

# 客户服务多次来电预警管理研究

门 萍

(国网客服中心南方分中心,江苏 南京 211000)

**摘要:**我们基于95598海量工单大数据,对重复来电行为预测作为分类问题,进行客户服务多次来电预警管理研究,最终目标是为了提取用户行为特征,建立重复来电行为预警模型,用于识别和筛选特定类型的目标用户,提高用户满意度,提高95598来电接听效率。

**关键词:**客户服务;多次来电;管理

**【DOI】**10.12231/j.issn.1000-8772.2021.27.159

## 1 研究背景

城市经济的高速发展,对电力的需求在达到峰值之前会越来越,由此产生的数据也会呈现几何级数爆发式增长。大数据时代定义——《大数据时代》说明数据价值[1]。在数据作为重要生产资料的当下,如何利用好数据,用数据分析结果去驱动业务,大数据三化引用。

2020年9月以来,习近平总书记就全球气候变化发表一系列指导意见,明确我国实现“碳达峰、碳中和”的具体目标;2021年全国两会更被首次写入政府工作报告;2021年3月1日,国家电网有限公司发布《国家电网公司发布“碳达峰、碳中和”行动方案》。

传统电力企业为实现碳达峰、碳中和目标,电力行业的管理、运营、服务、交易模式正发生巨大变化,各大电力企业均展开数字化战略部署,长期、持续投入数字化转型建设,从组织架构、资源投入、商业模式等多方面进行变革与探索。根据数字化转型成熟度模型,电力企业的数字化转型被分为单点试验、局部推广、扩展复制、运营管理、优化创新五个阶段<sup>[2]</sup>。

与此同时,聚焦到95598供电服务过程中,服务风险管控是日常业务运营管控中的难点问题,客户轨迹分析也较为复杂,给服务风险事件的筛查带来了非常大的工作量,导致该项工作无法有效开展,急需依赖大数据平台形成自动化工具。根据客户多次来电轨迹进行服务风险识别,形成高、中、低三级风险预警,对于同一客户下派量及催办量大的问题,形成服务预警图,协同进行该风险的消除。可以看到每个客户的风险等级及服务问题,该数据每日数据进行更新。主要应用在于:(1)减轻多次来电管控压力,及时进行风险控制,降低服务压力。(2)根据预警结果,及时干预,降低服务升级风险及一线人员处理压力。(3)根据预警结果,定期总结重要服务事项报备,最终答复、特殊客户等疑难业务处理方面的难点问题,协同优化处置方法或升级沟通,降低业务处理难度。(4)挖掘某类问题长期会引发多次来电的情况,进行双方深度沟通,提升业务处置效率和质量,提升客户满意度。

## 2 客户行为分析

### 2.1 基于行为数据的投诉倾向预测

用户行为是用户在具体时间、地点,基于某个事件产生的交互行为。从某种程度上,一个完整的事件等价于对用户行为的一个定义。通过事件将用户行为连贯观察并进行深入分析,即为用户行为分析<sup>[3]</sup>。同时,用户行为受到外在环境和内在心里判断的共同作用,会形成如图1的反馈机制。

分析用户来电行为,设计出符合用户需求的功能,推进产品的迭代;另一方面,挖掘出行为数据背后的规律,为用户提供更为专业化、个性化的服

务,提升用户体验,促进产品生态的良性循环。

## 2.2 数据来源及预处理

### 2.2.1 数据准备

(1)剔除重复的简单诉求、异常来电;(2)剔除客户催办、补充信息、撤销业务;(3)剔除12345、110来电号码;(4)剔除要求“保密”的客户。

### 2.3 规则设计

#### 2.3.1 高风险客户

通过对来电次数、业务类型、问题描述、回访结果等关键信息进行提取、分析,研判出后期将大概率出现工单升级、投诉属实等情况的来电,对于高风险预警,应安排专人进行分析和跟踪,加快处理速度、提升优质服务水平,防止出现事件升级。

(1)超期风险客户规则。a.识别目的:主要用于对服务处理部门进行预警,对快到时限的在途工单和时限即将过半的在途工单进行管控,及时干预,安抚客户情绪。(2)情绪升级客户规则。a.识别目的:通过媒体等表达投诉意愿并在1天内再次来电的进行预警,用于防范事件未处理完全。b.识别目的:通过每日报修在途工单的预警,主要用于防止客户停电问题未彻底解决。c.识别目的:海量语音诉求识别系统中,多次来电服务风险识别的情绪升级服务风险客户。(3)疑难业务客户规则。识别目的:主要用于复杂业务,容易造成客户多次来电,提示尽快解决。(4)处理不满客户规则。识别目的:主要用于定位出对处理结果不认可的客户。

