



**中国支付清算协会**

Payment & Clearing Association of China

# 大数据在金融领域的典型应用研究

中国支付清算协会金融大数据应用研究组  
中国信息通信研究院云计算与大数据研究所

2018 年 3 月

---

## 版权声明

---

本报告版权属于中国支付清算协会金融大数据应用研究组，并受法律保护。转载、摘编或利用其它方式使用本白皮书文字或者观点的，应注明“来源：中国支付清算协会金融大数据应用研究组”。违反上述声明者，本单位会将追究其相关法律责任。

编写指导：王素珍、何宝宏

编写小组：丁华明、赵计博、韩涵、赵博、何阳

# 前 言

近年来，我国金融科技快速发展，在多个领域已经走在世界前列。大数据、人工智能、云计算、移动互联网等技术与金融业务深度融合，大大推动了我国金融业转型升级，助力金融更好地服务实体经济，有效促进了普惠金融发展。在这一发展过程中，又以大数据技术发展最为成熟、应用最为广泛。从发展特点和趋势来看，金融云快速建设落地奠定了金融大数据的应用基础，金融数据与其他跨领域数据的融合应用不断强化，人工智能正在成为金融大数据应用的新方向，金融行业数据的整合、共享和开放正在成为趋势，给金融行业带来了新的发展机遇和巨大的发展动力。

为促进大数据技术在金融领域的创新和安全应用，中国支付清算协会在金融科技专业委员会的基础上，成立了金融大数据应用研究组，依托金融科技专业委员会开展相关研究验证和推广交流活动，充分发挥行业协会贴近市场和研究机构的优势，深入研究金融大数据应用理论和实践问题。研究组成立以来，在组长单位中国信息通信研究院云计算与大数据研究所的带领下，在广大成员单位的支持和配合下，积极开展市场调研，努力搭建交流平台，探索行业标准建设，开展了许多富有成效的工作，取得了积极成果。

其中一项重要工作就是面向成员单位征集金融大数据

创新应用案例，开展重点课题研究。该项工作自 2017 年 11 月启动以来，得到了广大成员单位的积极响应，共征集到 40 多个有效案例。经专家严格评审，最终有 24 个案例获评“金融大数据创新应用优秀成果奖”。研究组进一步整合获奖案例内容，结合课题研究成果撰写完成《大数据在金融领域的典型应用研究》白皮书并公开发布，借此机会希望促进行业交流和优秀成果的学习借鉴，为推动金融大数据更好发展应用贡献一份智慧和力量。

## 大数据在金融领域的典型应用研究

数据是数字经济时代的新型生产资料，基于数据的生产变革和业务模式创新正驱动着全球范围内经济社会各个领域的数字化、智能化转型，发展大数据已经成为国家战略。十九大报告明确指出，要推动“互联网、大数据、人工智能和实体经济深度融合。”

金融行业数据资源丰富，数据应用由来已久。从发展特点和趋势来看，金融云快速建设落地奠定了金融大数据的应用基础，金融数据与其他跨领域数据的融合应用不断强化，人工智能正在成为金融大数据应用的新方向，金融行业数据的整合、共享和开放正在成为趋势。

随着大数据技术的广泛普及和发展成熟，金融大数据应用已经成为行业热点趋势，在交易欺诈识别、精准营销、黑产防范、消费信贷、信贷风险评估、供应链金融、股市行情预测、股价预测、智能投顾、骗保识别、风险定价等涉及银行、证券、保险、支付清算和互联网金融等多领域的具体业务中，得到广泛应用。涌现出一大批技术创新、业务突破的应用案例。总结来看，对于大数据的应用分析能力，正在成为金融机构未来发展的核心竞争要素。

毋庸置疑，金融大数据拥有着广阔的发展前景。然而，金融大数据应用也面临着数据资产管理水平不足、技术改造难度大、行业标准缺失、安全管控压力大和政策保障仍不完善等一系列制约因素。为推动金融大数据更好发展应用，必须从政策扶持保障、数据管理能力提升、行业标准规范建设和应用合作创新等多个方面入手，不断强化应用基础能力，持续完善产业生态环境。

## **一、大数据的兴起与金融大数据的发展**

### **（一）大数据的兴起**

近年来，摄像头、可穿戴设备、GPS 等传感器收集着大量音频、视频、图像等各类结构化和非结构化数据，随着电子商务、社交、综合信息网站等互联网应用的发展，数据基于网络大量产生并存储，信息量爆发式增长。据 IDC 的研究显示，全球数据总量年复合增长率 50%。这种增长速度意味着未来两年，全球新增的数据量将超过人类有史以来积累的数据总和。IDC 预测，到 2020 年，全球数据总量将达到 40ZB（400 亿 TB），代表地球上每个人平均会产生 5TB 的数据。

