



GOPS2017
Shanghai



GOPS

全球运维大会

2017

上海站

指导单位:  质量中心联盟

主办单位:  信息技术协会  数据中心联盟

大会时间: 2017年11月17日-18日

大会地点: 上海光大会展中心国际大酒店 (上海徐汇区漕宝路67号)





GOPS2017
Shanghai

AI浪潮下的高效运维思考及实践

赵建春@腾讯



GOPS2017
Shanghai

目录



1

AI 是如何工作的？

2

AI 和运维工作结合的思考

3

我们的一些实践案例



GOPS2017
Shanghai

从大量输入中总结出准确预测的规律（模型）

$$y = f(x) = ax + d$$

$$y = f(x) = ax^2 + bx + d$$

$$y = f(x_1, x_2, x_3) = ax_1 + bx_2 + cx_3 + d$$

$$a_1^{(2)} = f(W_{11}^{(1)}x_1 + W_{12}^{(1)}x_2 + W_{13}^{(1)}x_3 + b_1^{(1)})$$

$$a_2^{(2)} = f(W_{21}^{(1)}x_1 + W_{22}^{(1)}x_2 + W_{23}^{(1)}x_3 + b_2^{(1)})$$

$$a_3^{(2)} = f(W_{31}^{(1)}x_1 + W_{32}^{(1)}x_2 + W_{33}^{(1)}x_3 + b_3^{(1)})$$

$$h_{w,b}(x) = a_1^{(3)} = f(W_{11}^{(2)}a_1^{(2)} + W_{12}^{(2)}a_2^{(2)} + W_{13}^{(2)}a_3^{(2)} + b_1^{(2)})$$

x, x_1, x_2, x_3 是我们的输入

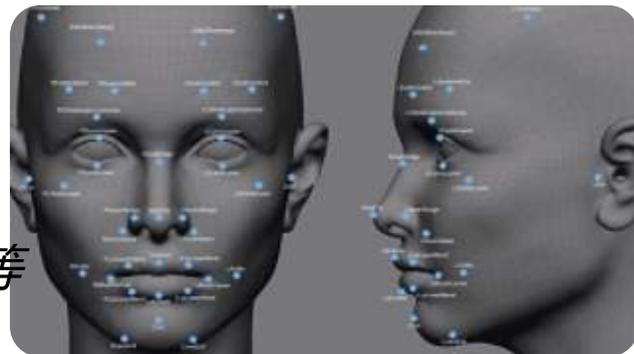
y 是期望的输出

a, b, c, d (W, b) 是我们要求解的参数

f 是转换函数（公式、函数、算法、模型...）也就是我们要找的“规律”



数值型预测
0/1型预测
概率型预测等





GOPS2017
Shanghai

AI和机器学习的一些分类

按学习机制

有监督学习
无监督学习
强化学习

.....

按作用

回 归
分 类
聚 类

.....

是否深度网络

非深度学习
深度学习

.....

AI 从大量输入中通过监督、非监督或强化学习的方式找到对应输出的规律（也就是模型、公式、算法）。

主要算法作用有分类、聚类、回归等。

按照是否使用多层神经网络，又分为传统机器学习和深度学习

监督学习和样本标注



GOPS2017
Shanghai

训练集	正样本	特征
验证集	负样本	
测试集		

通过输入计算机的正、负样本，定义一个**损失函数**，让机器在不断调整模型参数中，达到让损失函数值最小（多数情况下），从而拟合出一个算法模型来进行准确的结果预测。



GOPS2017
Shanghai

损失函数及常见损失函数

0-1损失函数	平方损失函数	绝对损失函数	对数损失函数
0-1 loss function $L(Y, f(X)) = \begin{cases} 1, & Y \neq f(X) \\ 0, & Y = f(X) \end{cases}$	quadratic loss function $L(Y, f(X)) = (Y - f(X))^2$	absolute loss function $L(Y, f(X)) = Y - f(X) $	logarithmic loss function $L(Y, P(Y X)) = -\log P(Y X)$

AI和机器学习提升的指南针；衡量损失和错误程度的函数

预测值和真实值偏差的累积；多数情况损失函数越小越好

一些常见算法工作机制简述



GOPS2017
Shanghai

十大机器学习算法

编号	算法	说明	编号	算法	说明
1	C4.5	分类	6	PageRank	搜索
2	K-means	聚类	7	AdaBoost	分类
3	SVM	分类	8	KNN	分类
4	Apriori	频繁项集	9	贝叶斯	分类
5	EM	最大似然估计	10	CART	分类和回归

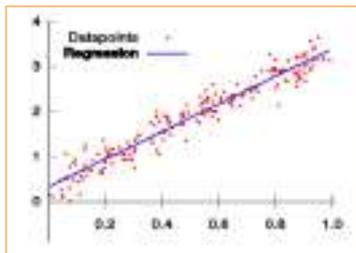
推荐类如逻辑回归、决策树、矩阵分解、CF、Word2Vector、强化学习等
神经网络DNN、CNN、RNN

.....

