



中国电信大数据普惠金融应用实例

I. 数据来源

天翼大数据，主要涵盖了中国电信运营服务中所产生的各类基本数据及通信娱乐、交互行为、位置轨迹等全域数据。

II. 概念定义

- **普惠金融**：指以可负担的成本为有金融服务需求的社会各阶层和群体提供适当、有效的金融服务，小微企业、农民、城镇低收入人群等弱势群体是其重点服务对象。
- **大数据**：指为了适应海量、高增长率和多样化的信息资产，采用各类软硬件技术优化处理模式，使数据能够具备更强的决策力、洞察发现力和流程优化能力。
- **金融科技**：指将大数据、云计算、区块链、人工智能等多项高新技术整合创新，从而使其可以全面应用于支付、清算、融资租赁、保险、互联网金融等方面，提升金融产业的效率。
- **权益方案**：指根据人群的特性兴趣，设定一系列不同的权益吸引受众。
- **特征映射**：对字段进行相关的特征处理，映射成为建模所需要特征变量。
- **训练集**：指在利用机器学习建模过程中，用来估计模型的样本集。
- **测试集**：指在利用机器学习建模过程中，用于测试训练好的模型的分辨能力的样本集。
- **原始变量**：直接将初始字段用于建模的变量。

- **衍生变量**：通过对原生变量进行加工而产生的变量。

III. 数据定义

- **网络行为覆盖率**：指具有相关行业网络行为的人群占总体人群的比例。
- **ROC曲线**：受试者工作特征曲线（ receiver operating characteristic curve ），是反映敏感性和特异性连续变量的综合指标,是用构图法揭示敏感性和特异性的相互关系，它通过将连续变量设定出多个不同的临界值，从而计算出一系列敏感性和特异性，再以敏感性为纵坐标、（ 1-特异性 ）为横坐标绘制成曲线，曲线下面积越大，诊断准确性越高。
- **AUC**：Area Under Curve，ROC曲线下面积与整个图形的占比，用来评定分类模型区分度的值。通常，AUC的值介于0.5到1.0之间，较大的AUC代表了较好的模型效果。
- **正样本**：还款时间小于等于N天（约定还款期限）的用户，N值会根据业务和建模需求进行调整
- **负样本**：还款时间大于N天（约定还款期限）的用户，N值会根据业务和建模需求进行调整。
- **短信响应率**：短信响应是指用户接收到短信后，按照短信内容提示点击链接或者回复相关内容；短信响应率=响应短信的人数/发送短信的人数。

- **进件率**：进件是指用户提交业务申请并填写相关资料，进件率=有相关操作的人数/短信响应的人数。
- **批核率**：批核是指银行审核用户的资质后，准予对用户进行发卡，批核率=批核人数/进件人数。
- **一卡批核率**：一卡批核是指首次申请该行信用卡的批核人数，一卡批核率=一卡批核人数/批核人数。
- **综合营销效果**：综合营销效果=短信响应率*进件率*批核率*一卡批核率=一卡批核人数/发送短信人数。
- **促活率**：促活率=响应行方活动的人数/筛选出需要营销的人数。
- **有值**：指匹配到相关数据的客户样本。
- **无值**：没有匹配到相关数据的客户样本。
- **准确率**：准确率=预测正确（将正样本预测为正和将负样本预测为负）的数量/所有样本数。
- **召回率**：召回率=将正样本预测为正的数/（将正样本预测为正的数+将正样本预测为负的数）。
- **KS**：KS值表示模型将正样本及负样本区分开来的能力，值越大，模型的预测准确性越好。一般，KS>0.2即可认为模型有比较好的预测准确性。
- **区间违约率**：表示该评分区间内，坏样本的数量占该评分区间内所有样本数量的比例。

目录

CONTENTS

01 我国普惠金融发展现状

02 电信大数据的特点及优势

03 电信大数据金融行业应用案例解析

04 数据融合为普惠金融发展提供空间

国家大力推进普惠金融的发展

联合国会议

- 完善金融基础设施
- 成本可负担
- 向欠发达地区和社会低收入人群扩展
- 价格合理、方便快捷

习主席阐述经济新常态

- 普惠金融关乎经济结构调整和发展方式转型、关系人民福祉

“一带一路”开幕式主旨演讲

- 发展普惠金融；
- 完善金融服务网络

2005年

2013年

2014年

2015年

2017年

- 发展普惠金融
- 鼓励金融创新
- 丰富金融市场层次和产品

- 《推进普惠金融发展规划（2016—2020年）》
- 从国家层面确立普惠金融的实施战略
- 各大金融机构也相继成立普惠金融事业部

规划引领

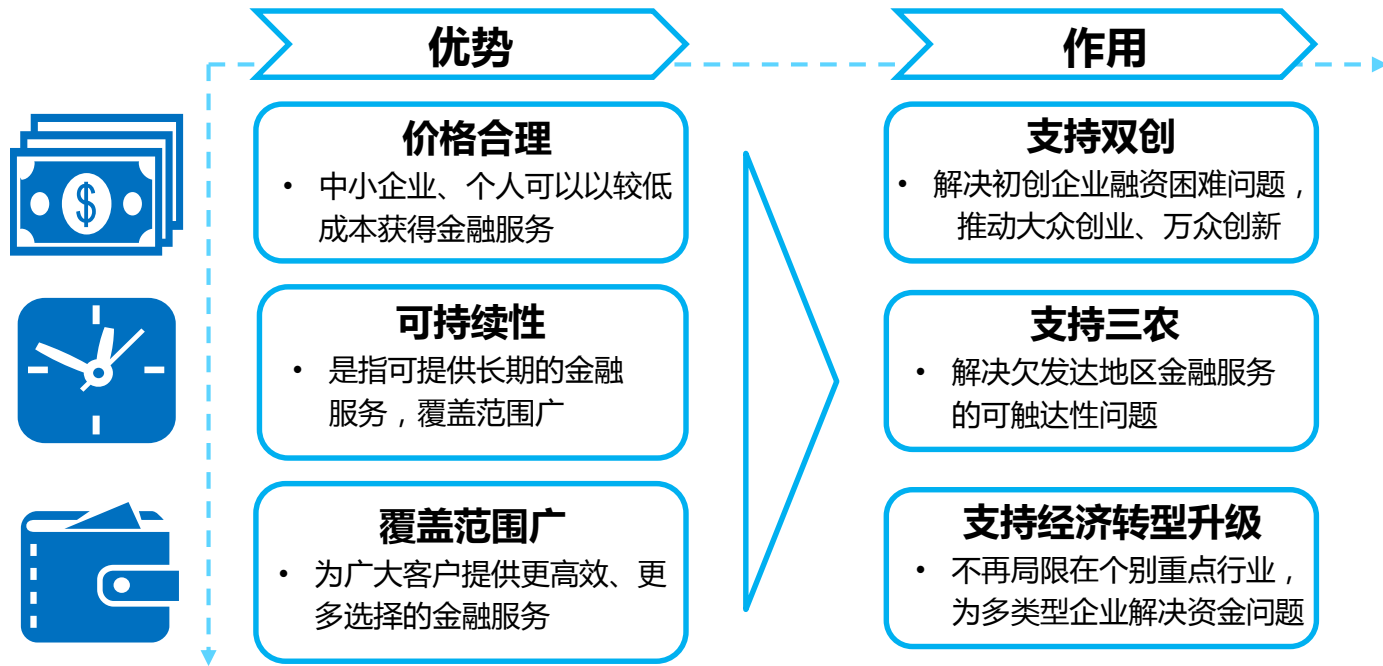
规划引领发展
战略路线明确
地位日益重要

我国十八届三中全会

国务院正式发布规划

普惠金融的优势和作用

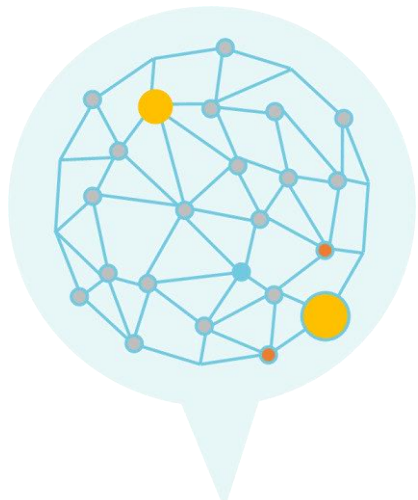
- 普惠金融为经济发展的转型升级提供积极的促进作用，同时增强了金融服务的可得性、获得感以及满意度。