#### 2.3.2 中风险客户

中风险预警主要用于识别客户有投诉升级表述的来电或在来电中提及现场服务问题、现场服务规范等有一定升级可能的来电,及时发现服务风险和隐患。

(1)情绪升级客户。主要用于通过其他途径等表达投诉意愿的关键词预警,客户未在1天内再次致电,但仍存在隐患。(2)风险高分客户。主要用于通过风险系数评分7以上的事件进行预警,出现多次来电情况,及时提示处理部门。(3)线下诉求客户。主要用于定位服务诉求超出已有服务流程,需发起临时线下流程客户。

#### 2.3.3 低风险客户

(1)疑似非客户本人。主要用于按照业务规则,筛选疑似套取用户信息。(2)报备特殊客户。主要用于对已有报备影响服务事项的特殊事件进行筛选。

## 3 算法应用

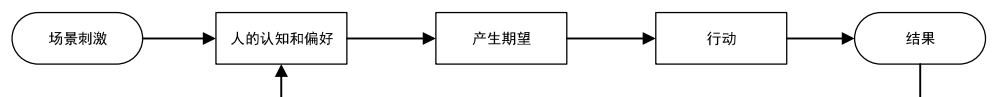


图 1

### 3.1 Logistic 模型

Logistic 回归模型常用于描述分类响应变量与解释变量之间的关系,属于广义线性回归模型中的一类,具体可分为二分类的回归模型和多分类的回归模型。其解释变量可以是连续型变量和离散型变量,如属性数据,计数数据等,响应变量则是两个或两个以上的离散型变量<sup>9</sup>。建模程序如下:(1)构建预测函数,记为  $h$  函数,该函数就是我们需要找的分类函数,它用来预测输入数据的判断结果。(2)构造一个 Cost 函数(损失函数),该函数表示预测的输出( $h$ )与训练数据类别( $y$ )之间的偏差,利用最大似然估计函数<sup>9</sup>。

考虑所有训练数据的“损失”,将 Cost 求和或者求平均,记为  $J(\theta)$  函数,表示所有训练数据预测值与实际类别的偏差。(3)  $J(\theta)$  函数的值越小表示预测函数越准确(即  $h$  函数越准确),所以这一步需要做的是找到  $J(\theta)$  函数的最小值。通常我们使用的是梯度上升法/梯度下降法。

### 3.2 GBDT 模型

DT-Decision Tree 决策树,GB 是 Gradient Boosting,是一种学习策略,GBDT 的含义就是用 Gradient Boosting 的策略训练出来的 DT 模型<sup>9</sup>。GBDT 是为了解决一般损失函数的优化问题,方法是用损失函数的负梯度在当前模型的值来模拟回归问题中残差的近似值。基于 boosting 算法的思想,每次迭代不断提高预测的准确性。由于 GBDT 能够发现多种有区分性的特征以及特征组合,决策树的路径可以直接作为其他模型的输入特征使用,省去了人工寻找特征、特征组合的步骤。

GBDT 是决策树的组合模型,使用 GBDT 构造组合特征是指将 GBDT 中所有决策树每一个叶子节点为一个新的特征,因此构造得到的特征数目与 GBDT 叶子节点的数目相同,每一个特征取值为 0 或者 1。对于每一棵决策树,若输入的样本落入到某个叶子节点,则该叶子节点的取值为 1,否则为 0。模型的结果是一组回归分类树组合(CART Tree Ensemble): $\dots, \dots$ 。其中学习的是之前  $j-1$  棵树预测结果的残差。

### 3.3 GBDT 与 Logistic 结合模型

Logistic 回归在使用中有相当大的缺陷,模型本身容易欠拟合,会造成准确度不太高;不能很好地处理大量多类特征或变量;Logistic 只能处理两分类问题,且必须线性可分,对于非线性特征,需要进行转换。GBDT 模型单独使用也会以下的缺陷,由于弱学习器之间存在依赖关系,难以并行训练数据。如果数据维度较高时会加大算法的计算复杂度<sup>9</sup>。

我们通过 GBDT-LR 结合来最大可能规避以上的风险,其核心思想利用 GBDT 生成特征在利用 LR 对特征值进行拟合。LR 模型十分适合并行化,因此对于大数据的训练十分有效。但是对于线性模型而言,学习能力是有限的,因此需要大量的特征工程预先分析出有效的特征或者是特征组合,从而去间接的增强 LR 的非线性学习能力。特征组合,是通过特征的一些线性叠加或者非线性叠加得到一个新的特征,可以有效的提高分类效果,常见的特征组合方式有笛卡尔积方式。如图 2,GBDT 是梯度提升决策树,由多棵树组成。构造一个决策树,根据已有的模型和实际样本输出的残差上再构造一颗决策树,不断地进行迭代。每一次迭代都会产生一个增益较大的分类特征,因此 GBDT 树有多少个叶子节点,得到的特征

