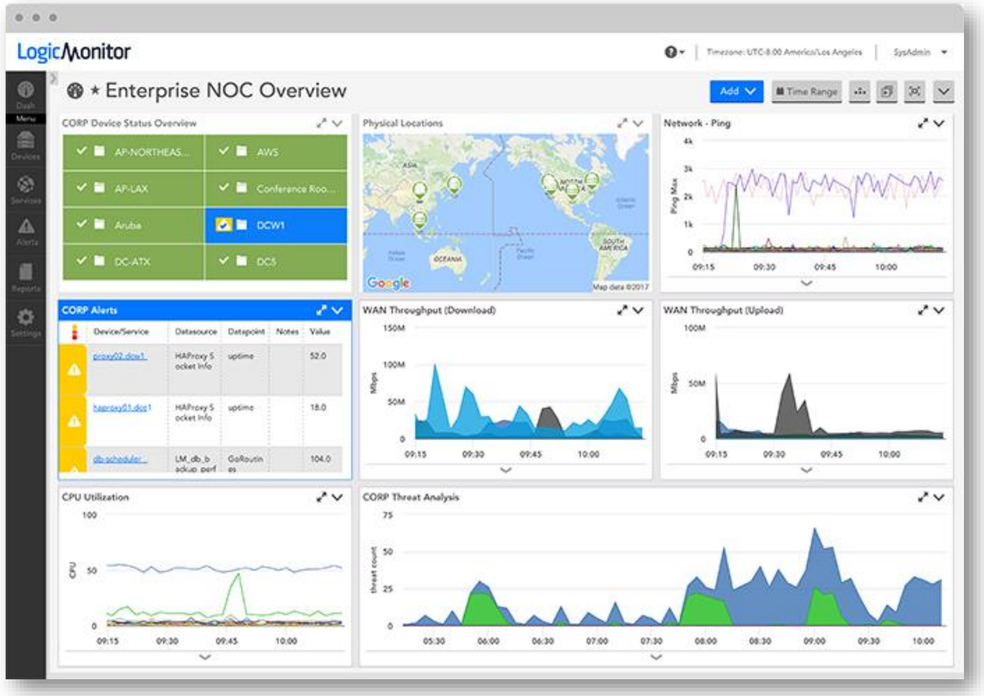
团队介绍

云智易控科技（LogicMonitor）

LogicMonitor是业界领先的基于SaaS模式的企业级IT性能监控平台

目前服务1000多家客户，遍布全球

Monitor Everything



LogicMonitor-AI Team

专注于AIOps产品研发

推动AIOps在监控领域的落地

LogicMonitor

Content

LogicMonitor

需求分析与设计原则

时间序列异常检测需求分析

- 有标注
- 禁用手工干预
- 禁用未来信息
- 时效性

设计原则

- 监督学习
- 自动化
- 流处理
- 普适性

LogicMonitor

方案总览

流程

预处理

特征提取

模型训练

异常检测

技术栈

Keras

Numpy

Pandas

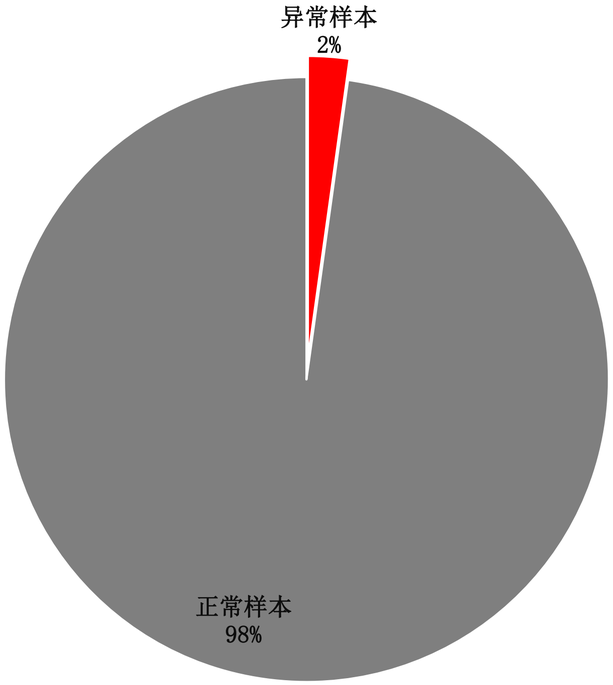
Sklearn

LogicMonitor

预处理：平衡类别