当前普惠金融的制约因素

- 我国普惠金融虽然取得了一定成果，但也受到传统方式的各种制约，从而影响了普惠金融的发展。



制约因素



信息不精准

- 无法有效掌握新客户的各类偏好信息
- 无法进行有针对性的营销信息触达



覆盖不充分

- 线下门店无法覆盖每一个区域
- 线下门店无法覆盖每一个人群



营销成本高

- 线上广告难以准确触达，浪费较高
- 线上单个用户营销成本越来越高昂

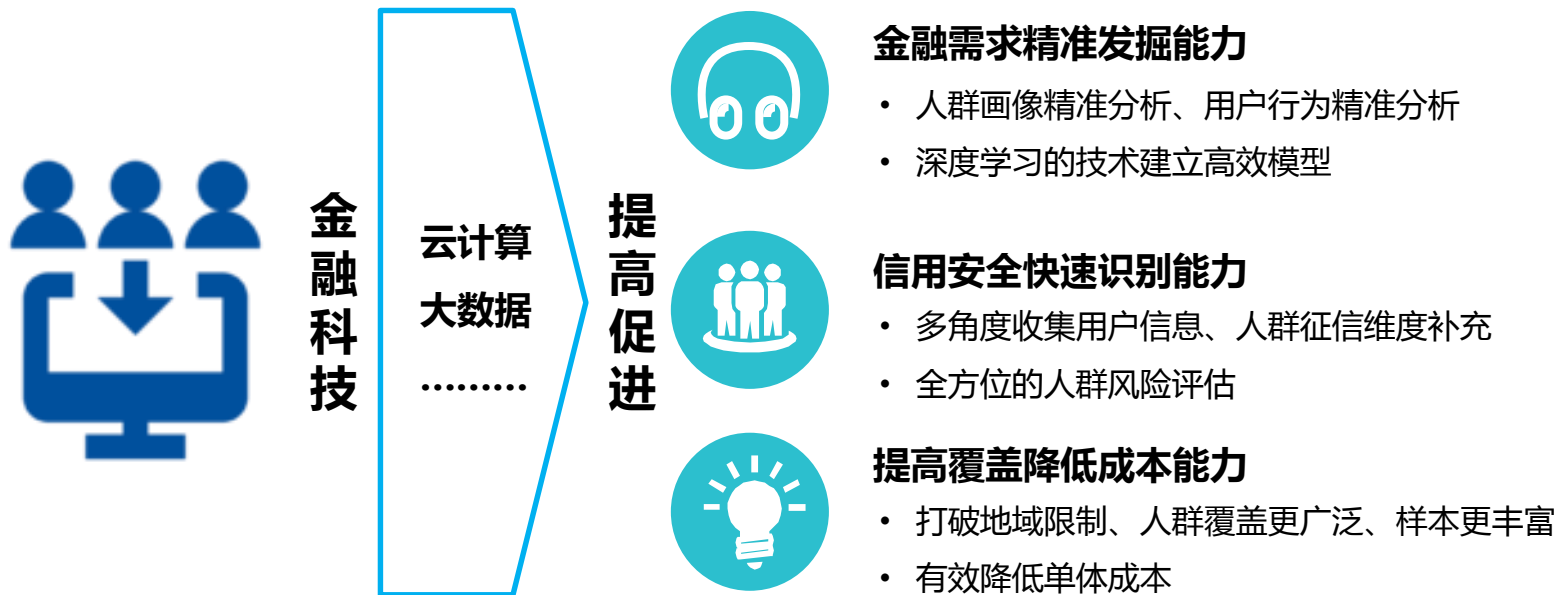


征信识别难

- 央行征信系统人群覆盖较少
- 央行征信系统接入成本较高

金融科技从三方面赋能普惠金融

- 利用金融科技特别是大数据手段能够提高金融机构的市场洞察能力
- 有效解决传统金融业务中的营销和风控问题，实现精准营销并且有助于完善风控机制



目录

CONTENTS

01 我国普惠金融发展现状

02 电信大数据的特点及优势

- 电信大数据的特点
- 提高营销的精确度和成功率
- 提高信贷周期的风控水平和能力

03 电信大数据金融行业应用案例

04 数据融合为普惠金融发展提供空间

中国电信大数据的优势和特点

中国电信大数据资产涵盖了电信运营服务中所产生的各类基本数据及通信娱乐、交互行为、位置轨迹等数据，数据资产具有**中立、横跨移固两网、多屏幕、高价值密度、基础信息完善、天然位置关联**等优势。