表 1

风险级别	客户性质	详细类别	客户数	正确数	预警准确率
高风险	超期风险客户	4 个工作日(最短工单处理时限)后工单在途且出现客户催办 1 次及以上	8	8	100.00%
	情绪升级客户	命中关键词,1 天内来电 3 次及以上	60	57	95.00%
		停电或复电事件后,1 天内出现工单升级或 1 天后仍在途且有催办	159	103	64.78%
		内部投诉来电,30 天内来电 2 次及以上	39	30	76.92%
	投诉来电,30 天内来电 2 次及以上	7	6	85.71%	
	疑难业务客户	疑难业务事件后来电 3 次及以上	25	19	76.00%
中风险	处理不满客户	出现回访不满意,并且出现【回访派单】1 次及以上	30	30	100.00%
	无报备特殊客户	30 天内,来电 15 次及以上,并没有重要服务事项、最终答复、特殊先生台账	8	7	87.50%
	情绪升级客户	命中关键词,1 周内来电 3 次及以上	33	29	87.88%
	风险高分客户	风险评分处于[7,10],1 天内来电 4 次及以上	26	19	73.08%
低风险	线下诉求客户	超出现有服务流程的其它服务诉求客户	5	5	100.00%
	疑似非户主本人客户	查询户号 5 次及以上	472	-	-
		查得不同户号 3 次及以上	550	-	-
	报备特殊客户	30 天内,来电 15 次及以上,并有重要服务事项	4	-	-
总体预警准确率(剔除疑似非户主本人)			392	305	77.81%
高风险预警准确率			328	252	76.83%

空间就有多大,并将该特征作为 LR 模型的输入。

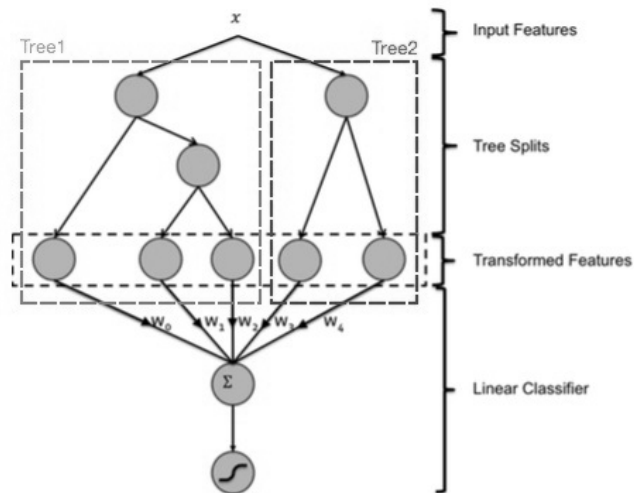


图 2

## 4 模型设计及结果验证

选取 2021 年某月期间某公司的数据进行验证。

周期内,某公司来电数 20.63 万个,共产生工单数 28.93 万件;其中 3 次及以上来电号码 14546 个。

应用本产品检出各类风险 1418 件,具体如表 1。

### 参考文献

[1] 新华网. 电力企业面向未来数字化转型五大关键因素不可或缺 [EB/OL]. [http://www.xinhuanet.com/info/2020-09/25/c\\_139395676.htm](http://www.xinhuanet.com/info/2020-09/25/c_139395676.htm).  
 [2] 曾浩峰. 社交网络用户行为的分析和预测[J]. 科学经济导刊, 2017(06):198.  
 [3] 周志华. 机器学习[M]. 北京:清华大学出版社, 2016.  
 [4] 王治. Logistic 回归系数极大似然估计的计算[J]. 数学理论与应用, 2009, 29(4):87-90.  
 [5] 李航. 统计学习方法[M]. 北京:清华大学出版社, 2012.  
 [6] Xinran He, Junfeng Pan, Ou Jin, Tianbing Xu, Bo Liu. Practical Lessons from Predicting Clicks on Ads at Facebook[M]. ACM, 2014.  
 [7] Lunardon N, Menardi G, Torelli N. ROSE: A package for binary imbalanced learning[J]. The R Journal, 2014, 6(1):82-92.

作者简介:门萍(1976,11-),江苏南京人,工作单位:国家电网公司客户服务中心南方分中心,职称:副高级工程师,学历:本科,研究方向:呼叫中心客户服务。