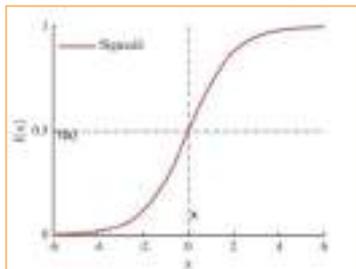
以分类算法居多（聚类也是种特殊分类）

NLP自然语言处理技术，也属于机器学习的一个领域

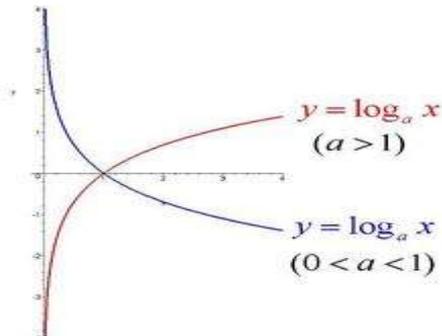
一些常见算法机制简述



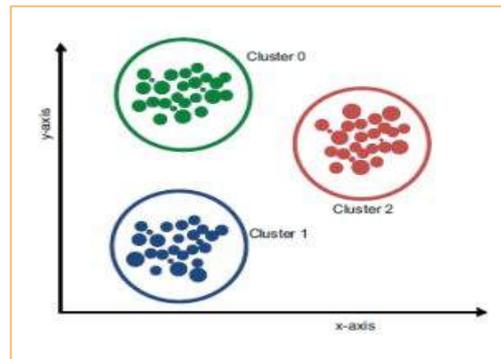
$$y = \zeta(x) = \frac{1}{1+e^{-x}}$$



线性 / 逻辑回归



决策树



- 1、随机选取K个聚类质心点 $\mu_1, \mu_2, \mu_3 \dots \mu_k$
- 2、重复下面过程直到收敛{

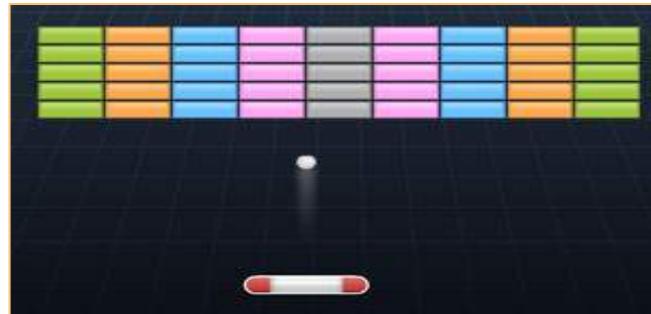
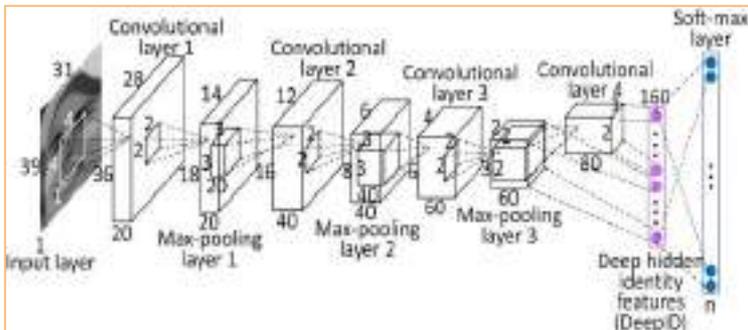
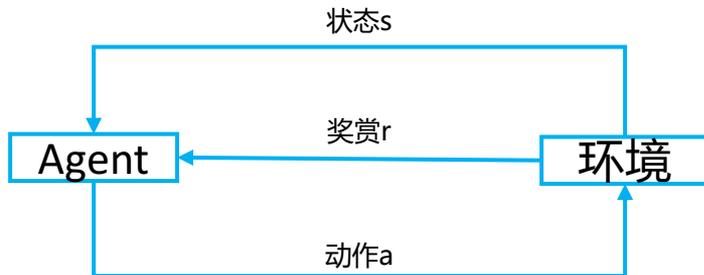
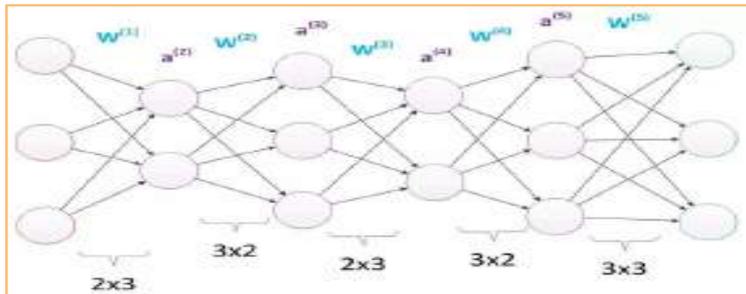
对于每一个样例 i , 计算期应该属于的类
 $c^{(i)} := \arg \min \|x^{(i)} - \mu_j\|^2$
 对于每一个类 j , 重新计算其质心

$$\mu_j := \frac{\sum_{i=1}^m 1\{c^{(i)} = j\}x^{(i)}}{\sum_{i=1}^m 1\{c^{(i)} = j\}}$$
 }

聚类



一些常见算法工作机制简述



深度学习

强化学习

通过文档分类简要认识NLP



GOPS2017
Shanghai

中文分词

- 1、字符串匹配法（正向最大匹配法；逆向最大匹配法；最少切分）
- 2、基于理解分词法（分词的同时进行句法、语义分析）
- 3、基于统计分词法（对相邻共现的各个字的组合的频度进行统计，计算它们的互现信息；HMM和CRF）
- 4、基于规则分词法（与一个“充分大的”机器词典中的词条进行匹配）

关键词提取

- 1、使用TF-IDF算法，找出两篇文章的关键词；（或LDA主题模型）
- 2、每篇文章各取出若干个关键词（比如10个），合并成一个集合，计算每篇文章对于这个集合中的词的词频 / 相对词频

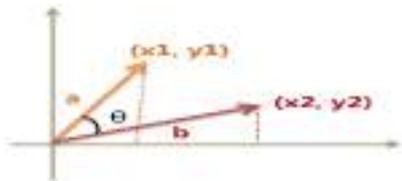
词向量化

文章A : [1, 2, 2, 1, 1, 1, 0]
文章B : [1, 2, 3, 1, 1, 2, 1]

TF-IDF

$$\text{词频}(TF) = \frac{\text{某个词在文章中的出现次数}}{\text{文章的总词数}}$$

关系计算



$$\text{逆文档频率}(IDF) = \log \left(\frac{\text{语料库的文档总数}}{\text{包含该词的文档数} + 1} \right)$$

$$TF - IDF = \text{词频}(TF) \times \text{逆文档频率}(IDF)$$

文档分类

- 1、计算向量之间的余弦相似度，值越大就表示越相似



GOPS2017
Shanghai

目录

1 AI 是如何工作的？



2 AI 和运维工作结合的思考

3 我们的一些实践案例

运维和AI的结合



GOPS2017
Shanghai

- 当前基于人工智能的“自动化运维”某种角度看，很像无人驾驶，还需要很长一段路要走
- 短期基于机器学习的监控和分析 + 基于规则的自动化，可挖掘的空间很
- 利用无监督机制，可以更好的提高效率
- 在对数据的特征建设和归档中走向“智能”，自己解放自己的好帮手



GOPS2017
Shanghai

易于接入AI的运维场景的特点

- 足够的数据量级
- 特征（因素）较齐备
- 特征质量高（提取、清洗容易等）
- 正负样本易抽取
- 积累了正负样本库（标注）
- 有持续的正负反馈

因为它们便于分类



GOPS2017
Shanghai

分类算法应用

- 本质是通过大量数据中的“线索”，找到分类的依据
- “线索”即特征
- 分类算法基本都对特征有较强依赖
- 运维环境有大量相应的场景
- 运维同学最熟悉这些“线索”
- 做AI的“领域专家”但不一定要是AI专家
- 运维场景下的“特征工程”和标注



运维和AI可能的结合点



GOPS2017
Shanghai

1 智能告警

2 网络异常分析

3 程序异常分析

4 关联异常分析

5 变更体检报告

6 硬件故障预测

7 投诉文本聚类等

8 咨询客服机器人

9 数据库参数调优

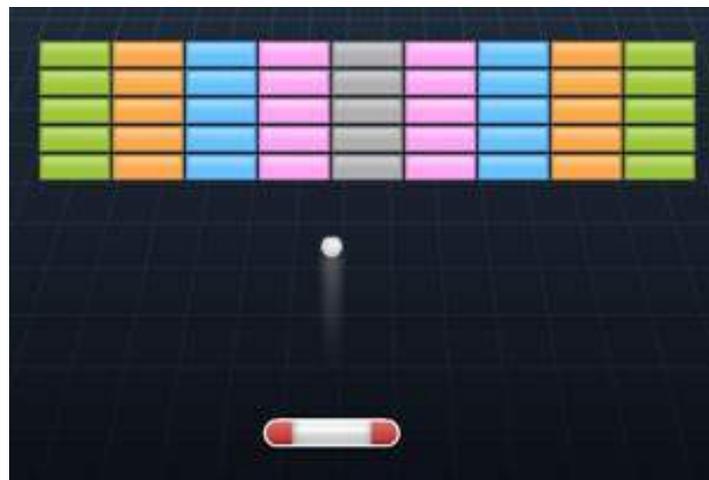
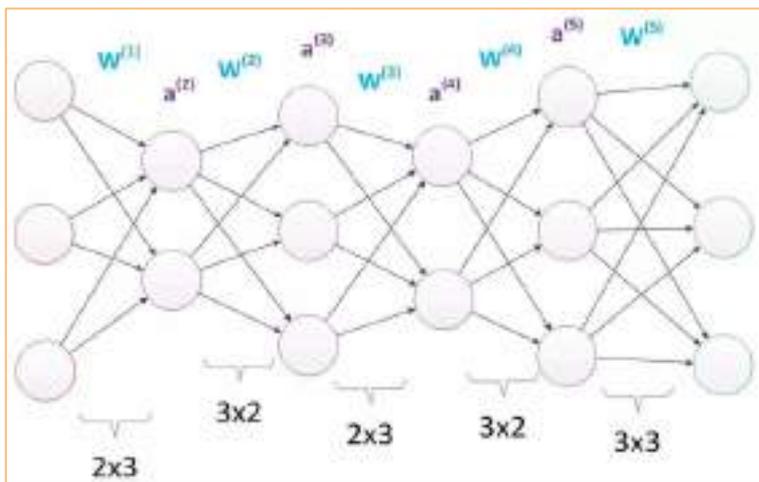
.....