1.数据中立性

- 区别于互联网公司，数据最大特点是中立

2.数据跨多行业

- 横跨垂直网站、电商、政府、金融、社交、服务多领域数据

3.移固数据融合

- 通过宽带账号、手机号码、终端串码、Cookie实现用户全网多屏幕收敛

4.内容独特准确

- 现实生活交往圈、全天候位置轨迹、准确的用户信息等；多维度数据可准确关联并高效处理和传输



基本信息



语音通信



宽带网络



移动通信/移动网络



用户位置



WiFi网络



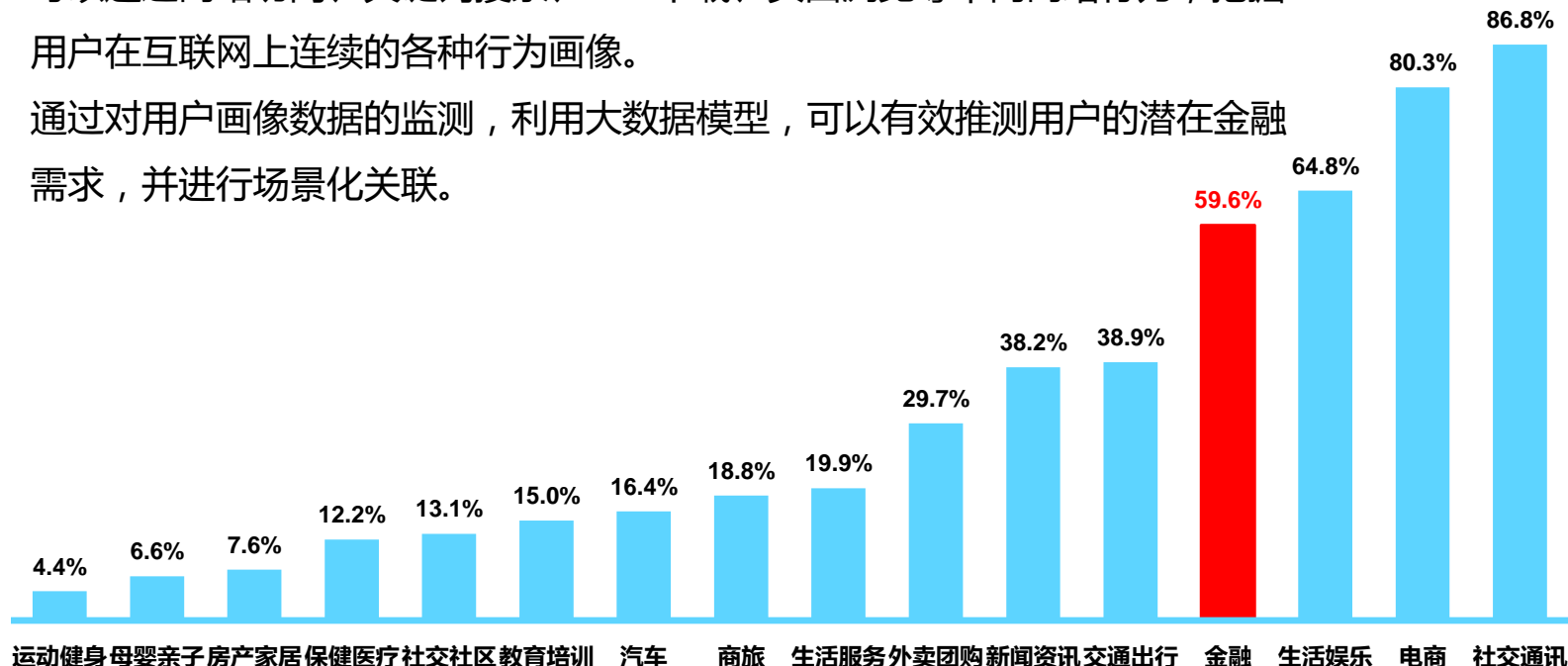
用户触点



场景挖掘提升用户价值

电信移动端用户各类网络行为覆盖率

- 可以通过网站访问、关键词搜索、APP下载、页面浏览等不同网络行为，挖掘用户在互联网上连续的各种行为画像。
- 通过对用户画像数据的监测，利用大数据模型，可以有效推测用户的潜在金融需求，并进行场景化关联。



