

运营商大数据 标签体系加速行业数字化转型实践

Accelerating Industry Digital Transformation Practices with Labeling System Based on Telecom Operator Big Data

李大中,张倬荣,闫 龙,王功举,宋雨伦(联通数字科技有限公司,北京 100010)

Li Dazhong,Zhang Zhuorong,Yan Long,Wang Gongju,Song Yulun(Unicom Digital Technology Co.,Ltd.,Beijing 100010,China)

摘要:

“数据二十条”的发布加速了政务运行、经济活动、社会服务与民生等领域拓展数据应用的步伐,将推动更多领域实现数据驱动与数字化转型。运营商数据具备超大规模、全面连续、时效性强、真实度高、关联价值显著等特征。运营商大数据标签体系可以拓展数据应用场景,提高数据资源利用率,推进数据要素市场化配置,实现数据价值流通。结合中国联通数据市场化服务实践,设计并提出了一套集约化的跨行业、分场景的标签体系。

关键词:

电信运营商;标签体系;大数据

doi:10.12045/j.issn.1007-3043.2023.06.002

文章编号:1007-3043(2023)06-0008-07

中图分类号:TP391

文献标识码:A

开放科学(资源服务)标识码(OSID):



Abstract:

The release of the "Twenty Data Measures" has accelerated the pace of data application expansion in government operations, economic activities, social services and people's livelihood, which will drive more fields to achieve data-driven and digital transformation. Telecom operators' big data has the characteristics of ultra-large scale, all-round continuity, strong timeliness, high authenticity and significant associated value. The telecom operator's big data labeling system can expand data application scenarios, improve data resource utilization, promote the market-based allocation of data elements, and realize the circulation of data value. Combining China Unicom's data marketization service practice, it designs and proposes a set of cross-industry, intensive labeling system under different scenarios.

Keywords:

Telecom operator; Labeling system; Big data

引用格式:李大中,张倬荣,闫龙,等. 运营商大数据标签体系加速行业数字化转型实践[J]. 邮电设计技术,2023(6):8-14.

1 概述

标签体系,即用户信息标签化,是真实用户的虚拟代表,是根据用户的社会属性、生活习惯、消费行为、偏好特征等数据,挖掘出的有价值的标签化用户模型^[1]。

标签体系的应用领域包含图书馆、医疗、金融等传统行业以及社交网络、电商平台等互联网行业^[2]。从微观个人层面到宏观产业层面,标签体系在推动数字化转型和实现数据价值方面都起着重要作用。在

微观层面,标签体系可精准定位和识别用户,满足用户个性化需求,帮助企业提高用户体验,推动产品和服务升级^[3]。在中观层面,标签体系有助于公司企业和社会组织的运营优化,具体体现在产品功能优化、精准营销和个性化推荐3个方面^[4]。从宏观层面看,标签体系有助于构建公共数据资源,实现跨行业与跨领域的数据连接与交流,从而推动数字经济发展,实现产业协同创新,释放数据红利。

运营商作为各种信息的传输枢纽,位于信息传递与分发的核心地位,掌握了丰富的数据资源^[5]。运营商数据资源包含用户通信行为、设备信息、位置信息、互联网浏览、社交偏好等数据维度,具备超大规模、全

收稿日期:2023-04-28

面连续、时效性强、真实度高、关联价值显著等特征。运营商丰富的数据资源为构建高质量标签提供了基础保障。运营商标签为存量用户经营中的用户留存、用户增值、生命周期管理等场景提供了精细化、差异化和个性化的能力。在市场化服务过程中,运营商标签在智能推荐、内容管理、知识图谱构建、网络舆情监测、产品设计优化等场景也均有广泛应用。基于行业场景需求与运营商数据资源,通过构建统一加工、服务、运营的标签体系,实现对内数据治理与对外高效服务,加速企业内外部数字化转型。