GOPS2017
Shanghai

是否可以更简单些？

深度学习&强化学习



面临的问题和挑战？



GOPS2017
Shanghai

目录

1 AI 是如何工作的？

2 AI 和运维工作结合的思考

➔ 3 我们的一些实践案例



GOPS2017
Shanghai

Monitor时序数据智能监控



Monitor时序类数据智能监控

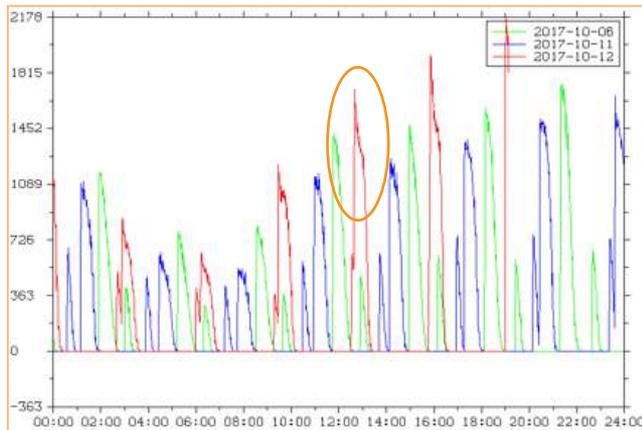
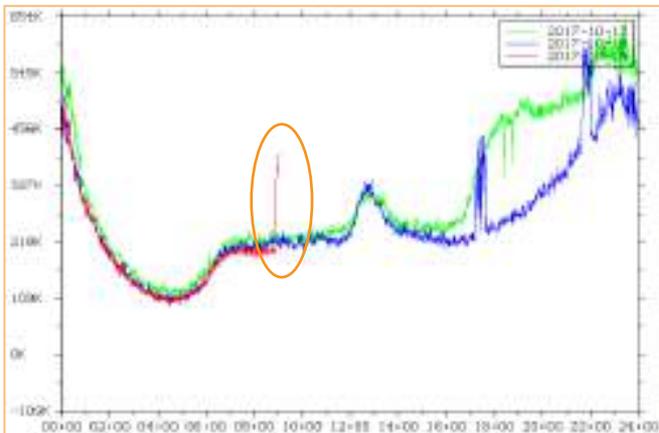
monitor监控数据的特点：

量大：125W个监控点数据，形成125W个视图。

复杂：曲线形态各异。

数据正负样本极不平衡、数据量级大，工作量大。

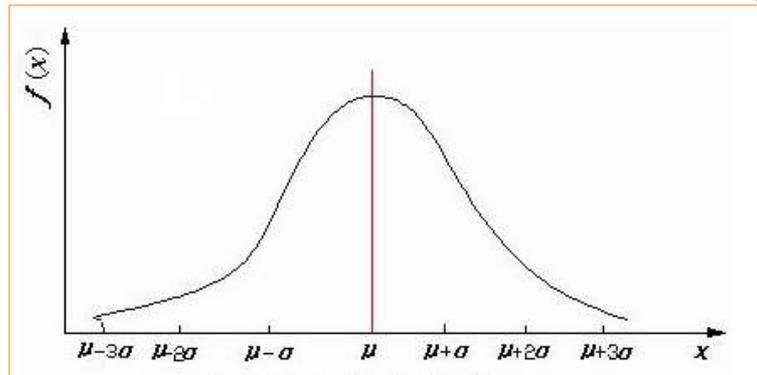
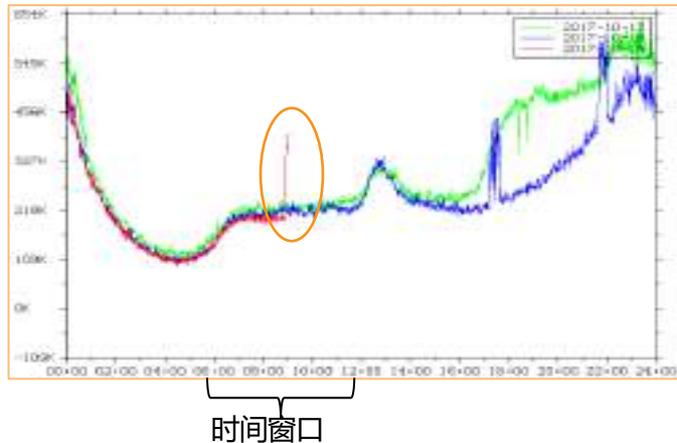
如何高效的去除干扰，“准确”的按照“希望”的方式告警？



第一步：统计判别法



GOPS2017
Shanghai



突破

1. 取消监控告警的阈值、波动率等人为设置规则

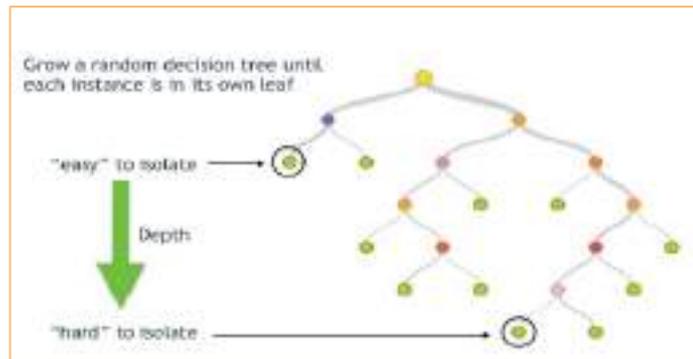
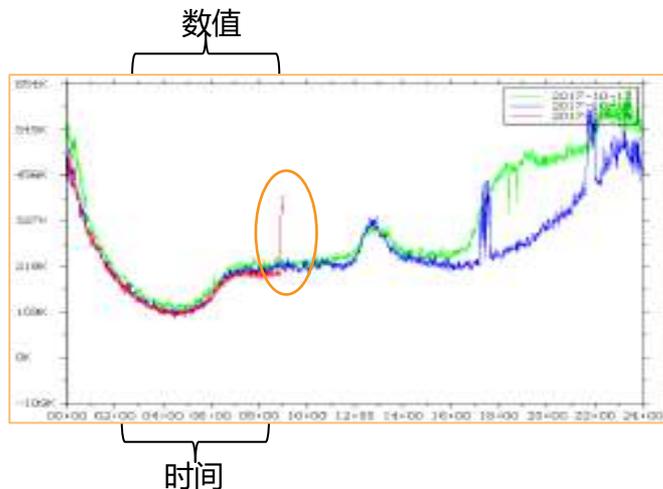
问题

1. 非正态分布（混合高斯分布）
2. 纵向的周期波动

算法过程

1. 取当前点**时间窗口**内的环比、周同比N个数值，总样本为M
2. 计算样本M的均值 (μ)、标准方差 (σ)、当前点的残差 ($|x_i - \mu| >$)
3. 当 $|x_i - \mu| > Z_c * \sigma$ ，判定为异常，否则判定为正常

第二步：无监督算法



孤立森林

优点

- 1 取消阈值告警，无监督，工作量小
- 2 避免样本非正态分布情况下，统计判别失效的场景

问题

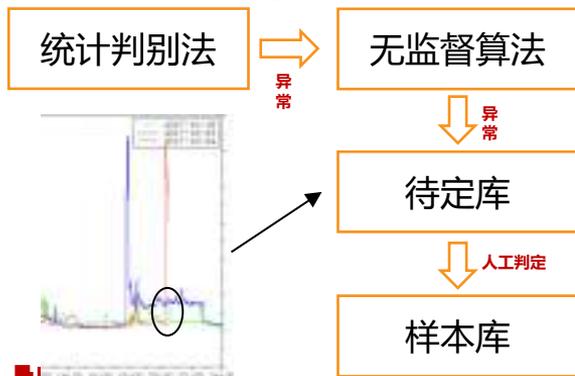
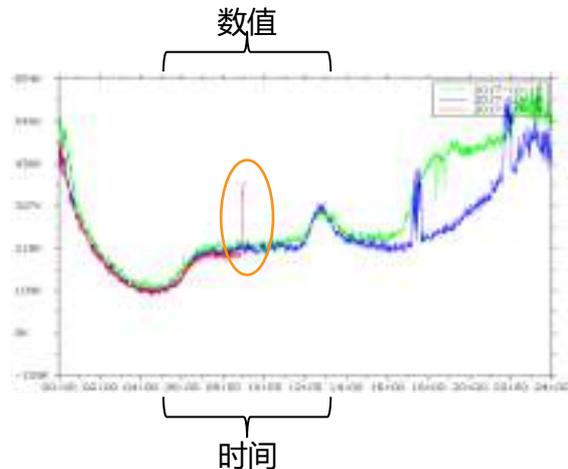
无校正学习能力