数据来源：电信天翼大数据

目录

CONTENTS

01 我国普惠金融发展现状

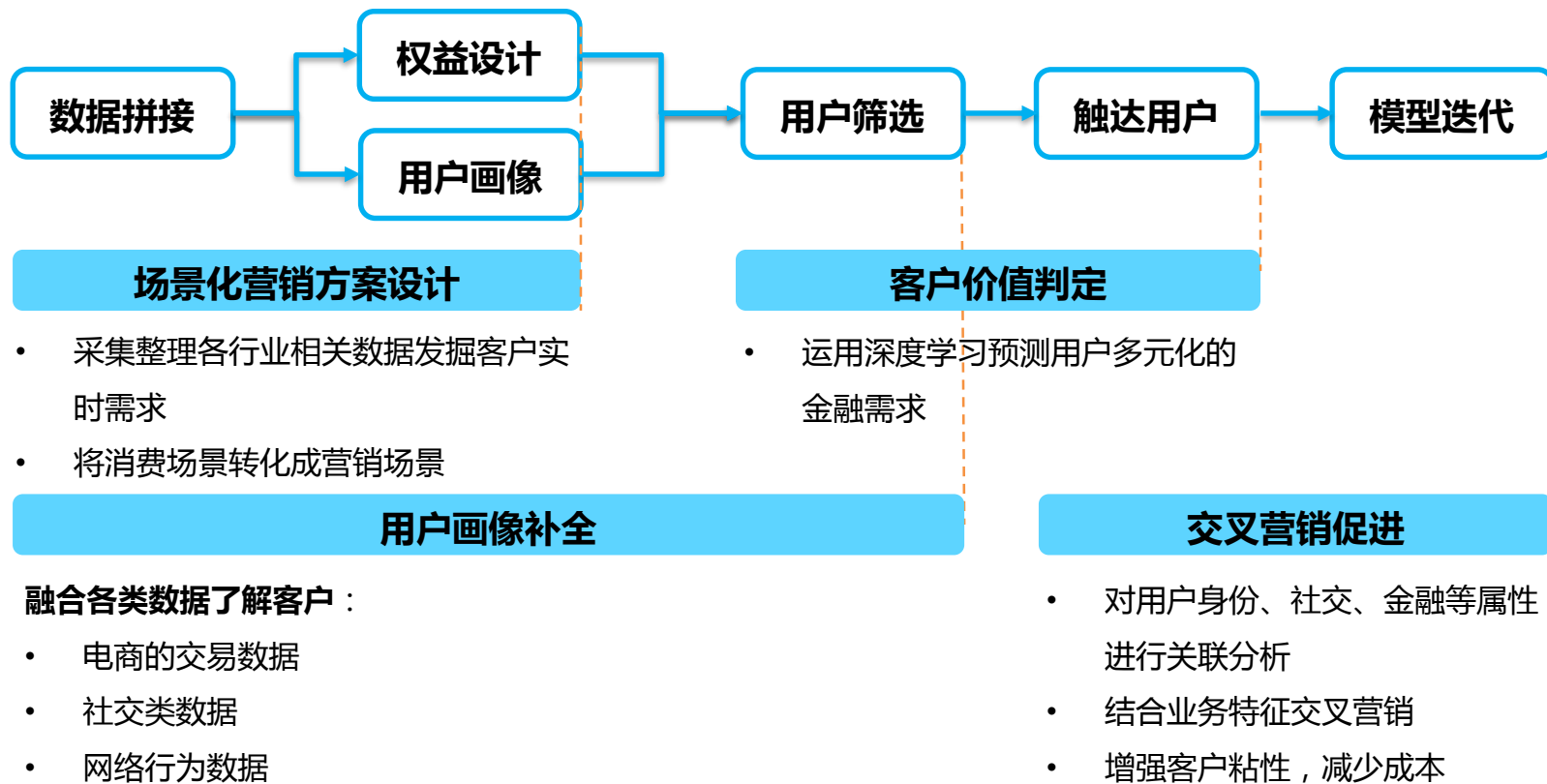
02 电信大数据的特点及优势

- 电信大数据的特点
- 提高营销的精确度和成功率
- 提高信贷周期的风控水平和能力

03 电信大数据金融行业应用案例

04 数据融合为普惠金融发展提供空间

数据融合可有效的发掘潜在营销需求点



精准用户画像能有效提高新客获取的成功率

重点依赖运营商数据

01



潜在新客识别

- 对样本客户分析
- 补全新客用户画像
- 挖掘需求人群

用户画像
用户筛选

02



行为分析及场景化分类

- 分析用户互联网金融行为
- 把目标客户分群
- 采取不同场景进行营销

权益设计
用户筛选

03



线索触达用户

- 不同用户采用不同权益
- 不同用户采用不同场景
- 多样渠道精准触达用户

触达用户

04



数据反馈及模型迭代

- 反馈各种方式营销的结果
- 进一步优化营销模型
- 提升转化率

模型迭代

场景化挖掘可有效激活存客用户的消费需求

01

数据拼接



- 金融机构数据敏感度高
- 金融机构数据合规要求高
- 由运营商将数据导入银行内部进行数据融合挖掘

02

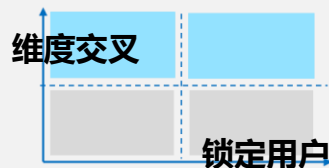
场景挖掘



- 确定营销方向
- 确定营销场景
- 优化不同场景下的营销方案和权益设计

03

价值名单



- 根据模型进行维度选择
- 多维交叉分析，细分客户价值
- 锁定场景营销的目标客户

04

数据反馈



- 营销方式持续反馈
- 根据结果循环优化模型
- 模型再指导营销，迭代推进
- 最终持续提升营销效果

目录

CONTENTS

01 我国普惠金融发展现状

02 电信大数据的特点及优势

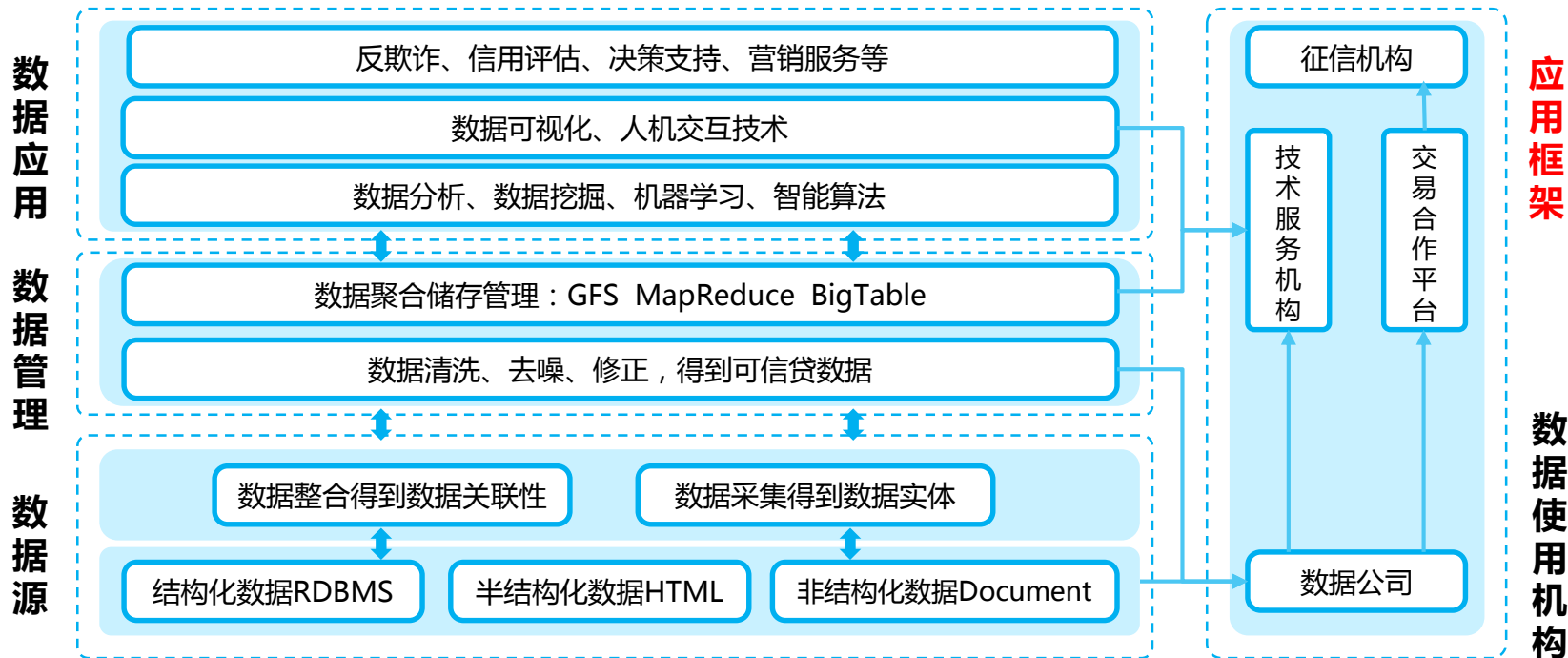
- 电信大数据的特点
- 提高营销的精确度和成功率
- 提高信贷周期的风控水平和能力

03 电信大数据金融行业应用案例

04 数据融合为普惠金融发展提供空间

中国电信大数据征信应用技术框架

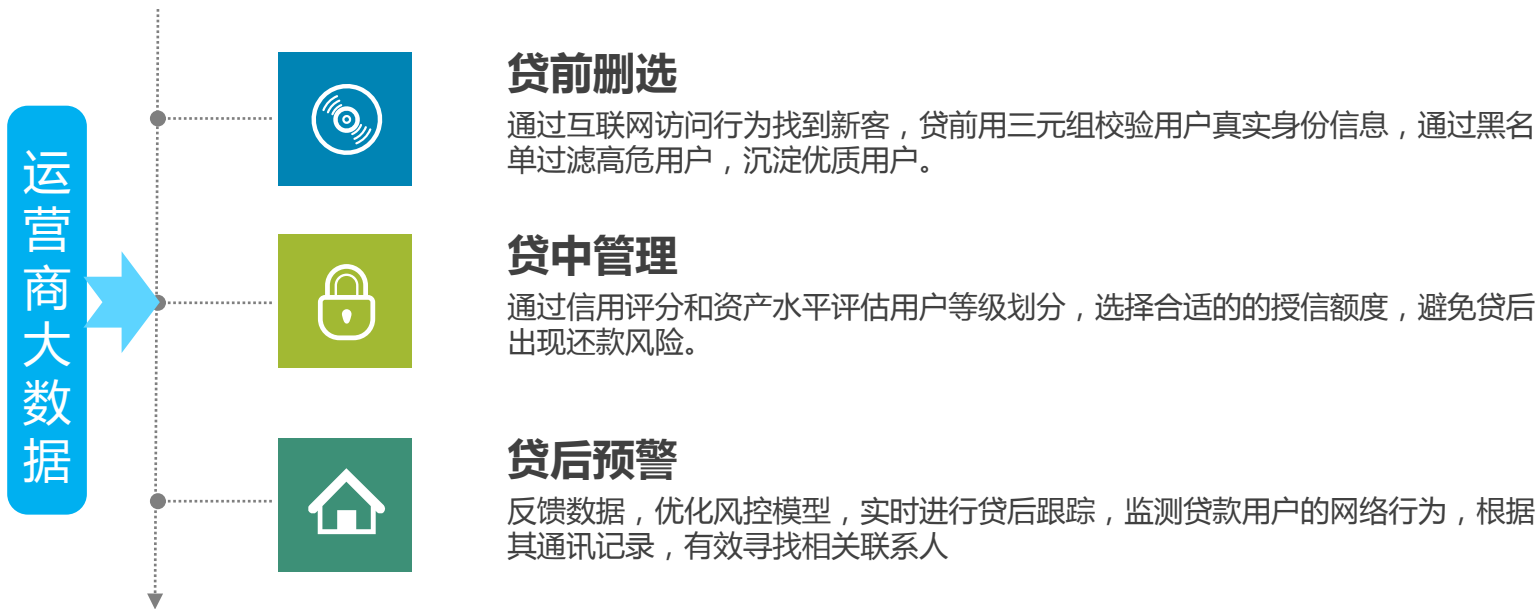
- 依托于运营商大数据和云计算的技术优势，可挖掘大量数据碎片中的关联性，推动征信模型不断完善，提升风控力度。