本文完整展现了标签体系从规划设计到实施落地的过程,包括标签体系业务架构设计、技术实现流程以及行业应用案例3个模块。

2 运营商标签体系业务架构设计

根据国内外研究综述,标签体系主要分为用户行为、兴趣偏好、文本主题、人格特性和情绪等主题。基于运营商数据特性,本文在业务理解和技术创新的基础上,采用事实抽取、统计分析、学习类算法等方式对多源合规数据进行标准化、集约化、流程化的设计与研发,形成了一套跨行业、分场景的标签体系,包括通用标签和金融、汽车、文旅3套行业标签。通用标签包含了人口属性、通信属性、区域洞察、兴趣偏好等维度;汽车领域能力覆盖留资、线索评级、到店、成交、保客增换购与引流等方面;金融领域能力包括贷前风控、贷中评估、贷后管理等方面;文旅领域能力涵盖旅游管理、旅游服务和旅游推荐等方面。

2.1 通用标签

人口属性包含基础属性、地域属性、人生阶段、经济能力、社交关系、身份职业六大二级分类。这些分类跨领域通用性较强,可在个性化推荐、广告投放、用户画像与市场研究等场景提供基础数据支撑。

深入分析用户的通信行为数据至关重要,这能够帮助运营商发现潜在的风险和欺诈,优化运营业务,提高用户满意度。通信属性类别涵盖通信行为模式、基础信息、终端设备信息和基于用户通信行为的风险评估标签与社交属性标签。这些标签主要用于信用评级、风险评分、黑灰产识别与欺诈防范等场景。

区域洞察一级分类下涵盖旅游、交通、兴趣和通勤4个二级分类,从不同角度洞察区域内人口的空间活动规律。具体而言,旅游类标签用于洞察区域内用户的旅游行为模式,统计旅游景点与资源,为旅游管

理、服务和推荐提供参考。交通类标签用于洞察区域内的交通设施,提供交通运输规划与出行路线推荐等服务。兴趣类标签用于洞察区域内的休闲消费等场所,为商业推荐与营销决策提供支持。通勤类标签用于分析区域人口的流动模式,为城市交通规划提供数据支持。

构建兴趣偏好标签需要大量的行业调研和数据探索分析。为满足不同应用场景,本文区分了长期兴趣和短期兴趣2类标签。长期兴趣稳定且持久,如音乐、阅读与影音等;短期兴趣常见于消费场景。

2.2 行业标签

在通用标签的基础上,结合汽车、金融和文旅行业知识,逐步衍生出了3个行业标签体系。

数字化是汽车行业服务提质增效的重要抓手,以客户体验为中心的全渠道、连续性、定制化服务同样需要数字化来实现。鉴于汽车主机厂难以全面构建存量客户画像,也难以多层次实时感知客户动态,本文提供了一套汽车行业标签体系。该类标签有助于主机厂完善客户画像,建立动态客户价值评估框架,增强对用户的认知和精细化运营能力,包括但不限于私域维系、流失预警和客户激活等服务范畴。

金融行业标签中用户财富属性标签,将有助于金融机构更好地识别和管理风险,例如客户的风险偏好、借款能力和信用状况。同时,金融机构可以结合工作生活属性标签(例如商旅人士、出行偏好、外卖达人等),推出定向权益活动,以激活沉默存量用户。在安全合规的前提下,金融机构可以综合使用内外部标签,全面完善用户画像,以便为存量用户提供更优质的服务。

文旅行业标签有助于企业深入理解用户需求和旅游偏好,以提供个性化的旅游服务和推荐。例如,OTA企业可以根据用户标签推荐匹配的景点与酒店,实现差异化定制服务。景区也可以基于不同类型的游客标签,提供多种主题的游记、旅游攻略与视频内容,提高用户体验。

3 运营商标签体系技术实现流程

基于《标签类目体系:面向业务的数据资产设计方法论》中关于标签体系构建的方法论与《用户画像:方法论与工程化解决方案》中关于标签体系工程化的解决方案,本文提出了一套运营商标签体系的实施路径,包含标签设计、开发与使用3个环节。

3.1 标签设计

标签设计作为标签体系构建的起点,其合理性与完备性决定了后续开发和应用的效率与质量。标签设计阶段包含标签类目设计、元标签设计和标签存储设计3个环节。

3.1.1 标签类目设计

标签体系设计应根植于业务需求,同时考量数据采集的可行性与商业应用的价值。随着标签种类的增加,构建一套清晰且可扩展的标签分类框架尤为必要。标签类目体系应易于理解,例如以用户为对象的标签体系类目设计,可以从用户静态属性入手,如性别、年龄与常住地等,再拓展至动态行为,如通信记录、用户兴趣与偏好,如图1所示。