算法过程

1. 取当前点**时间窗口**内的环比、周同比N个数值特征与对应时间点，组成总样本M
2. 随机选取一个维度划分，之后对左右孩子继续重复执行直到所有孩子都为叶子
3. 选择树深度小于一定阈值的叶子结点标记为异常



第三步：有监督的特征与打标工程



特征过程

- 1.取当前点**时间窗口**内今天、昨天、周同比N个数值
- 2.Tsfresh计算N个数值的时序特征值（32个）

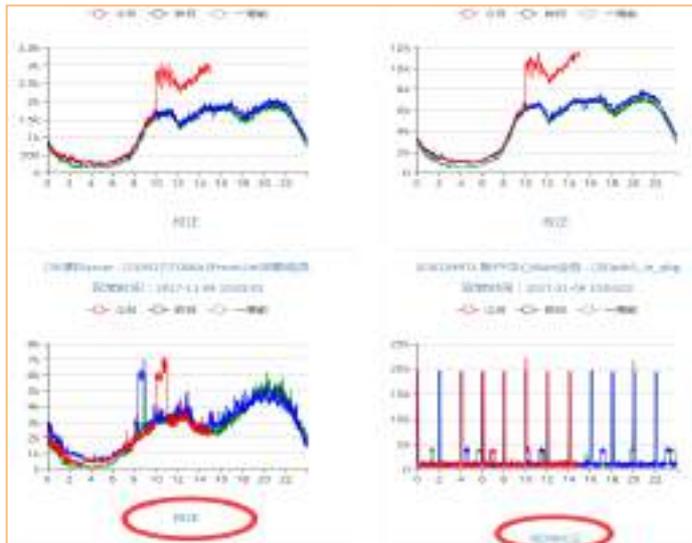
打标过程

- 1.统计判别法+无监督算法都判定异常输出到异常库
- 2.人工校正异常库的样本点，根据今天、昨天、周同比确认该样本：异常标记为0，正常标记为1

第三步：有监督的特征与打标工程



GOPS2017
Shanghai



alarm_id	att_id	log_time	end_time	sender	alarm_content	is_check
1337	2151646	2017-11-07 19:02:03	2017-11-07 19:18:03	0	0	1
1405	142637	2017-11-07 19:59:02	2017-11-07 20:04:02	0	0	1
15888	222152	2017-11-08 00:00:02	2017-11-08 10:03:02	0	0	1
6327	180161	2017-11-08 00:04:02	2017-11-08 00:11:02	0	0	1
6328	180161	2017-11-08 00:04:02	2017-11-08 00:12:02	0	0	1
6327	180161	2017-11-08 00:04:02	2017-11-08 00:12:02	0	0	1
6327	180162	2017-11-08 00:04:02	2017-11-08 00:11:02	0	0	1
6327	180162	2017-11-08 00:04:02	2017-11-08 00:12:02	0	0	1
6327	180162	2017-11-08 00:04:02	2017-11-08 00:12:02	0	0	1
6341	251212	2017-11-08 05:48:02	2017-11-08 05:53:02	0	0	1
6328	180162	2017-11-08 00:04:02	2017-11-08 00:12:02	0	0	1
16297	2189117	2017-11-08 00:03:02	2017-11-08 00:08:02	0	0	1
4392	20	2017-11-08 10:04:02	2017-11-08 10:09:02	0	0	1
4851	54839	2017-11-08 00:04:02	2017-11-08 00:09:02	0	0	1
11295	129854	2017-11-08 00:03:02	2017-11-08 00:08:02	0	0	1
15213	2189117	2017-11-08 00:03:02	2017-11-08 00:08:02	0	0	1
18018	825985	2017-11-08 00:03:02	2017-11-08 00:08:02	0	0	1
80541	10780	2017-10-25 11:04:02	2017-10-25 11:05:02	1	0	0
80541	10780	2017-10-25 11:04:02	2017-10-25 11:05:02	1	0	0
80541	10780	2017-10-25 11:04:02	2017-10-25 11:04:02	1	0	0
80541	10780	2017-10-25 11:04:02	2017-10-25 11:05:02	1	0	0
80541	10780	2017-10-25 11:04:02	2017-10-25 11:05:02	0	0	0
80541	10780	2017-10-25 11:04:02	2017-10-25 11:05:02	0	0	0

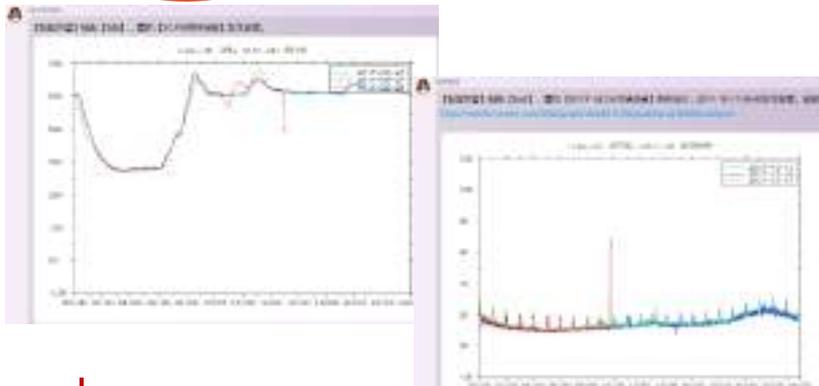
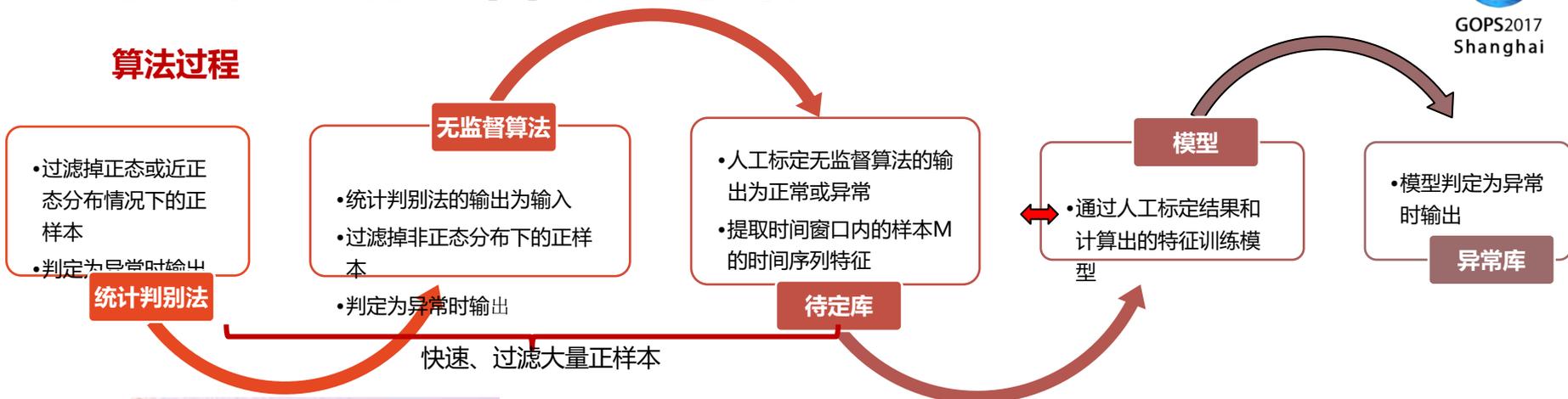
1. 输出异常视图到前端页面
2. 人工确认是否真的异常，假异常则校正
3. 后台根据人工校正的结果，存下校正后的所有结果：正常记为1，异常记为0