数据来源：电信天翼大数据

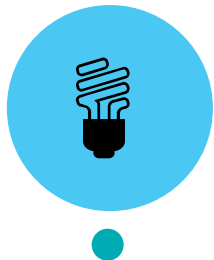
提高信贷全生命周期风控质量

- 普惠金融的本质是金融，但核心是风控，良好的风控体系可以使机构规避许多风险，而全生命周期的渗透可以帮助企业保持良好的运行状态，实现健康有序的发展。



贷前：高价值用户挖掘，降低风险成本

负面信息核查



- 黑名单查询
- 不良信息核验
- 多头申请
- 异常线上行为

身份核验



- 冒名信息识别
- “养卡” 欺诈识别

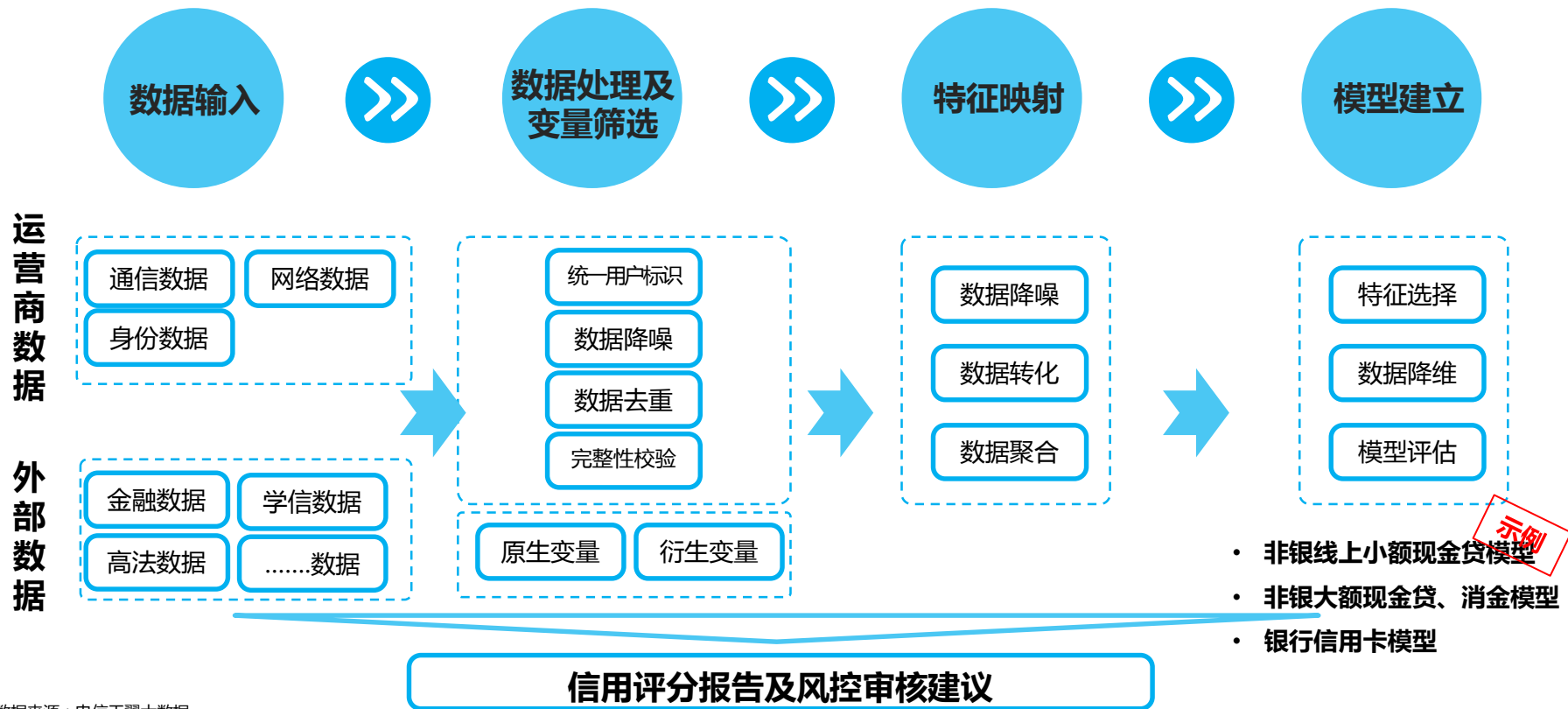
通信行为欺诈



- 夜间通话行为分析
- 通话静默情况分析
- 联系人号码数量分析
- 特殊号码通话情况分析
- 近三个月停机次数查询

贷中：风险评分模型精准快速识别逾期风险

- 以多维数据联合，建立合适的贷中评分模型，加速放款速度。



贷后：可实现贷后风险有效监控预警

- 运营商通过合规的手段，帮助金融机构获取更高价值的数据，实现更好的贷后预警



通信行为

- 通话行为
- 上网情况
- 人员属性
- 联系时间
- 联系频率
-



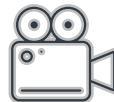
社交关系

- 社交属性
- 偏好分析
-



消费能力

- 消费层级
- 消费方式
-



生活方式

- 出行方式
- 职住区域特征
-

目录

CONTENTS

01 我国普惠金融发展现状

02 电信大数据的特点及优势

03 电信大数据金融行业应用案例

- 案例1——某银行信用卡新客户推广
- 案例2——某银行信用卡存客用户精准营销
- 案例3——某小额贷款产品的贷中评估产品应用

04 数据融合为普惠金融发展提供空间

步骤1：构建信用卡新客获取营销模型

- 利用运营商数据进行存量客户画像，根据画像匹配人群属性、预测借贷意愿、匹配消费场景，促进信用卡的营销推广。

选择推广信用卡范围



10款主题信用卡

- 标准，白金，境外游，车主卡，女人卡.....