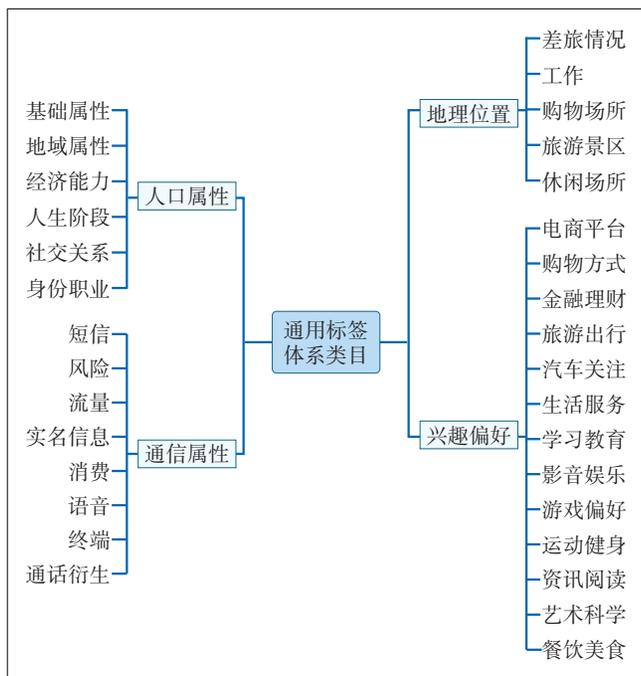


图1 运营商通用标签体系类目设计

然而,标签类目体系的设计完全依赖于业务场景,并不存在一套通用的体系模板。在运营商环境下,除用户维度外,标签体系还可以根据区域、设备和产品等维度开展设计。此外,标签类目设计还需要考虑数据采集的完备性与标签应用的成熟度,逐步拓展标签类型与层级。

3.1.2 元标签设计

元标签是对标签的进一步描述,其信息的准确全面直接决定了业务人员是否能正确理解与运用标签。主要的元标签有标签描述、标签名称、标签加工类型等,这些元标签可以帮助业务人员理解标签信息、选

择并使用标签。而字段名、负责人、完成时间等元标签更偏向技术方向,主要记录一些开发实施过程中的相关指标,与标签治理、运营等管理技术有关。标签名称的设计尤为关键,一定要避免产生侵犯隐私的误解,尽量不使用如“轨迹”“定位”“GPS”等敏感词汇。对于算法模型产出的标签,建议增加“预测”二字。同类标签应使用同类语句结构,例如,“近2个月买车类APP关注度”和“买车类APP近2个月的关注度”是不同的表达方式,应当统一使用一种语句结构。

3.1.3 标签存储设计

数据存储技术的选取至关重要,不同的数据存储方案各具特点,适用于标签体系建设中不同环节与应用场景的需求。

在标签开发阶段,一般采用Hive作为数据仓库,用于存储用户特征、标签与维度表等相关历史数据,为标签设计与开发提供支撑。对于标签体系元数据等小规模数据,MySQL拥有更快的读写速度,性能稳定,适用于元数据管理以及ETL结果监控与标签生成情况的跟踪。HBase具有高吞吐量与低延迟的特点,能满足大规模实时查询与分析的需求,能支持单用户标签查询与多维度组合查询,常用于风控点管理等场景。HAWQ结合了MPP数据库的技术优势与Hadoop的高扩展性,性能表现优异,适用于实时围栏选择、时空分析建模等高性能计算场景。

综上,这4种存储方式存储技术各具特色,能全面支撑标签体系从开发到应用的全部流程。运营商在规划构建标签体系时,需要根据自身数据与业务发展现状,选择最适宜的存储方案组合与架构。

3.2 标签开发

标签开发是构建标签体系过程中最关键的阶段。从标签生成方式上看,标签可分为事实类、规则类和预测类。事实类标签直接来源于原始数据字段;规则类标签的开发依赖业务规则与数据验证,更取决于人工推理与判断;预测类标签开发需要运用机器学习模型与算法,通过训练与评估实现自动化生成。这3种标签相互补充,共同构建全面智能的标签体系。

3.2.1 事实类标签

事实类标签是直接对底层数据进行加工和处理得到的标签,在加工时,需要结合业务知识,去除其中的脏数据和重复数据,并进行必要的数据格式化和标准化。根据使用需求,将原始数据转化为离散型标签和连续型标签这2种形式。针对连续型标签,还可以