GOPS2017
Shanghai

第三步：有监督算法方案

算法过程



实际成果

1. 秒级监控
2. 统计判别+无监督算法：**100%**覆盖、**125万**次/分

多维下钻智能分析



GOPS2017
Shanghai

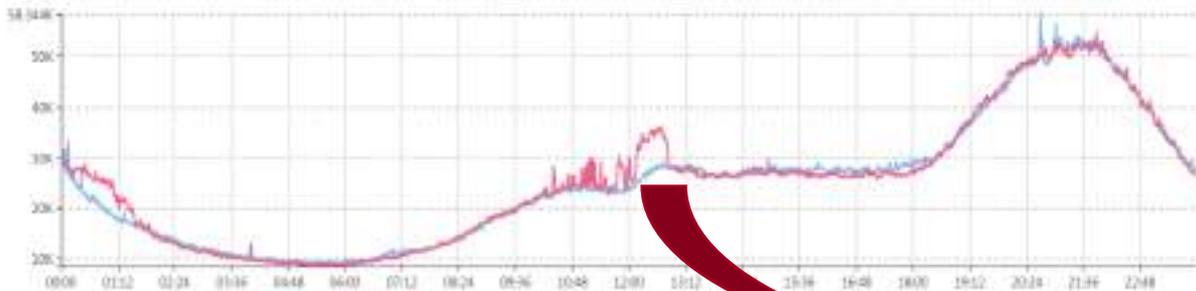
多维下钻分析的问题

例如：视频服务业务的KPI指标总卡顿次数统计维度由以下多个维度组成

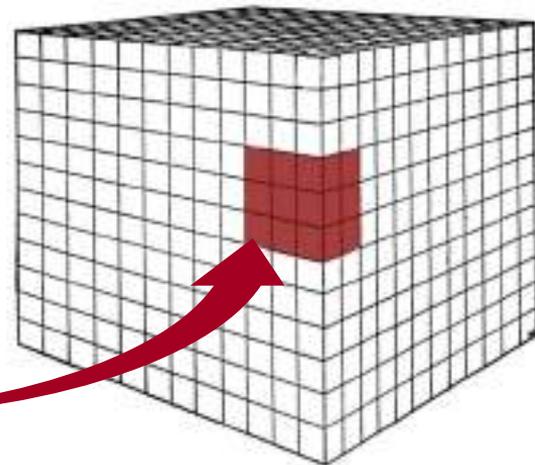


如何快速找到大量组合中最核心的影响因素？

异常KPI曲线



异常维度组合



多维下钻分析算法流程



GOPS2017
Shanghai

数据预处理

- 清理冗余维度、噪声数据
- 选取KPI正常时刻作为参考点

压缩搜索空间

- 移除与异常无关维度和特征
- 选择最有可能聚集的异常维度特征集

搜索异常维度组合

- 精确定位异常聚集维度组合



GOPS2017
Shanghai

多维分析方法

目标：搜索出对异常波动和贡献度最大的维度组合

差异度：（ Jensen-Shannon (JS) divergence ）（ 找出维度 ）

$$D_{JS}(P, Q) = 0.5 \left(\sum_i p_i \log \frac{2p_i}{p_i + q_i} + \sum_i q_i \log \frac{2q_i}{p_i + q_i} \right)$$

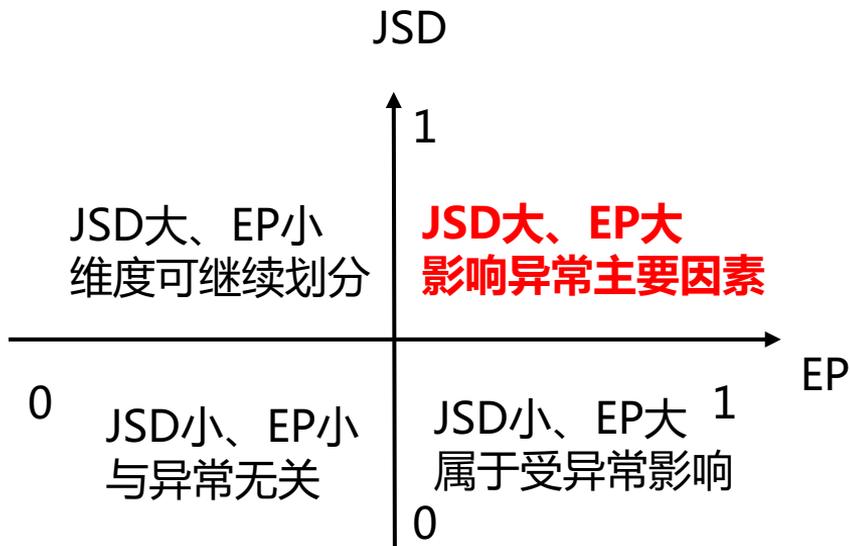
（ 预测值和观测值 ）

贡献度： $EP_{ij} = (A_{ij}(m) - F_{ij}(m)) / (A(m) - F(m))$ （ 找出子项 ）

搜索策略



GOPS2017
Shanghai



效果与应用场景



GOPS2017
Shanghai

异常分析时间缩短：

人工分析：30至60分钟

算法分析：10秒至30秒

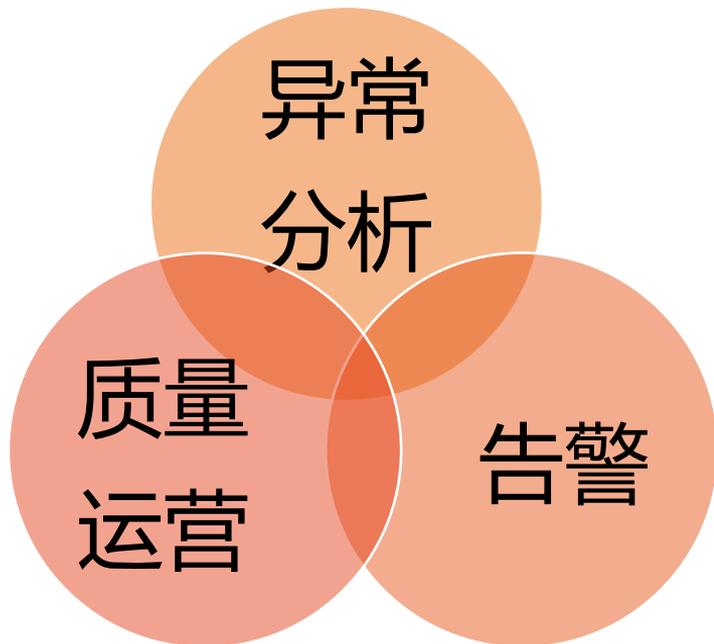
快速定位根因，避免故障扩大

准确率提升：

人工分析：容易错漏

算法分析：准确率近95%

精准定位故障根因，避免遗漏





GOPS2017
Shanghai

关联告警智能分析

Apriori算法机制简述

著名的超市里啤酒+尿布算法 **频繁项集算法**，目标是找到最大的K项频繁集

支持度就是几个关联的数据在数据集中出现的次数占**总数据集的比重**。
或者说几个数据关联出现的概率

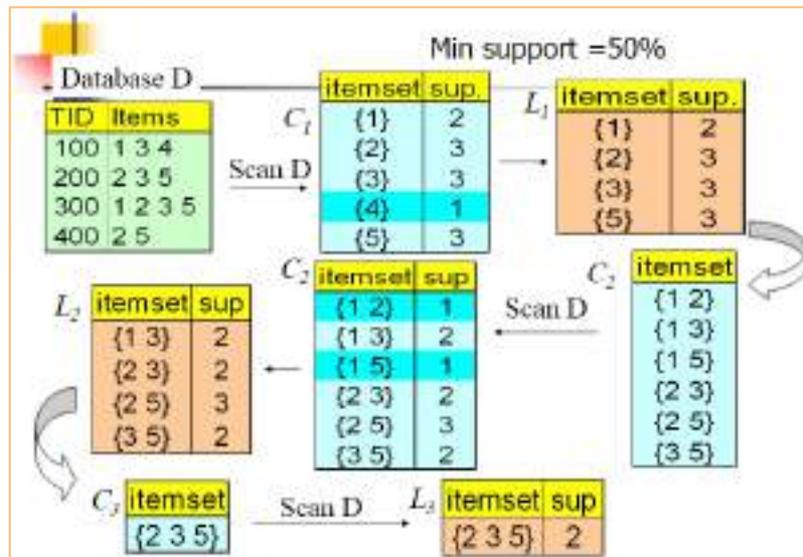
$$\text{Support}(X, Y) = P(XY) = \text{num}(XY) / \text{num}(\text{总样本})$$

$$\text{Support}(X, Y, Z) = P(XYZ) = \text{num}(XYZ) / \text{num}(\text{总样本})$$

置信度体现了一个数据出现后，**另一个数据**出现的概率

$$\text{Confidence}(X \leftarrow Y) = P(X|Y) = \frac{P(XY)}{P(Y)}$$

$$\text{Confidence}(X \leftarrow YZ) = P(X|YZ) = \frac{P(XYZ)}{P(YZ)}$$



图片来源: <https://www.evget.com/article/2017/4/12/26086.html>

Apriori 频繁项集



关联规则结果 举例说明



GOPS2017
Shanghai

tnm2`字符型业务特性告警`[CVM设备区][其他] ==> tnm2`Ping告警`[CVM设备区][其他]	preCount:8276	prePro:0.96	ItemSetCount:8183	itemSetPro:0.95	Conf:0.968762687289
tnm2`Ping告警`[CVM设备区][其他] ==> tnm2`字符型业务特性告警`[CVM设备区][其他]	preCount:8522	prePro:0.99	ItemSetCount:8183	itemSetPro:0.95	Conf:0.960220605492
tnm2`硬盘告警`[CVM设备区][其他] ==> tnm2`Ping告警`[CVM设备区][其他]	preCount:6461	prePro:0.75	ItemSetCount:6389	itemSetPro:0.74	Conf:0.968856214208