挑战

异常检测中正负样本极度不平衡



方案1：正常样本欠采样

方案2：欠采样 + 集成学习

方案3：异常样本过采样

LogicMonitor

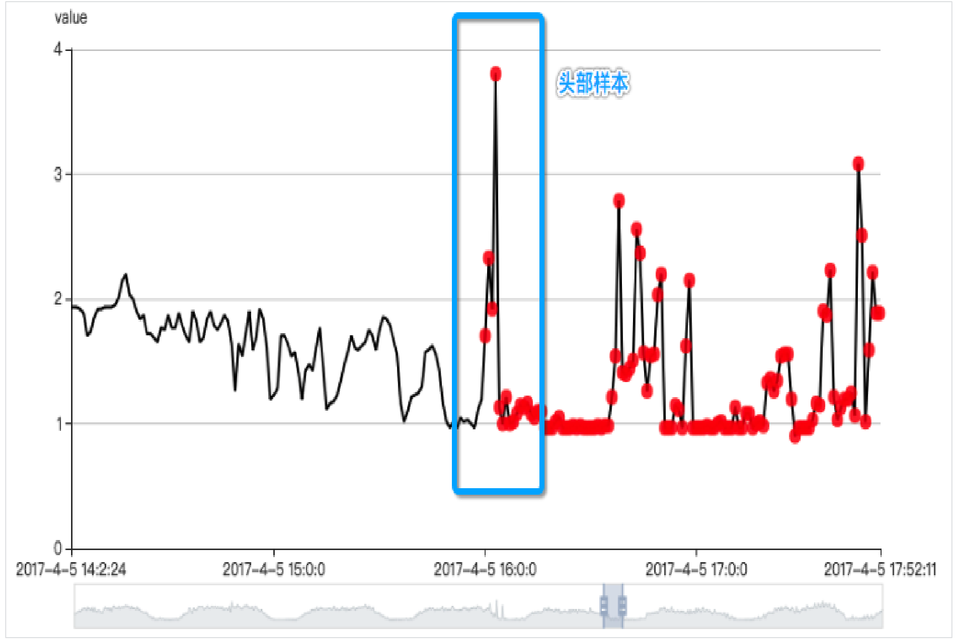
预处理：样本权重调整

时效性需求

需要在首个异常点出现的后续N个数据点内检出异常

思考

对于一段连续异常段，起始段异常点样本价值最大



方案

增加头部异常样本的权重

LogicMonitor

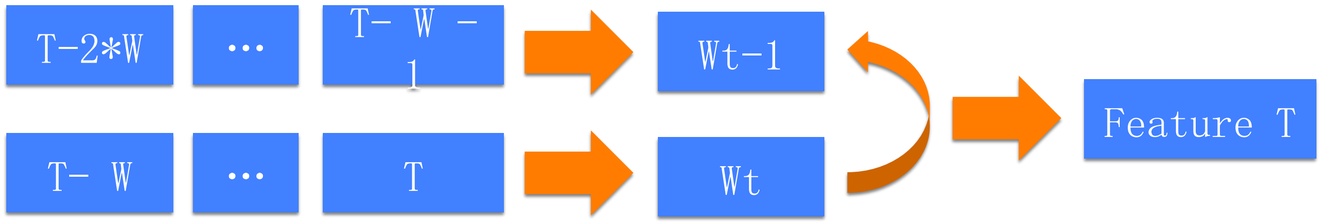
特征提取

时间序列特征提取方案

滑动窗口统计特征

对比特征

滑动窗口 + 对比



LogicMonitor

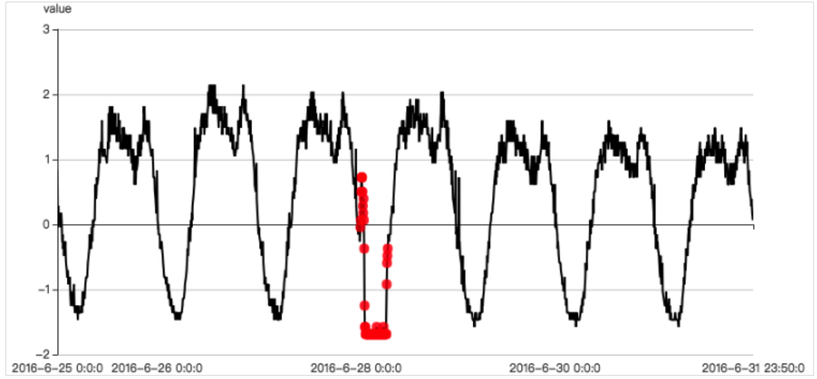
特征提取

统计特征
均值、方差等

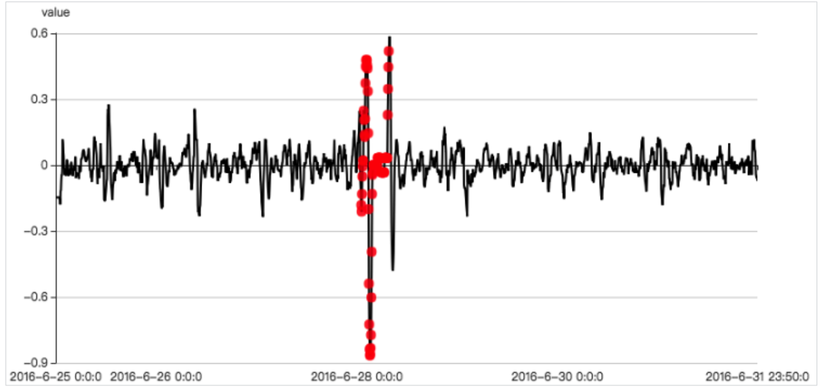
对比特征
差分、变化比例等

组合特征
统计特征 X 对比特征

特征集
多重窗口宽度 X 统计特征 X 对比特征