考虑对标签内容进行归一化处理,以符合统一表示规则,方便后续标签的应用和管理。

3.2.2 规则类标签

规则类标签是通过定义一系列规则,根据底层数据的特征进行匹配得到的标签。这些规则可以基于人工经验、行业知识、专家建议等多种来源。在规则类标签定义过程中,需要在规则权重、规则范围划分、逻辑冲突、时效性更新、可解释性等方面深入思考和论证,保证标签的精准性和实用性。下文以如何利用客观赋权法和时间衰减系数解决兴趣偏好分标签开发中遇到的规则权重、规则范围划分问题为例,简述规则类标签开发中的一般流程。

3.2.2.1 CRITIC 客观赋权法

构建规则类标签时,开发人员经常面临数据特征赋权的问题。主观赋权过程中决策者完全脱离实测数据,只凭借经验和专业知识,极易造成主观偏好过强的情形出现。

CRITIC 法是 Diakoulaki 提出的一种客观赋权方法,其主要思路是利用指标的对比强度和冲突性来衡量指标权系数。对比强度是指同一指标在各决策目标间取值的差异性大小,用标准差来度量。冲突性以指标间的相关性为基础,若指标间有较强的正相关性,它们之间的冲突性就较低。CRITIC 法既兼顾到指标数据的相关性,又考虑到数据之间的对比强度。

$$w_j = \frac{\sigma_j \sum_{i=1}^n (1 - r_{ij})}{\sum_{j=1}^m \sigma_j \sum_{i=1}^n (1 - r_{ij})} \quad (1)$$

其中, w_j 是第 j 个指标权重, r_{ij} 是指标 i 与 j 的相关系数, σ_j 是指标 j 的标准差。

本文定义,长期兴趣偏好分的计算是基于多个兴趣偏好指标及其对应权重的加权求和。首先,计算每个兴趣偏好指标的数据值与其权重的乘积,然后将所有乘积之和进行加权求和(S_L),如式(2)所示。其次,将加权求和值在 1~100 内进行归一化,得到初步的长

期偏好分值 S_{Lstd} 。最后,取最近 3 个月的长期偏好分平均值,得到最终的长期兴趣偏好分 P_L 。

$$S_L = \sum_{j=1}^n w_j \times w_j \quad (2)$$

$$P_L = \frac{1}{3} \sum_{i=1}^3 S_{Lstd,i} \quad (3)$$

3.2.2.2 时间衰减系数

部分兴趣偏好具有强烈的时效性,如旅游,用户在计划旅游与出行期间会大量使用相关 APP,但旅游结束后使用量会急剧下降。这种兴趣周期通常在几天到几十天,因此,使用长周期数据(3个月)的偏好分析难以准确刻画用户的旅游兴趣偏好。

基于 CRITIC 的客观赋权法没有考虑数据的时间性。因此,在该算法的基础上增加时间衰减系数,模拟用户兴趣随时间的衰减情况。不同场景下,时间衰减系数由相应人群的 APP 使用数据确定。与固定的短期窗口(如近 1 个月)相比,这种方法可以更精准地量化用户的短期兴趣偏好。

不同领域的短期兴趣偏好存在各自的衰减周期,假设用户兴趣最强烈时,其程度可量化为 $N(0) = 1$ 。当兴趣衰减至一半,即 $N(t) = 1/2$ 时,所用时长为 Δt ,参考牛顿冷却定律公式:

$$N(t) = N(0) \times \exp(-\lambda \times \Delta t) \quad (4)$$

在已知 $N(t)$ 和 $N(0)$ 的情况下,如果知道某种兴趣的半衰期,便可推算出衰减系数 λ 。将 λ 代入式(4),即可确定这种兴趣在不同时间的衰减情况,并用于计算时间权重。

3.2.2.3 旅游短期偏好分标签实现

本文提出了一种估算用户短期兴趣衰减周期的方法,通过统计用户在近 3 个月内访问旅游出行类 APP 的访问次数分布,并观察用户访问次数的变化趋势,估算出旅游兴趣的衰减周期。这一方法也适用于其他兴趣领域。图 2 展示了短期偏好标签整体加工流程。