人类正从 IT 时代走向 DT 时代，数据是数字化时代的“石油”，大数据就是数字化时代的“冶炼工艺”。通过数据的收集、存储、分析和可视化技术，解决大数据海量、高速、多变、低密度的问题，使数据从散乱的信息，变成知识和智慧，帮助组织解决发展中遇到的实际问题。

麦肯锡公司早在 2011 年就已经预见到大数据时代的到来，并提出：“各个行业和领域都已经被数据渗透了，目前数据已成为非常重要的生产因素。对于大数据的处理和挖掘将意味着新一波的生产率不断增长和消费者盈余浪潮的到来。”人们已经意识到，通过数据给社会创造价值的能力和用数据盈利的能力将成为所有组织的核心竞争力。

纵观金融行业的发展史，每次都是科技的创新推动着金融行业的发展与变革。电报技术、互联网技术的推出都对金融机构的服务模式和风控方式产生了重大影响。近几年，各国政府不断加大对科技创新的重视程度。科技创新的速度不断加快，并逐步与金融业务深度融合，以大数据、云计算、人工智能、区块链等为代表的新技术已逐渐成为金融发展的新动力。

普华永道调研显示，在所有金融科技中，大数据是金融行业投资和应用的的首选。首先，从内在需求看，在互联网金融模式的冲击下，整个金融业的运作模式正在重构，行业竞争日益激烈，基于数据的精细化运营需求日益迫切。其次，从应用基础上看，金融行业拥有海量数据资源。金融业是最有意愿进行信息化投入的行业之一，经过多年的信息沉淀，各系统内积累了大量高价值的数据，拥有用于数据分析的基础资源。最后，从产品供应上看，大数据产品已经越来越成熟，技术供给越来越丰富，部署成本直线下降。此外，部分先行者为大数据部署提供了宝贵的应用案例，使得金融大数

据解决方案越趋完善。

## （二）金融大数据的产生

金融数据产生的主体有三种：“人”、“机”、“物”。“人”指的是人类活动的数据，它是人类在活动过程所产生的各类数据，包括评论、通话记录、照片、网页浏览痕迹、交易记录等信息。“机”指的是信息系统产生的数据，这些信息主要以文件、多媒体等形式存在，包括审计、日志这样自动生成的信息。“物”指的是物理世界产生的数据，是通过摄像头、传感器等数字设备在监测中采集的数据，例如服务器运行监控数据、押运车监控数据等。

## （三）金融机构数据获取方式

金融机构有三种数据获取的方式：在自有系统中沉淀、在网上采集和从第三方购买。

**1. 在自有系统中沉淀数据。**金融机构通常会部署数百个应用系统，这些系统在日常经营中持续产生和储存数据，经过长期的数字化运营积累，数据的规模已经较为庞大。以银行业为例，目前中国单家股份制商业银行累积的数据已经达到上百 TB。波士顿咨询的调研表示，银行业每创收 100 万美元，平均就会产生 820GB 的数据。

**2. 在网上采集数据。**金融机构在网上主要采集企业的舆情数据和个人的行动数据。企业舆情数据包括两大方面：一是政府公开数据：工商、司法、行政和一行三会的处罚/涉诉数据等；二是企业经营动态数据：资产重组、投融资、高



管变动、员工招聘、新产品发布和产品销售情况等。个人行动数据也包括两大方面：一是基本属性数据：性别、年龄、学历、职业等；二是喜好数据：浏览页面、浏览商品、页面停留时间、关注的商品、支付的商品、产品评分、产品投诉、产品建议、加入的社群、经常互动的话题等。

**3. 从第三方购买数据。**金融机构购买的主要是企业的数  
据，在个人数据购买上比较谨慎。在个人数据交易上，大型  
金融机构普遍认为比较敏感，要谨慎对待。目前，大型金融  
机构从第三方购买个人数据的行动基本停滞，主要从政府侧  
购买公共数据，例如公积金、社保和税务数据等。

#### **（四）金融大数据的发展特点**

**1. 金融云快速落地奠定大数据应用基础。**金融云具备的  
快速交付、高扩展、低运维成本等特性，能够在充分考虑金  
融机构对信息安全、监管合规、数据隔离和中立性等要求的  
情况下，为机构处理突发业务需求、部署业务快速上线、实  
现业务创新改革提供有力支持。因此，金融业一直较为积极  
的推动云计算的落地。