原始序列数据



特征空间（方差、差分组合特征）

LogicMonitor

模型 选择

需求

- 适应较大的样本量：最低1.7万，最高29万，平均18万
- 能较好的控制过拟合

候选模型

Isolation Forest

对局部异常不敏感

Random Forest

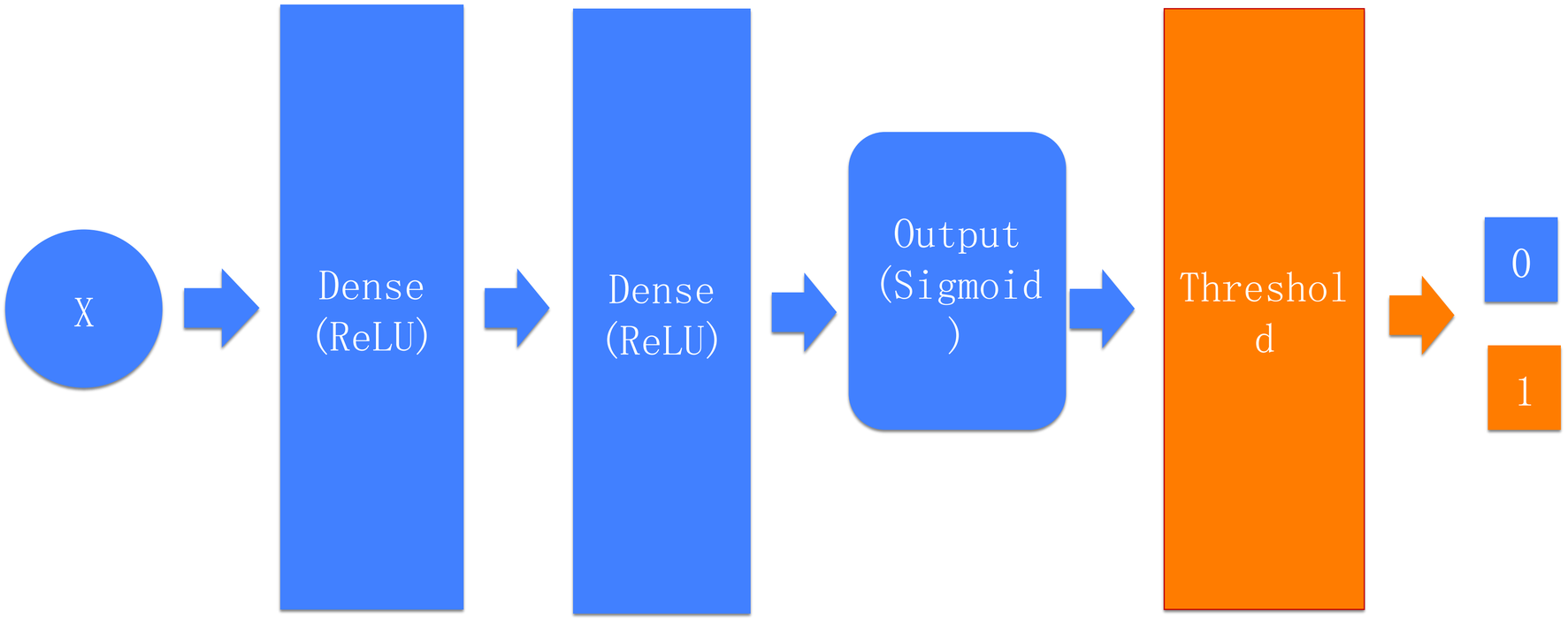
总体表现较好，实测泛化能力略低于DNN

DNN 

模型表达能力强，能适应大数据，泛化能力强

LogicMonitor

模型 DNN



- 通过正则化和Dropout控制模型的泛化性能
- 通过阈值化方法修正样本不均的问题

LogicMonitor

普适性

通用框架适用于所有KPI数据

完全地自动化

不针对单独KPI进行特殊优化

LogicMonitor

评估

不同数据集，模型表现稳定一致，泛化能力优秀

LogicMonitor

改进和展望

引入时间序列周期性特征

集成无监督模型

集成业务规则

Q & A

Thanks