以 2021 年 4 月为基准月份,提取活跃用户在 2021

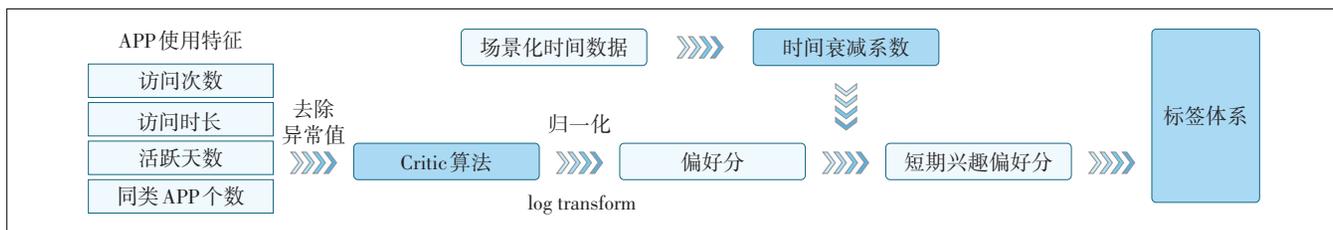


图 2 短期偏好标签加工流程

年4月及前2个月内每个月访问旅游出行类APP的次数,并统计其分布情况,如表1所示。

表1 2021年用户旅游出行APP月度使用情况统计分析

月份	次数	均值	标准差	下四分位数	中位数	上四分位数
2月	99 429 294	131.33	616.01	4	15	70
3月	107 882 924	167.73	776.09	5	17	86
4月	113 151 452	1 628.56	96 436.44	6	26	175

通过统计用户访问次数的分布,选择特定样本进行短期兴趣衰减周期的分析与量化估算。提取2021年4月访问旅游类APP在100次及以上,且2021年2月与3月访问均17次及以下的用户作为研究样本,共计362万用户。计算这部分用户每日访问旅游类APP的平均次数,结果如图3所示,其横坐标代表从2021年4月1日开始的日期,纵坐标代表当日样本人群访问旅游类APP的平均次数。

若以用户4月30日的日访问次数(图3中为520.6次)作为最高兴趣度的位置,则以其后访问次数减半的位置作为半衰期位置。从图3可以看出,用户日访问次数在5月6日(314.0次)与5月7日(186.3次)之间减半,因此半衰期可定为两者的中间位置,约为260次。根据半衰期理论,兴趣衰减的时间跨度为兴趣度

从最高点下降到半衰期位置的天数。因此,可判断用户对旅游类APP的短期兴趣衰减周期为6.5天左右。

将半衰期带入牛顿冷却定律公式,得到前3个月的时间权重分别为 $w_1=0.9593$, $w_2=0.0391$, $w_3=0.0016$,短期偏好分的计算方法如下:

$$P_s = \frac{1}{3} \sum_{i=1}^3 w_i \times S_{Lstd_i} \quad (5)$$

3.2.2.4 旅游短期偏好分标签验证

为验证短期偏好分的效果,本文选取KL散度(Kullback-Leibler Divergence)作为评判2个数据集分布相似度的指标。KL散度是从概率分布的角度衡量2个变量间的距离,如在连续区间D中有2个概率分布 P_1 和 P_2 ,则其KL散度计算如下:

$$D(P_1 \| P_2) = \sum_{x \in D} P_1(x) \log_2 \frac{P_1(x)}{P_2(x)} \quad (6)$$

其中, $P_1(x) > 0, P_2(x) > 0$ 。KL值越小,表示数据分布越相似。

根据与某OTA客户的合作项目,选择合规样本的意向度分布为目标分布,分别计算该样本的长、短期偏好分,数据结果分布如图4所示。为比较长期偏好分和短期偏好分的数据分布与目标分布的接近程度,计算它们与目标分布之间的KL散度。结果显示,短