目前，大型金融机构纷纷开启了基于云计算的信息系统  
架构转型之路，逐步将业务向云迁移。大型金融机构普遍青  
睐混合云架构，将非核心应用迁移到公有云上，再将部分核  
心应用迁移到私有云平台上，关键业务上继续使用传统架构。  
新兴金融机构如蚂蚁金服、微众银行等在诞生之初就把所有  
IT 系统架构在云上。

## **2. 实时计算分析能力是金融大数据应用的首要关注点。**

金融机构的业务要求大数据平台具有实时计算的能力。目前，金融机构最常使用的大数据应用场景为精准营销、实时风控、交易预警和反欺诈等业务都需要实时计算的支撑。

以精准营销和交易预警为例，精准营销要求在客户短暂的访问与咨询时间内发现客户的投资倾向，推荐适合的产品。交易预警场景要求大数据平台在秒级完成从事件发生、到感知变化、到输出计算结果的整个过程，识别出客户行为的异常，并做出交易预警。因此，流式计算框架的实时计算大数据平台目前逐渐在金融机构得到应用，以满足低延时的复杂应用场景需求。

## **3. 金融业务创新越来越依赖于大数据应用分析能力。**

客户对服务体验的要求越来越高，需要金融机构随时随地都能提供服务，产品设计的更易用、更直观，响应速度更快速。金融机构提供产品和服务的重点，也从简单的标准化，转变为个性化。

大数据能够在产品设计和客户服务两方面提高创新能力。在产品设计上，大数据能够更好的利用现有数据，为客户进行全面的客户画像，识别客户的需求。基于精准的客户认知，金融机构可以细分客户的需求，从而针对性的设计出符合客户个性化需求的、场景化的产品。在客户服务上，大数据可以提高产品的自动化程度，从而扩大产品和服务的范围、拓宽客户基础，使得金融机构得以覆盖以前服务不到的

长尾客户。此外，产品自动化还能够快速的对客户需求做出反应，提高客户粘性。

**4. 金融数据正在向金融科技行业巨头聚集。**互联网和科技行业存在的“赢家通吃”模式，在金融行业继续上演。随着行业的快速整合，原来分散在各家金融机构的数据正快速向金融科技行业巨头集中，从而形成数据寡头。

以支付行业为例，原来分散在各家银行手中的支付数据正快速向支付宝和财付通集中。目前，支付宝和财付通已经覆盖了绝大多数消费场景，包括电商购物、餐饮、出行、航旅、公共事业缴费、线下购物等几乎所有消费场景。过去银行可以通过借记卡和信用卡的消费记录来分析客户的消费行为，为金融企业的服务和产品设计提供支持。现在这些小额消费行为很多都通过第三方支付发生，银行无法拿到具体的消费数据。客户消费数据的缺少，正在影响银行对个人客户的了解和分析。

### **（五）金融大数据的发展趋势**

**1. 大数据应用水平正在成为金融企业竞争力的核心要素。**金融的核心就是风控，风控以数据为导向。金融机构的风控水平直接影响坏账率、营收和利润。经过长期的数字化改造，金融机构积累了大量的信息系统，通过这些系统积累了海量的数据，但是这些数据是分散在各个系统中，不能实现集中分析。金融机构已经意识到需要有效地管理其日益重要的数据资产，正在主动思考和实践数据资产治理的方法。

目前，金融机构正在加大在数据治理项目中的投入，结合大数据平台建设项目，构建企业内统一的数据池，实现数据的“穿透式”管理。大数据时代，数据治理是金融机构需要深入思考的命题，有效的数据资产管控，可以使数据资产成为金融机构的核心竞争力。

在国内，金融机构对大数据的认知已经从探索阶段进入到认同阶段。普华永道研究显示，83%的中国金融机构表示希望在大数据上进行投资。金融行业对大数据的需求属于业务驱动型。其迫切希望应用大数据技术使营销更精准、风险识别更准确、经营决策更具针对性、产品更具吸引力，从而降低企业成本，提高企业利润。随着更多金融机构基于大数据获得丰厚的回报，将进一步打消它们的顾虑，加速大数据的普及。