- 例1：日志变更`alarm_type`[w.qzone.qq.com][SET_D01][WEB] ==> tnm2`进程告警`[w.qzone.qq.com][SET_D01][WEB] , preCount:2535, prePro:0.29, ItemSetCount:1850, itemSetPro:0.21; Conf:0.73
- 例2：tnm2`字符型业务特性告警`[绑定逻辑][逻辑SPP] ==> tnm2`上报超时告警`[CVM设备区][其他] , preCount:4965, prePro:0.57, ItemSetCount:3269, itemSetPro:0.38; Conf:0.66
- 例3：告警流水`CoreDump监控`核心逻辑 ==> tnm2`Ping告警`[CVM设备区][其他] , preCount:2300, prePro:0.27, ItemSetCount:2290, itemSetPro:0.26; Conf:0.996

解释，如例3：

preCount/prePro为前者出现频次/频率，ItemSetCount/itemSetPro为两者出现频次/频率。

“告警流水`CoreDump监控`核心逻辑”（记为A）出现，“tnm2`Ping告警`[CVM设备区][其他]”（记为B）也随之在同一个时间窗口出现的概率为99.6%。即时间窗口出现A，同时包含B的条件概率为99.6%。其中，在0601~0630数据中，A出现次数2300次，AB共现次数2290次。

频繁项集关联告警智能分析

【告警智能分析】关联项告警详细列表

告警时间	源地址	告警内容	告警关联项名称	关联项发生时间	关联项发生地	关联项种类
2017-11-01 00:00:00	828	204908112*204908112*204908112		2017-03-11 00:22:00	%	Y
	929	204908112*204908112*204908112		2017-03-11 00:22:00	%	Y
	1281	21180000*21180000*21180000		2017-03-11 00:42:00	%	Y
		21180000*21180000*21180000		2017-03-11 00:42:00	%	Y
		21180000*21180000*21180000		2017-03-11 00:42:00	%	Y
	1484	21180000*21180000*21180000		2017-03-11 00:52:00	%	Y
		21180000*21180000*21180000		2017-03-11 00:52:00	%	Y
	1784	21180000*21180000*21180000		2017-03-11 00:42:00	%	Y
	1871	204908112*204908112*204908112		2017-03-11 00:01:00	%	Y
		204908112*204908112*204908112		2017-03-11 00:01:00	%	Y
	2085	21180000*21180000*21180000		2017-03-11 00:42:00	%	Y
		21180000*21180000*21180000		2017-03-11 00:42:00	%	Y
	2381	21180000*21180000*21180000		2017-03-11 00:52:00	%	Y
		21180000*21180000*21180000		2017-03-11 00:52:00	%	Y
	2584	204908112*204908112*204908112		2017-03-11 00:01:00	%	Y
		204908112*204908112*204908112		2017-03-11 00:01:00	%	Y
	2784	21180000*21180000*21180000		2017-03-11 00:52:00	%	Y

频繁项集算法有更多的应用场景.....