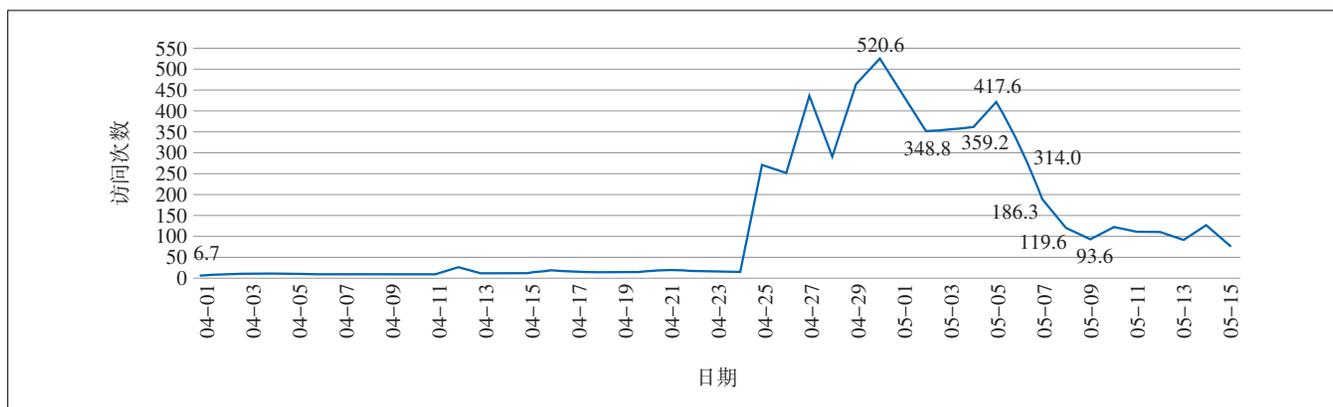


图3 旅游类APP用户日使用均值变化趋势

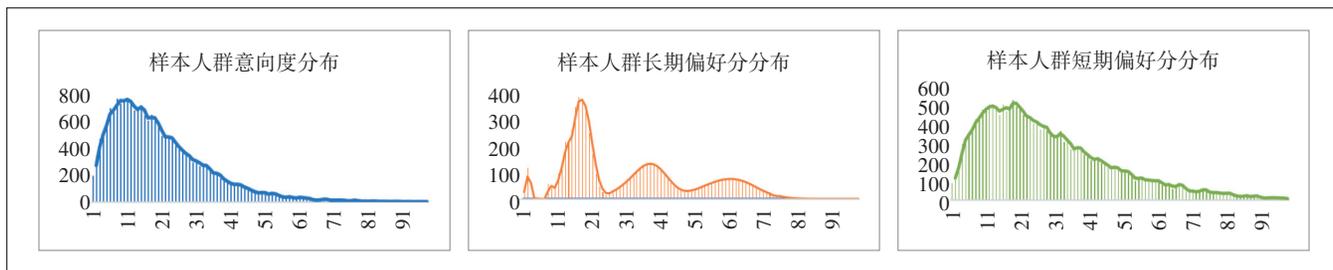


图4 样本人群意向标签验证对比

期偏好分与目标分布之间的KL散度更小。因此,可以判断短期偏好分的分布与目标分布最为接近,效果符合预期目标。

3.2.3 预测类标签

预测类标签是通过算法挖掘用户的相关特征生成的,如用户行为预测、意向程度预测等。预测类标签开发周期较长,需要经历多个环节,含样本构建、特征工程、算法调优与工程化部署,下面以贷款意向预测标签为例来介绍预测类标签的开发。

3.2.3.1 样本构建

根据与银行的合作项目,选择合规样本构建训练集。正样本为标记贷款意愿较强人群,共3万用户,负样本为贷款意愿较低人群,共5万用户。

3.2.3.2 特征工程

根据行业知识与业务场景选择与目标变量强相关的700余个用户多维特征,如身份特征、经济特征、社交特征等。通过特征工程发现,正负样本在经济水平、月出账金额、居住城市等级标签等特征分布差异较大,最终保留100余个入模特征。特征工程的具体实现步骤如图5所示。

3.2.3.3 算法调优与工程化部署

为构建贷款意向预测标签,采用XGBoost模型训练分类模型,并通过交叉验证评估与优化模型性能。在此基础上,使用网格搜索方法在训练样本上选择超参数,如学习率与叶子节点数等,提高模型准确度与泛化能力。最终模型AUC达到0.9,满足应用效果需求。