**2. 金融行业数据整合、共享和开放成为趋势。**数据越关联越有价值，越开放越有价值。随着各国政府和企业逐渐认识到数据共享带来的社会效益和商业价值，全球已经掀起一股数据开放的热潮。大数据的发展需要所有组织和个人的共同协作，将个人私有、企业自有、政府自有的数据进行整合，把私有大数据变为公共大数据。

目前，美欧等发达国家和地区的政府都在数据共享上做出了表率，开放大量的公共事业数据。中国政府也着力推动数据开放。一方面，国家带头着力推动政府数据公开。国务院《促进大数据发展行动纲要》提出：到 2018 年，中央政

府层面实现金税、金关、金财、金审、金盾、金宏、金保、金土、金农、金水、金质等信息系统通过统一平台进行数据共享和交换。另一方面，国家还通过推动建设各类大数据服务交易平台，为数据使用者提供更丰富的数据来源。在发改委发布的《国家发展和改革委员会办公厅关于请组织申报大数据领域创新能力建设专项通知》中明确提到要建设大数据流通与交易平台，用以支撑数据共享。

### **3. 金融数据与其他跨领域数据的融合应用不断强化。**

2015 年以前，金融机构主要基于金融业自有信息进行分析。金融机构主要基于自身静态数据通过人工对内进行经营分析、产品设计、营销设计等；对外进行客户分析和行情分析。从 2016 年开始，大数据技术逐渐成熟，数据采集技术快速发展，通过图像识别、语音识别、语义理解等技术实现外部海量高价值数据收集，包括政府公开数据、企业官网数据、社交数据。金融机构得以通过客户动态数据的获取更深入的了解客户。

未来，数据流通的市场会更健全。金融机构将可以方便的获取电信、电商、医疗、出行、教育等其他行业的数据，一方面会有力的促进金融数据和其他行业数据融合，使得金融机构的营销和风控模型更精准。另一方面，跨行业数据融合会催生出跨行业的应用，使金融行业得以设计出更多的基于场景的金融产品，与其他行业进行更深入的融合。

### **4. 人工智能正在成为金融大数据应用的新方向。新兴技**

术高速发展，大数据和人工智能技术正在快速融合。大数据技术强调数据的采集、存储、处理和展现。人工智能可以在各个阶段助力大数据发挥更大的作用。

在采集上，图像识别、语音识别、语义理解等人工智能认知技术实现海量非结构化数据采集。在数据的储存和管理上，人工智能技术可以实现自动为数据打标签，自动将数据归类。在数据处理上，人工智能深度学习、机器学习、知识图谱技术可以提高算法模型的数据处理的效率和准确度。数据展现上，智能可视化大屏技术可以实现数据实时监控和可视化呈现。大数据与人工智能正在进行多维度的深度融合，拓展了金融大数据的应用价值和应用场景。

**5. 金融数据安全问题越来越受到重视。**大数据的应用为数据安全带来新的风险。数据具有高价值、无限复制、可流动等特性，这些特性为数据安全带来了新的挑战。

对金融机构来说，网络恶意攻击成倍增长，组织数据被窃的事件层出不穷。这对金融机构的数据安全管理能力提出了更高的要求。大数据使得金融机构内海量的高价值数据得到集中，并使数据实现高速存取。但是，如果出现信息泄露可能一次性泄露组织内近乎全部的数据资产。数据泄露后还可能急速扩散，甚至出现更加严重的数据篡改和智能欺诈的情况。

对个人来说，金融信息的泄露会暴露出大量的个人基本信息和消费信息等，大数据技术可以便捷地大批量收集这些

信息并进行画像，这使得公民更容易受到欺诈，造成经济损失。

## **二、金融大数据的应用场景分析**

大数据技术的应用提升了金融行业的资源配置效率，强化了风险管控能力，有效促进了金融业务的创新发展。金融大数据在银行业、证券行业、保险行业、支付清算行业和互联网金融行业都得到广泛的应用。

### **（一）大数据在银行业中的应用**

**1. 信贷风险评估。**在传统方法中，银行对企业客户的违约风险评估多是基于过往的信贷数据和交易数据等静态数据，这种方式的最大弊端就是缺少前瞻性。因为影响企业违约的重要因素并不仅仅只是企业历史的信用情况，还包括行业的整体发展状况和实时的经营情况。而大数据手段的介入使信贷风险评估更趋近于事实。