用户属性分析

- 价值属性
- 兴趣偏好
- 信用卡需求
- 习惯偏好

- 分析各类信用卡人群的属性特征，建立筛选模型

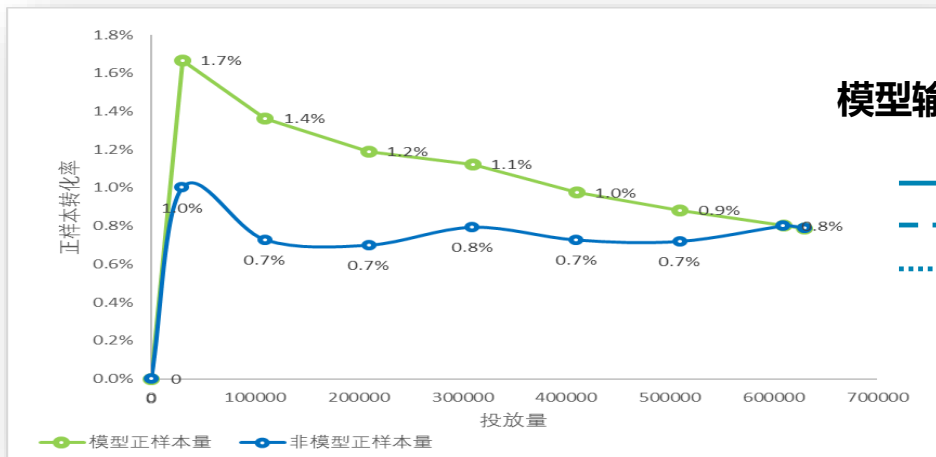
构建营销模型



- 通过模型输出营销指导建议

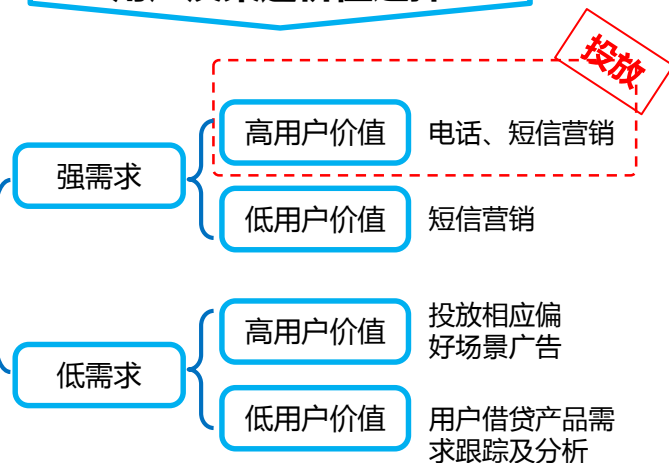
步骤2：根据数据反馈对模型进行迭代优化

模型训练及结果输出



模型输出指导

用户及渠道价值选择

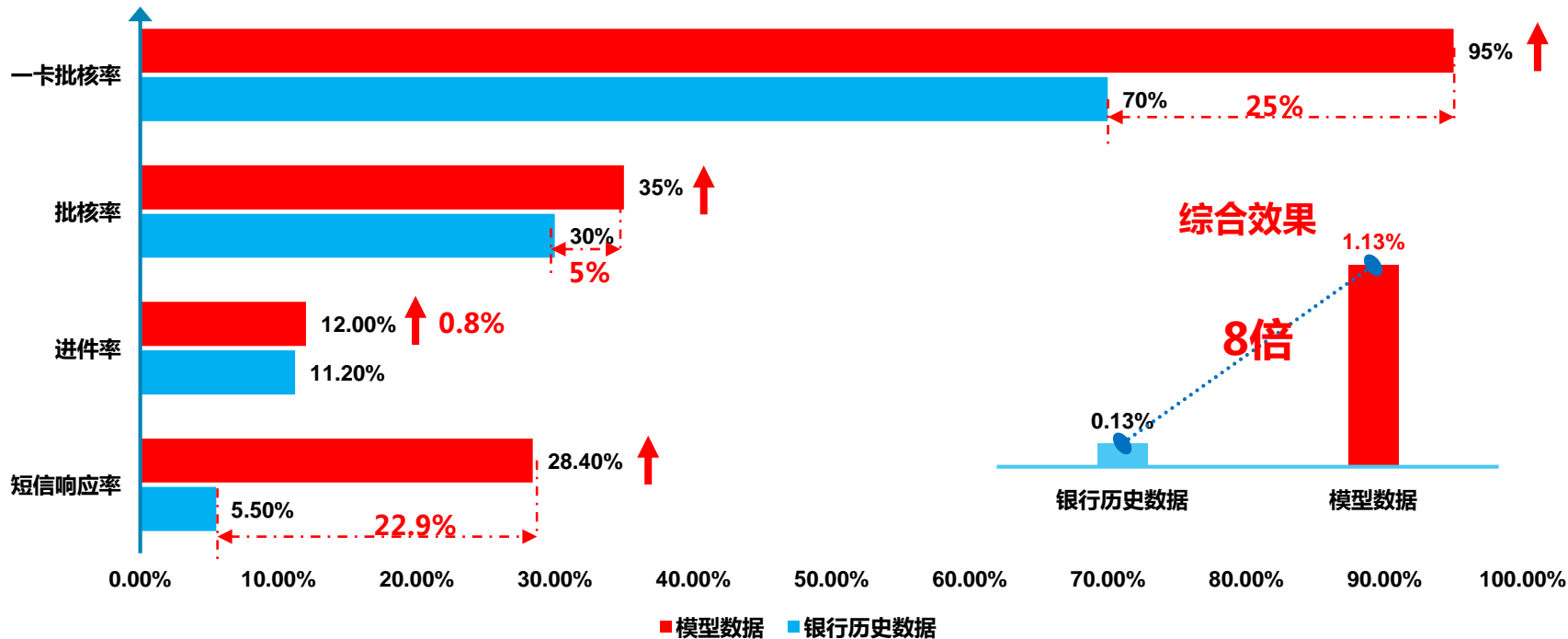


- 采用真实样本，根据实际业务场景训练模型；
- 训练集AUC表现为0.75，测试集AUC为0.68，发送相同的投放量，模型筛选出的正样本最高是旧规则的2倍。

根据投放结果反馈，不断迭代优化模型

步骤3：营销效果提升了8倍

通过模型筛选后，信用卡营销各阶段相关指标均高于历史水平



备注：

- 短信响应率：点击营销短信人数/营销执行人数
- 进件率：填写申请材料人数/点击营销短信人数

- 批核率：批核通过人数/填写相关申请材料人数
- 一卡批核率：首次在该行办理信用卡批核人数/批核通过人数
- 综合营销效果：响应率*进件率*批核率*一卡批核率

数据来源：电信天翼大数据

目录

CONTENTS

01 我国普惠金融发展现状

02 电信大数据的特点及优势

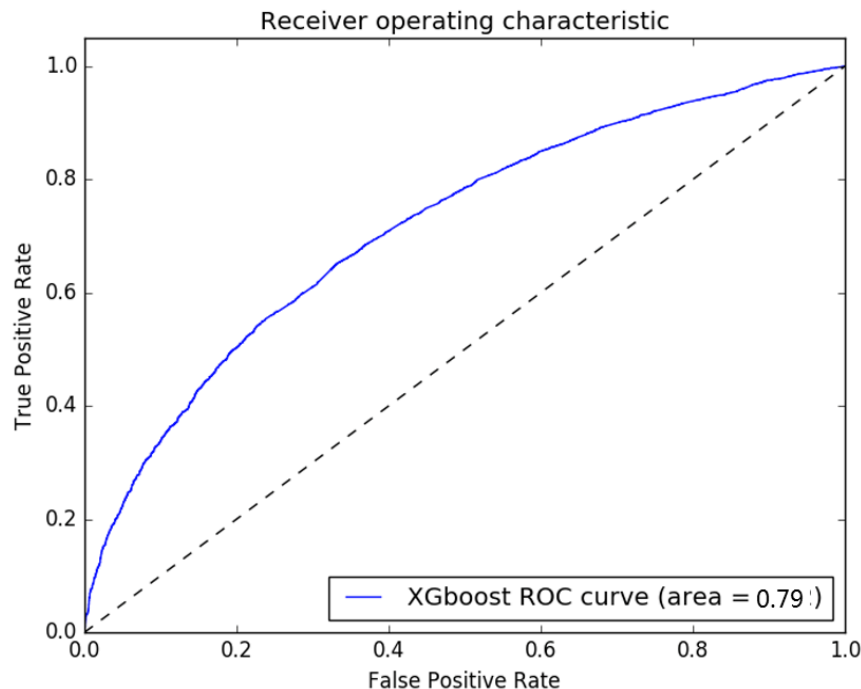
03 电信大数据金融行业应用案例

- 案例1——某银行信用卡新客户推广
- 案例2——某银行信用卡存客用户精准营销
- 案例3——某小额贷款产品的贷中评估产品应用

04 数据融合为普惠金融发展提供空间

步骤1：建立存客信用卡激活营销模型

营销模型特征曲线



模型相关说明

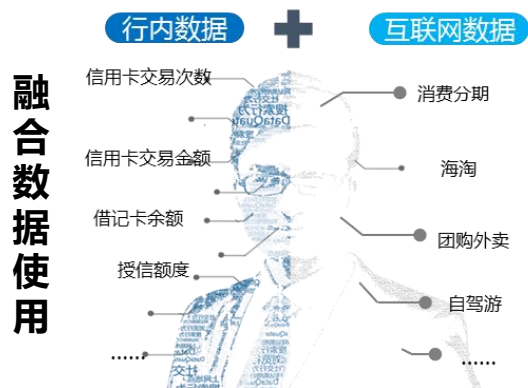
有效变量：

- 通过模型训练得出，用户的历史刷卡次数、刷卡金额、年龄、借记卡余额、客户等级，是影响客户能否被顺利激活的重要原因

效果呈现：

- 训练集：测试集 = 7:3