将经过训练与优化的XGBoost模型通过Hive-Transform工具部署至生产环境,实现全量贷款意向预测与标签生成。

3.3 标签使用

标签体系使用分为数据同步、数据应用与数据服务3种方式。

a) 数据同步是一种相对简单直接的应用方式,需使用方提前熟悉数据,难以追踪使用效果,适用于数据使用方对数据与标签高度熟悉,追求简单快速获取数据的场景。

b) 数据应用可以跟踪标签调用情况与评估效果,但与业务高度耦合,定制性需求多,扩展性有限,适用于对数据应用与效果有明确预期目标,并且使用数据范围集中的场景。

c) 数据服务是最理想的应用方式,业务人员可以

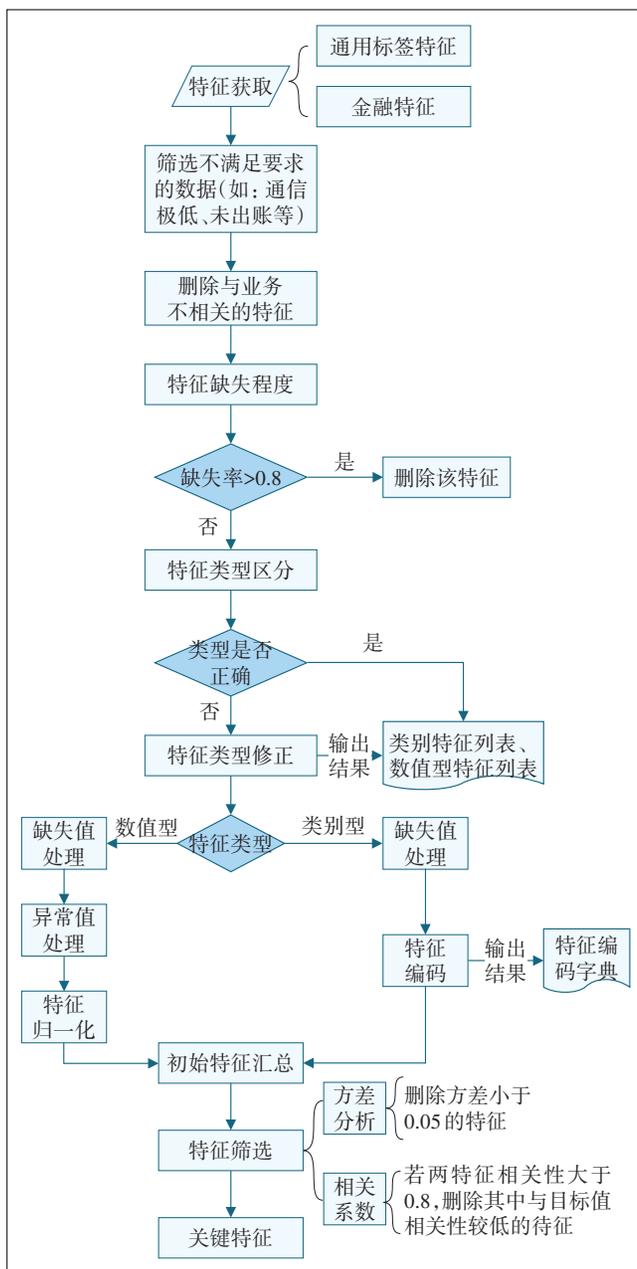


图5 贷款意向标签特征工程流程

灵活使用标签而不需要复制数据,适用于数据使用范围广、定制需求多变且追求高扩展性的场景。标签提供方可以在保证标签质量的前提下最大限度满足使用方的个性化需求。

4 运营商标签体系行业应用实例

4.1 标签体系辅助提升应急管理能效

利用运营商标签数据天然的时空连续性,在应急领域以人、地为分析主题,构建了10余个应急场景下的人群时空标签,如异常煤矿监控标签可为关停煤

矿提供监测与分析服务,有助于发现偷采等违规行为,预防煤矿事故;汛期人口转移、安置、回流等应急管理标签可为洪涝等自然灾害进行事前、事中、事后的转移安置,提供异常情况分析服务;地震灾害标签可为受灾人群判断、灾后评估与聚集地判定等提供实时与离线监测分析服务。

4.2 标签体系服务支撑交通运输分析

基于运营商全量数据进行游客识别、游客行为特征分析挖掘,结合信令数据时空连续性优势从客流、客源、画像、出游等维度进行深度分析,从管理、服务、推荐3个方面探究全域旅游态势及未来趋势,推进数据应用产品落地。