内外部数据资源整合是大数据信贷风险评估的前提。一般来说，商业银行在识别客户需求、估算客户价值、判断客户优劣、预测客户违约可能的过程中，既需要借助银行内部已掌握的客户相关信息，也需要借助外部机构掌握的人行征信信息、客户公共评价信息、商务经营信息、收支消费信息、社会关联信息等。该部分策略主要目标为数据分析提供更广阔的数据维度和数据鲜活度，从而共同形成商业银行贷款风险评估资源。

### 信贷风险评估的步骤：

(1) 以客户级大数据为基础，为存量客户建立画像，使银行能够向各管辖机构、各业务条线、各产品条线进行内容全面、形式友好、敏捷的客户级大数据集中供给。

(2) 建立专项集中的企业及个人风险名单库，统一“风险客户”等级标准，集中支持各专业条线、各金融产品对高风险客户的过滤工作。

(3) 统筹各专业条线、各业务环节对大数据增量信息的需求优先序列，对新客户、高等级客户、高时效业务、高风险业务实现大数据实时采集式更新；对存量、一般、普通时效业务、低风险业务实现大数据集中、批量、排序、滚动更新。

**2. 供应链金融。**供应链金融的风险控制从授信主体向整个链条转变。供应链核心企业拥有良好的资产、充足的资金和高额的授信额度。而依附于核心企业的上下游企业可能需要资金，但是贷不到款。供应链金融可以由核心企业做担保，以产品或应收账款做质押，帮助上下游企业获得资金。

利用大数据技术，银行可以根据企业之间的投资、控股、借贷、担保以及股东和法人之间的关系，形成企业之间的关系图谱，利于关联企业分析及风险控制。知识图谱在通过建立数据之间的关联链接，将碎片化的数据有机的组织起来，让数据更加容易被人和机器理解和处理，并为搜索、挖掘、分析等提供便利。



在风控上，银行以核心企业为切入点，将供应链上的多个关键企业作为一个整体。利用交往圈分析模型，持续观察企业间的通信交往数据变化情况，通过与基线数据的对比来洞察异常的交往动态，评估供应链的健康度及为企业贷后风控提供参考依据。

## （二）大数据在证券行业中的应用

**1. 股市行情预测。**大数据可以有效拓宽证券企业量化投资数据维度，帮助企业更精准的了解市场行情。随着大数据广泛应用、数据规模爆发式增长以及数据分析及处理能力显著提升，量化投资将获取更广阔的数据资源，构建更多元的量化因子，投研模型更加完善。

证券企业应用大数据对海量个人投资者样本进行持续性跟踪监测，对账本投资收益率、持仓率、资金流动情况等一系列指标进行统计、加权汇总，了解个人投资者交易行为的变化、投资信心的状态与发展趋势、对市场的预期以及当前的风险偏好等，对市场行情进行预测。

**2. 股价预测。**证券行业具有自身的特点，与其他行业产品与服务的价值衡量普遍存在间接性的特点不同，证券行业客户的投资与收益以直接的、客观的货币形式直观的呈现。受证券行业自身特点和行业监管要求的限制，证券行业金融业务与产品的设计、营销与销售方式也与其他行业具有鲜明的差异，专业性更强。

诺贝尔经济学奖得主罗伯特•席勒设计的投资模型至今

仍被业内沿用。在他的模型中，主要参考三个变量：投资项目计划的现金流、公司资本的估算成本、股票市场对投资的反应（市场情绪）。他认为，市场本身带有主观判断因素，投资者情绪会影响投资行为，而投资行为直接影响资产价格。然而，在大数据技术诞生之前，市场情绪始终无法进行量化。大数据技术可以收集并分析社交网络如微博、朋友圈、专业论坛等渠道上的结构化和非结构化数据，了解市场对特定企业的观感，使得市场情绪感知成为可能。

**3. 智能投顾。**智能投顾是近年证券公司应用大数据技术匹配客户多样化需求的新尝试之一，目前已经成为财富管理新蓝海。智能投顾业务提供线上的投资顾问服务，能够基于客户的风险偏好、交易行为等个性化数据，采用量化模型，为客户提供低门槛、低费率的个性化财富管理方案。智能投顾在客户资料收集分析、投资方案的制定、执行以及后续的维护等步骤上均采用智能系统自动化完成，且具有低门槛、低费率等特点，因此能够为更多的零售客户提供定制化服务。随着线上投顾服务的成熟以及未来更多基于大数据技术的智能投资策略的应用，智能投顾有望从广度和深度上都将证券行业带入财富管理的全新阶段，为未来政策放宽证券公司投资顾问从前端佣金收费向后端的管理费收取模式转变进行探索准备。