GOPS2017
Shanghai

智能运维客服



GOPS2017
Shanghai

根据运维客服需求特点定制的三种模式

智能模式

- 默认模式
- 自动匹配业务FAQ库
- 置信度高的匹配直接给出答案，置信度不高则给出相关度Top5的问题链接

人工模式

- 需要用户手动点击链接进入
- 不活跃时间超过两小时的人工会话，会自动切换回智能模式

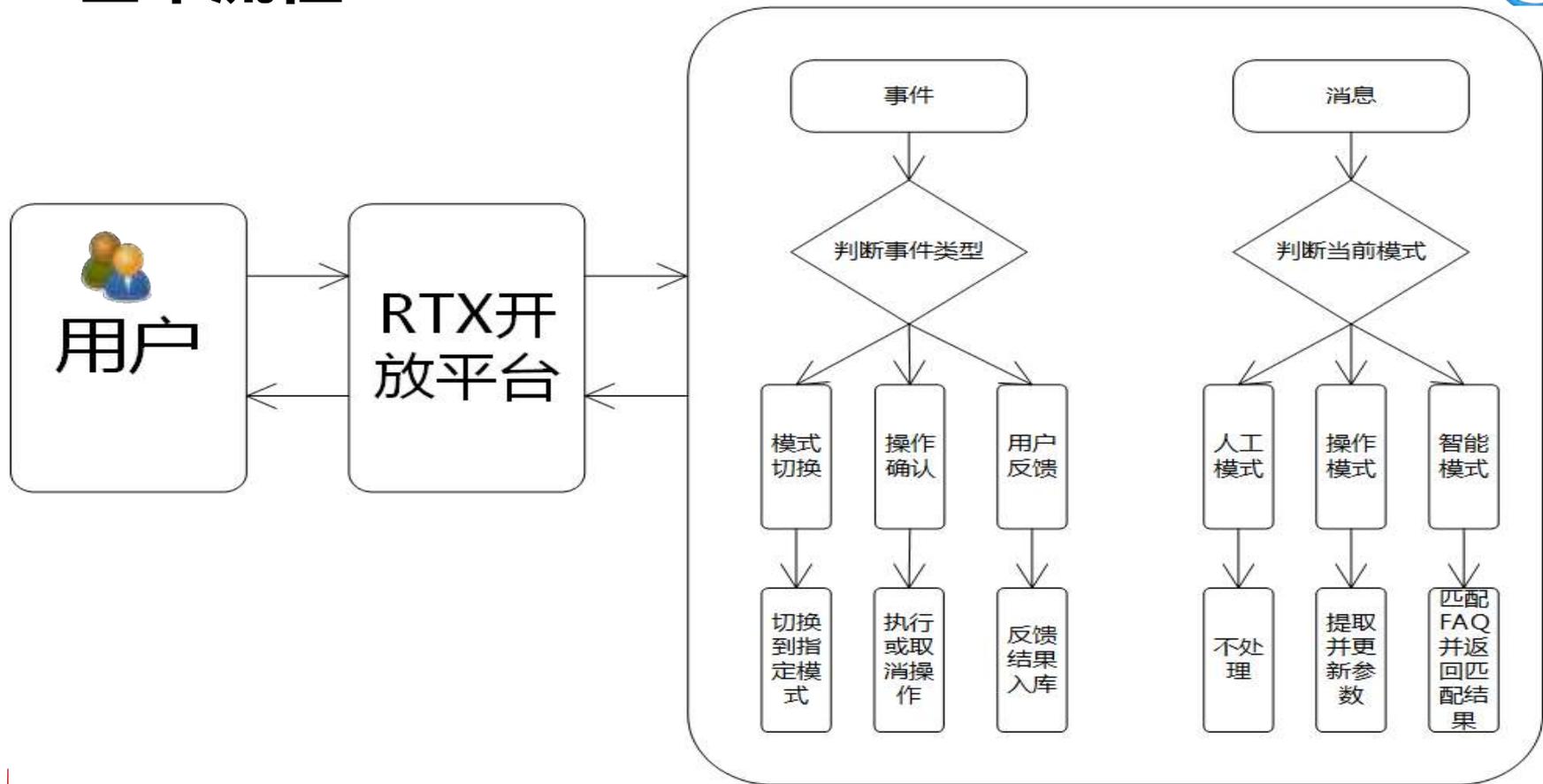
操作模式

- 默认不开启
- 只适用于参数简单的操作
- 有后台鉴权

基本流程



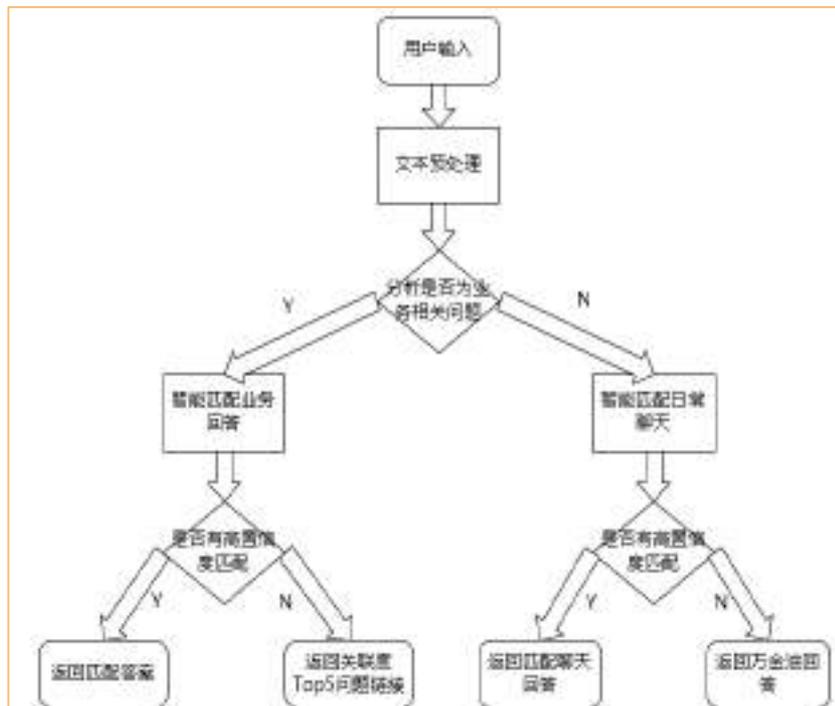
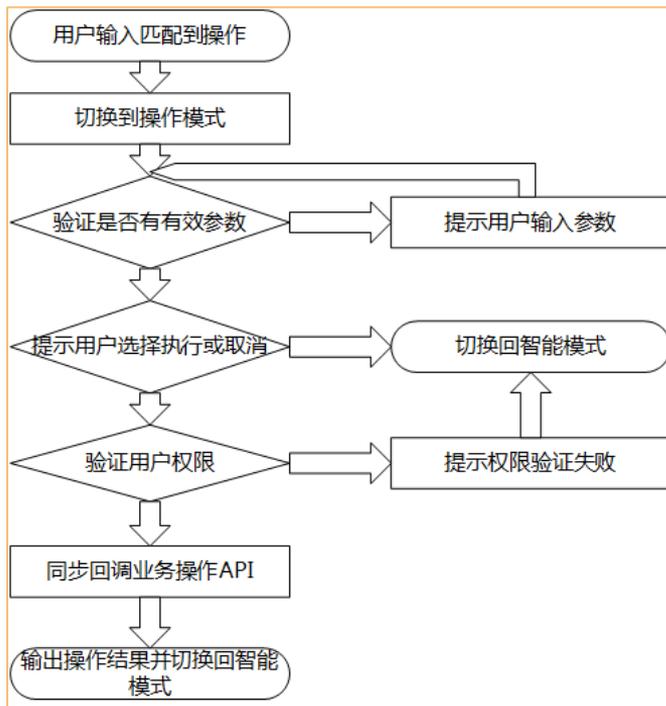
2017
ghai