- 训练集AUC：0.81
- 测试集AUC：0.79

步骤2：构建多个用户参照组进行对比实验



引入运营商数据

- 从银行现有客户出发，结合银行内数据和运营商数据；
- 以用户使用场景为切入点进行场景化营销。

促活的可能性概率： <0.3

促活的可能性概率： ≥ 0.3

行内对照组

低分模型组

高分模型组

参照组细分

- 仅仅使用银行内部数据进行筛选。
- 主要包含用户的历史刷卡次数、刷卡金额、年龄、借记卡余额、客户等级等
- 模型组除采用行内相关数据字段外，还使用了各类互联网场景数据，匹配后进行综合建模
- 通过机器学习方法建立的模型，候选池样本超过总人群的20%，可营销范围极大扩充

数据来源：电信天翼大数据

步骤3：针对两个模型组用户进行精准权益设计

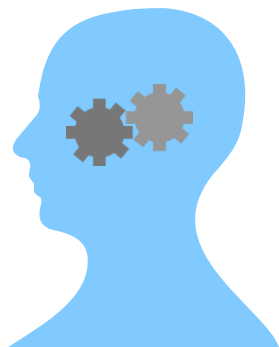
权益匹配：

- 根据我方数据匹配行内促活权益（吃的开心、玩的开心、淘的开心、购的开心）

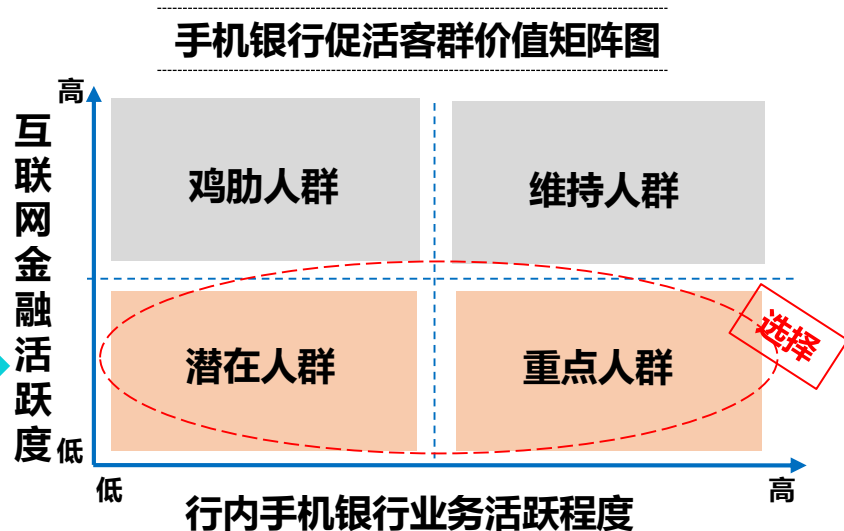


数据来源：电信天翼大数据

步骤4：选定高价值的用户触达群体



针对模型组中的用户群体，筛选高价值的人群作为本次活动的触达用户

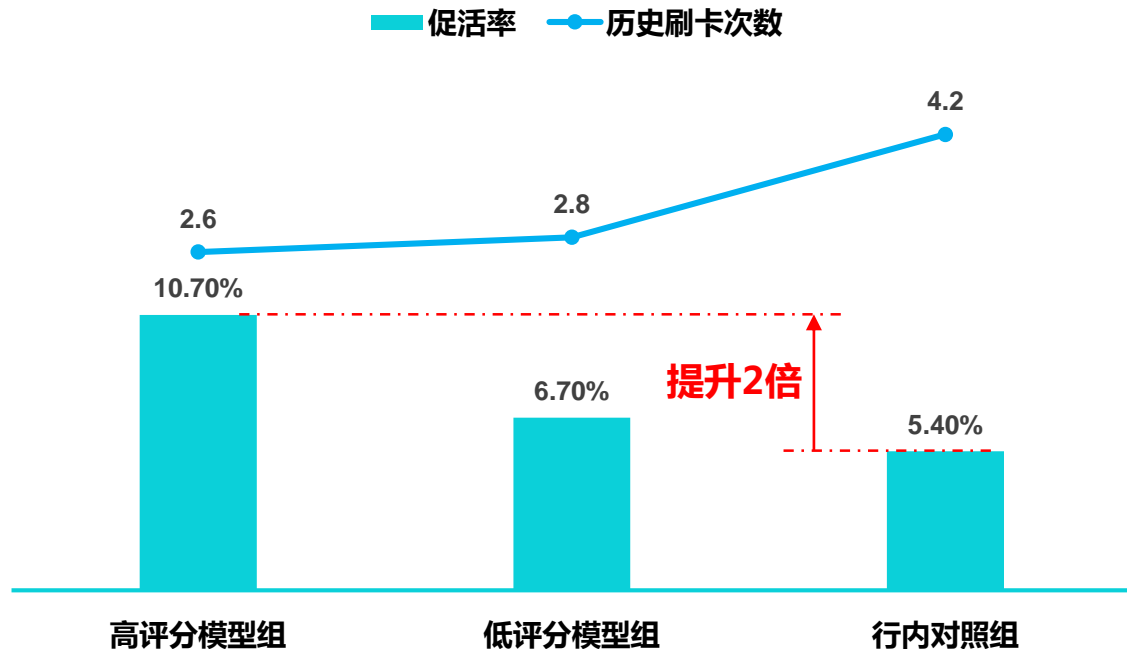


- **维持人群**：该类人群金融需求多样，数量占比一般在20%以内，投入一定资源维持即可；
- **鸡肋人群**：该类客户占比在30-40%，金融需求基本在其他渠道实现，抢回客户较为困难；
- **潜在人群**：该类客户占比在20-30%，目前金融需求不足，继续培育用户的使用习惯
- **重点人群**：一般占比在20-30%，已经建立一定使用习惯，需要通过手段强化使用习惯

***备注：**互联网金融活跃度根据用户互联网相关金融行为建模判别

步骤5：模型组促活率提高了2倍

三个用户组促活率对比图



图表结果分析

- 对照组人群的历史人均刷卡次数最高，但促活率依旧低于联合模型筛选组
- 低评分模型组促活率高出对照组 **20%**，高评分模型组促活率达到对照组的 **2倍**