基于时空连续性信令OD分析模型及机器学习算法,对全国客流出入枢纽的轨迹数据进行统计分析,构建了交通出行监测模型。针对春运期间返乡、返程、复工、农民工等特殊行为和职业构建标签,进行专题人群画像分析,辅助春运期间交通运输分析及管理,全方位分析全国人口流动态势。

4.3 标签体系助力提升欺诈风险防控

利用中国联通号码、异网号码、脱敏身份等9类节点,月度同网通话关系、异网通话关系、中国联通用户对应职住社区编码等17类边,以及基础信息、通信偏好、金融偏好、风险评级等数据标签作为号码属性,构建了亿级节点、千亿级边规模的运营商通话关系图谱。依托图数据库强大的计算引擎和区别于传统关系型数据库的计算方式,该关系图谱可实现千亿级边秒级查询服务。

对于合规样本中的异常理赔人群,可基于图计算引擎快速对其内外部共有联系人进行研判分析,准确识别出56%的号码存在亲密关系以及15%的异常号码,进而采用数据挖掘、机器学习等技术,对风险团体成员的联系人、地点、时间等进行关联分析,找到团体活动的规律和特征。通过挖掘到的隐藏规律和模式,形成风险团体标签,并输出合规的定制化保险欺诈分析报告。

4.4 “标签体系+隐私计算”保障数据要素安全流通

基于互联网行业和电信运营商的海量用户数据,借助隐私计算“数据可用不可见”的能力,结合行业模型标签能力,形成“多方安全计算+标签”、“联邦学习+特征”的服务模式,并在项目实践中不断规范完善相关标准、开发流程和审核机制。

某互联网行业企业希望对旅游业务模型进行综

合治理与提升,需要补充外部数据与模型能力。基于数据安全合规要求,利用隐私计算技术拉通中国联通与客户数据,联合建模。该联合模型指标在该企业自有数据模型指标的基础上提升6个百分点,同时在业务侧原有下单意愿低于30%的用户量提升180%,原有下单意愿在30%~70%的用户提升61%。

5 结束语

数字化时代,数据资产已成为企业核心竞争力的源泉。运营商持有海量数据资源与技术优势,构建标签体系可以实现数据资产的全面产出与价值释放。本文深入研究了运营商构建标签体系的主体架构、技术路径与行业应用案例,系统地阐释了构建高质量标签体系的方法与步骤,为行业数字化转型与产业升级提供重要理论参考与实践指导。

随着5G、大数据、云计算与人工智能等技术的融合创新,数字经济进入快速发展期,为运营商构建标签体系提供了更加广阔的市场空间与技术条件。运营商应在数据安全合规的前提下,运用新技术手段不断优化标签体系,加速行业数字化转型进程。

参考文献:

- [1] 费鹏,林鸿飞,杨亮,等.一种用于构建用户画像的多视角融合框架[J].计算机科学,2018,45(1):179-182,204.
- [2] 宋美琦,陈焯,张瑞.用户画像研究述评[J].情报科学,2019,37(4):171-177.
- [3] 亓丛,吴俊.用户画像概念溯源与应用场景研究[J].重庆交通大学学报(社会科学版),2017,17(5):82-87.
- [4] 刘海,卢慧,阮金花,等.基于“用户画像”挖掘的精准营销细分模型研究[J].丝绸,2015,52(12):37-42,47.
- [5] 李斯.大数据背景下面向运营商精准营销的用户画像研究[D].辽宁:大连理工大学,2019.

作者简介:

李大中,毕业于北京邮电大学,高级工程师,硕士,主要从事一体化大数据平台、大数据能力开放、数据治理、数据可视化、数据科学等大数据全生命周期能力建设、实施交付、咨询服务等工作;张倬荣,毕业于香港中文大学,工程师,硕士,主要从事运营商大数据分析、模型设计与开发、运营商标签体系建设等工作;闫龙,毕业于利物浦大学,工程师,博士,主要从事运营商大数据数仓保障与优化、模型设计与开发、数据科学组件集成与交付、数据合规流通及各类数据科学咨询服务;王功举,毕业于中国人民解放军军事科学院,硕士,主要从事人工智能、机器学习、深度学习相关模型开发工作;宋雨伦,毕业于格拉斯哥大学,高级工程师,博士,主要从事核心数智链技术能力建设,行业解决方案及实施,打造数据治理、服务、应用、安全四大产品体系等工作。