操作模式和FAQ流程图



GOPS2017
hanghai



智能模式演示



GOPS2017
Shanghai



Demo1:
完全识别,直接给出答案



Demo2:
部分识别,给出相关问题

操作模式演示



GOPS2017
Shanghai





GOPS2017
Shanghai

天王星文本投诉智能检测系统

天王星文本投诉智能检测系统

数据简介：用户舆论反馈的自然语言

数据特点

量级较大（约5-10W条/天）

主题明确

短文本

反馈量及主题与故障事件相关性强

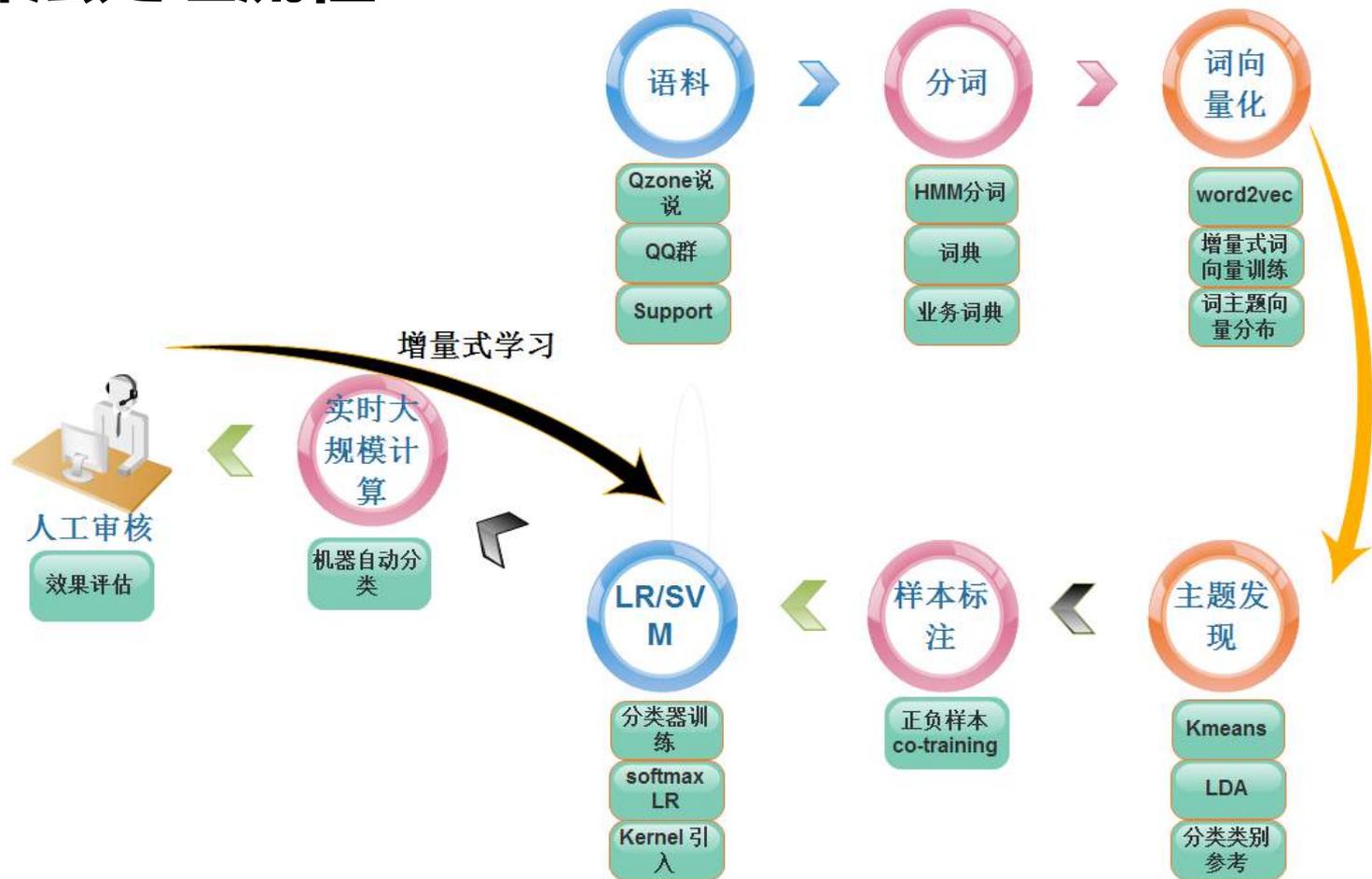
文本反映的信息量全面

思路

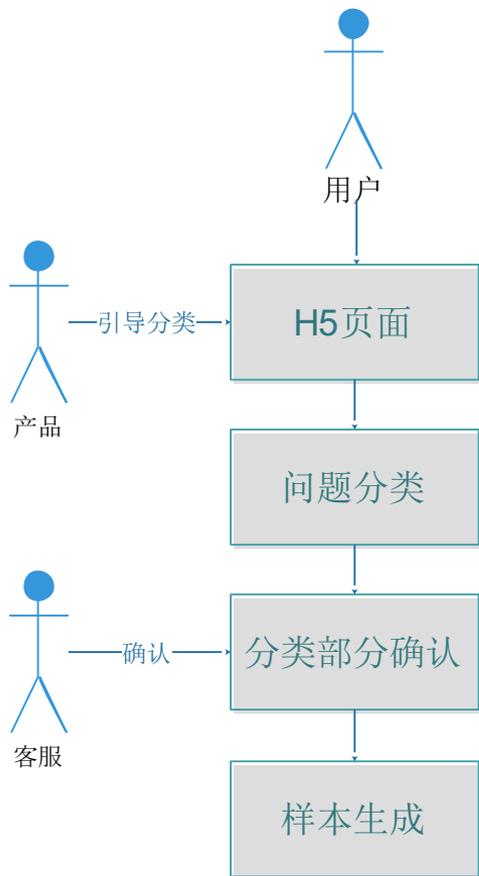
监控告警：利用自然语言文本挖掘技术



算法处理流程



产品策略



应用实践

内容	版本	自动分类*
现在登录QQ空间上应用时（比如QQ农场牧场），为什么要密码了，以前不用的。这到底什么意思？	V1_AND_Q Z_5.8.1_288_YB_D	玩吧问题 wbwt
情侣空间为什么没有呢？	V1_AND_Q Z_5.8.1_288_YB_D	情侣空间 qlkj
双击扩大屏幕能不能关掉？	V1_AND_Q Z_5.8.1_288_YB_D	建议需求 jyxw
空间说说设置部分人禁止评论！	V1_AND_Q Z_5.8.1_288_XM_A	体验反馈：说说
手机QQ空间无法进入农牧场	V1_AND_Q Z_5.8.1_288_OPP O_A	玩吧问题 wbwt
应用呢？怎么玩游戏了，不想用动画，	V1_AND_Q Z_5.8.1_288_YT D_D	玩吧问题 wbwt
被人举报 无法回复	V1_AND_Q Z_5.8.1_288_YB_D	体验反馈：禁言 jy
访空间，有好文章，原可收藏，现访空间，没有收藏提示！	V1_AND_Q Z_5.8.1_288_YB_D	分享收藏 fxsc

数据效果

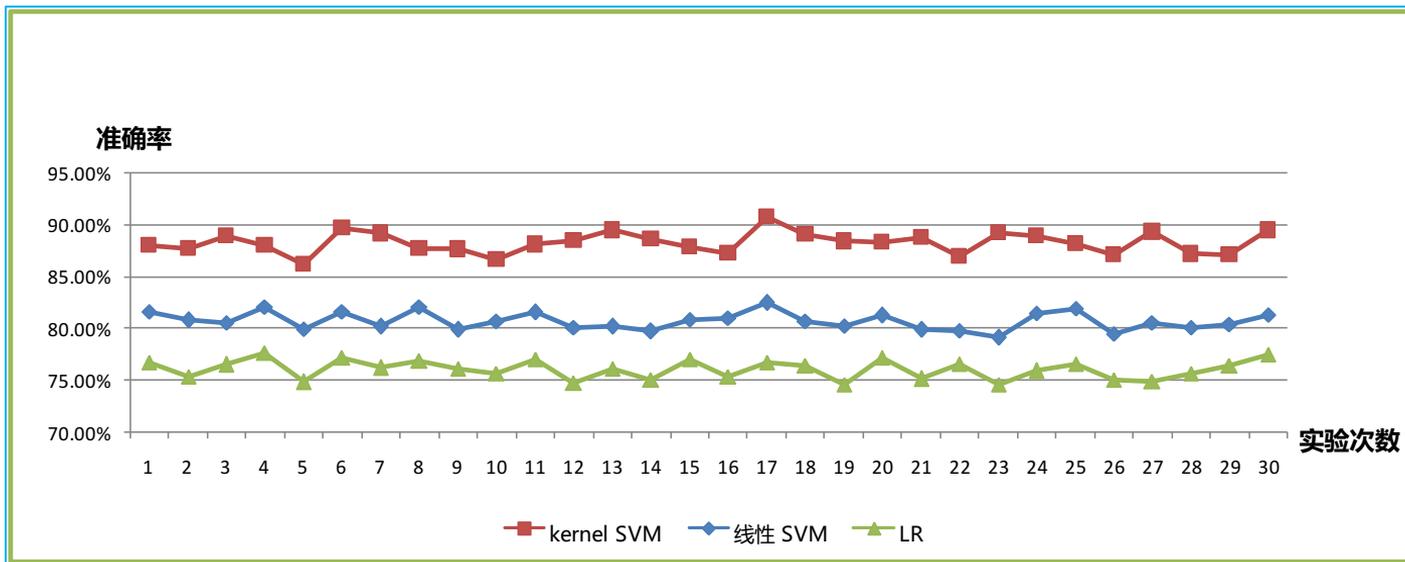
分类准确率：

平均75~90%

手机QQ空间独立版（约80+个类别）：90~95%

手机QQ结合版：85~90%

客服至少减少**75%**的工作量





标准 简单 智能 深耕运维

专注自动化运维，永不止步

织云已为设备数万规模的业务提供高效运维支持



了解、咨询、试用请扫描



腾讯云全球基础设施布局



增速领先行业，全球5大洲 31个可用区

中国带宽储备最大的CDN服务提供商 70T+CDN总带宽 800+全球节点

腾讯云服务能力图



GOPS2017
Shanghai

云应用API接口



腾讯云



数据库

关系型数据库CDB

文档数据库

数据传输服务

分布式云数据库



存储与 CDN

对象存储COS

私有云存储

存储网关

云硬盘



安全与大数 据/AI

大禹网络安全

天御业务安全防护

大数据处理套件

数据工坊

数字营销

大数据可视交互系统

智能分析

DI-P实施多维分析引擎

机器翻译

DI-x深度学习平台证



视频服务

移动直播

互动直播

短视频

美视优享频



游戏服务

游戏更新

全球应用加速

智营网优

游戏存储



计算&网络

私有网络

NAT网
关

跨地域互
联

弹性网卡

云服务器

GPU云
服务器

FPGA云
服务器

黑石



大数据与AI

万象优图

文智自然语言处
理

智能语音服务



管理运维

命令行工具

域名管理

蓝鲸 / 织云平台

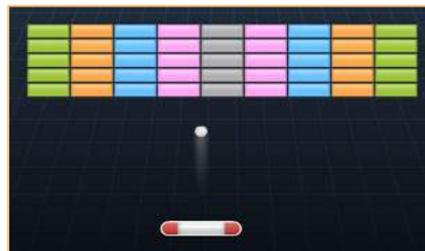
娱乐
教育
医疗
政府
金融
零售





GOPS2017
Shanghai

理想的AI全自动化运维



状态s



封闭问题空间
VS
开放问题空间

经验->规则->规则+ (算法->统计判别->机器学习
(无监督、有监督、深度学习、强化学习)) ->人工智能

我们该如何定义
AI全自动化运维的损失函数？

结束语



GOPS2017
Shanghai

理想的AI运维就像无人驾驶一样

似乎离我们还很远

但通过不断的进步

我们坚信它一定会到来



GOPS2017
Shanghai



Thanks

高效运维社区
开放运维联盟

荣誉出品



GOPS2017
Shanghai



想第一时间看到
高效运维社区公众号
的好文章吗？

请打开高效运维社区公众号，点击右上角小人，如右侧所示设置就好