目录

CONTENTS

01 我国普惠金融发展现状

02 电信大数据的特点及优势

03 电信大数据金融行业应用案例

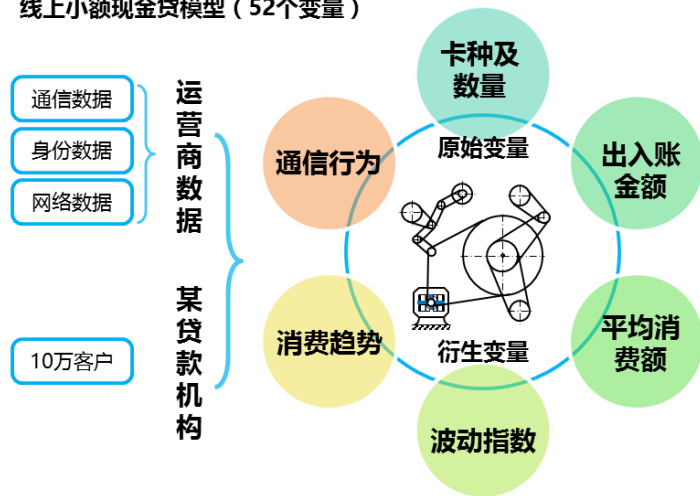
- 案例1——某银行信用卡新客户推广
- 案例2——某银行信用卡存客用户精准营销
- 案例3——某小额贷款产品的贷中评估产品应用

04 数据融合为普惠金融发展提供空间

步骤1：构建逾期风险评估模型

- 通过对用户身份数据的核验，并结合客户标签行为数据，挖掘高价值用户，降低风险成本

线上小额现金贷模型（52个变量）



与运营商数据总体匹配率80%，其中有值80281个，无值20288个

整体样本	总和	正样本	负样本	负样本比例	有值总和	有值正	有值负	有值负样本比例	无值总和	无值正	无值负	无值负样本比例
201606	14768	12615	2153	14.58%	11204	9905	1299	11.59%	3564	2710	854	23.96%
201607	13683	11840	1843	13.47%	11400	10201	1199	10.52%	2283	1639	644	28.21%
.....
201612	13310	10855	2455	18.44%	9381	8207	1174	12.51%	3929	2648	1281	32.60%
总和	100569	86123	14446	14.36%	80281	71479	8802	10.96%	20288	14644	5644	27.82%

备注：

- 样本来源：随机抽取某贷款机构16年6月-12月放款的10万客户；
- 样本比例：正负样本比例为8：2
- 负样本（坏客户）定义：逾期大于30天的用户；
- 正样本（好客户）定义：逾期小于等于30天

步骤2：模型达到了较好的预期效果

- 模型的主要指标KS值为0.342，AUC值为0.72，模型效果达到理想水平

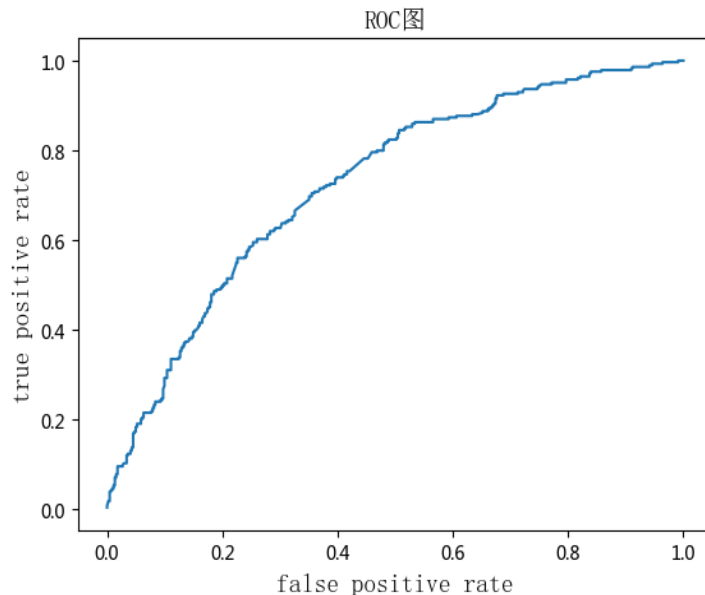
准确率：74% **召回率：51%**

KS=0.342

Ks检验：20%以上认为显著相关，40%以上认为极显著相关
该模型评分的KS值为34.2%，认为和用户是否逾期是强显著相关的。

AUC=0.72

总体来看，该模型评分的覆盖度比较高。从各方面检验来看，其与用户是否逾期也显著相关，具有很好的参考利用价值。



步骤3：坏账率降低了4.9%

- 若以505分作为通过阈值为例，与原模型对比，新模型将降低坏账率4.9个百分点，提升通过率2.2个百分点

提高了识别速度、增强了识别准确率，有效降低坏账损失、提高通过率并增加收入

评分	区间违约率		通过率		通过坏人占比		
(516,900]	<div></div>	0.096	<div></div>	0.1	<div></div>	0.096	评分较好，直接通过
(505,516]	<div></div>	0.178	<div></div>	0.2	<div></div>	0.137	
(495,505]	<div></div>	0.236	<div></div>	0.3	<div></div>	0.17	
(487,495]	<div></div>	0.274	<div></div>	0.4	<div></div>	0.196	评分适中，人工审核
(481,487]	<div></div>	0.361	<div></div>	0.5	<div></div>	0.229	
(477,481]	<div></div>	0.411	<div></div>	0.6	<div></div>	0.259	
(470,477]	<div></div>	0.486	<div></div>	0.7	<div></div>	0.291	
(460,470]	<div></div>	0.562	<div></div>	0.8	<div></div>	0.325	
(442,460]	<div></div>	0.611	<div></div>	0.9	<div></div>	0.357	
(300,442]	<div></div>	0.699	<div></div>	1	<div></div>	0.391	

评分较低，拒绝

评分较好，直接通过

评分适中，人工审核

评分过低，拒绝

目录

CONTENTS

01 我国普惠金融发展现状

02 电信大数据的特点及优势

03 电信大数据金融行业应用案例

04 数据融合为普惠金融发展提供空间

数据融合为普惠金融未来发展提供较大的优化空间

- 目前大数据金融应用还处于探索阶段，各类方案尤其是多维数据融合仍有较大的优化空间
- 在各方推动下，大数据应用整合也在不断加速，助力普惠金融更快发展



- 各方数据源统一整合应用
- 打破信息孤岛
- 关联数据能够得到有效应用

- 具备多方整合能力的大平台将提供多样化服务
- 为金融机构提供全面的一站式服务解决方案
- 助力金融服务发展

- 风控应用、营销应用相关应用还较为割裂
- 营销和风控一体化解决方案出现并更加成熟

中国电信股份有限公司云计算分公司（以下简称：云公司）成立于2012年3月，是中国电信旗下的专业公司，集约化发展包括互联网数据中心（IDC）、内容分发网络（CDN）等在内的云计算业务和大数据服务。云公司率先启动全网数据汇聚奠定大数据运营的坚实基础，已经走出了一条数据创新与孵化模式的商业运营道路。

天翼大数据作为云公司大数据品牌，天翼大数据资产涵盖了电信运营服务中所产生的各类基本数据及通信娱乐、交互行为、位置轨迹等全域数据，已汇聚四门14类数据，计划新接入16类数据，日均数据量已达200TB，日均记录条数已达2727亿。天翼大数据建设了“风险防控、精准营销、区域洞察、咨询报告、大数据云”的“4+1”大数据产品体系，提供大数据体系建设咨询规划能力、治理能力、平台能力、运营能力和基础网络五大服务力。同时在金融、广告、交通等行业市场实现业务突破，形成面向金融、数字营销、旅游、房地产等十大行业解决方案，实现了顺丰、招行等标杆客户突破，并以大数据业务为切入点带动传统基础通信服务的发展。



中国电信股份有限公司云计算分公司

天翼大数据官方微信：天翼大数据