### （三）大数据在保险行业中的应用

**1. 骗保识别。**赔付直接影响保险企业的利润，对于赔付

的管理一直是险企的关注点。而赔付中的“异常值”（即超大额赔付）是推高赔付成本的主要驱动因素之一。保险欺诈严重损害了保险公司的利益，为了识别可疑保险欺诈案件，需要展开大量专项调查，但往往需要耗费数月或数年的时间。

借助大数据手段，保险企业可以识别诈骗规律，显著提升骗保识别的准确性与及时性。保险企业可以通过建设保险欺诈识别模型，大规模的识别近年来发生的所有赔付事件。通过筛选从数万条赔付信息中挑出疑似诈骗索赔。保险企业再根据疑似诈骗索赔展开调查会有效提高工作效率。此外，保险企业可以结合内部、第三方和社交媒体数据进行早期异常值检测，包括了客户的健康状况、财产状况、理赔记录等，及时采取干预措施，减少先期赔付。

**2. 风险定价。**保险企业对保费的定义是基于对一个群体的风险判断，对于高风险的群体收取较高的费用，对于低风险群体则降低费用。通过灵活的定价模式可以有效提高客户的粘性。而大数据为这样的风险判断带来了前所未有的创新。

保险公司通过大数据分析可以解决现有的风险管理问题。比如，通过智能监控装置搜集驾驶者的行车数据，如行车频率、行车速度、急刹车和急加速频率等；通过社交媒体搜集驾驶者的行为数据，如：在网上吵架频率、性格情况等；通过医疗系统搜集驾驶者的健康数据。以这些数据为出发点，如果一个人不经常开车，并且开车十分谨慎的话，那么他可以比大部分人节省 30%~40%的保费，这将大大的提高保险产

品的竞争力。

（四）大数据在支付清算行业中的应用

**交易欺诈识别。**目前，支付服务操作十分便捷，客户已经可以做到随时、随地进行转账操作。面对盗刷和金融诈骗案件频发的现状，支付清算企业交易诈骗识别挑战巨大。

大数据可以利用账户基本信息、交易历史、位置历史、历史行为模式、正在发生行为模式等，结合智能规则引擎进行实时的交易反欺诈分析。整个技术实现流程为实时采集行为日志、实时计算行为特征、实时判断欺诈等级、实时触发风控决策、案件归并形成闭环。

欺诈方式	欺诈场景	处理方式
盗刷	客户账号于多个手机上登录	身份再验证
	客户从一个不经常出现的地区进行大额转账操作	
	在1小时内于不同的城市进行消费	提醒客户
诈骗	发现多个客户在短时间内向单个账户转账	
	发现钓鱼网站	

图 1 交易欺诈方式及场景示例

（五）大数据在互联网金融行业中的应用

**1. 精准营销。**在移动互联网时代，客户在消费需求和消费行为上快速转变。首先，在消费需求上，客户的需求更加细化，急需个性化的金融产品。另外，在消费行为上，互联网金融企业很难接触到消费者及了解客户的需求并推销产

品，营销资源和营销机会极其宝贵。因此，为了降低对用户打扰和营销成本，提高营销转化率以面对日趋激烈的行业内部竞争，互联网金融企业急需一种更为精准的营销解决方案。

具体来讲，银行业精准营销的主要应用目标主要为三点：一是精准营销的首要目的是寻找目标客户，精准定位营销对象。二是在获得客户挖掘结果后，精准营销应用的下一步功能应当是针对具体客户，提供一整套智能决策方案。三是配备完整的业务操作平台，实现整个精准营销从客户挖掘直至业务完成的各工作环节，最大化程度缩短业务操作流程，实现精准营销的“一站式”操作。

利用大数据平台的模型分析结果，挖掘出潜在客户，实现可持续的营销计划。银行业精准营销的技术流程包括以下五点：

（1）客户信息整合及验证。该步骤的核心为数据整合处理，即利用大数据平台打通内外部数据、不同业务数据、不同结构数据之间的壁垒，对数据格式进行规范化处理，形成以客户为中心的“一户一条”数据记录。

（2）客户及场景标签设定。根据精准营销的不同角度设定不同类别的场景标签，该类标签较用户标签通常具有更高的灵活性，以便随着业务发展和精准营销场景的变换随时增减或改变。

（3）客户类型初分。对于不同类型的目标客户，精准营销模型应当给出具有针对性的营销方案。因此需要对大数

据平台里的所有客户进行类型的区分。该部分可以通过用户画像技术提供的分类标准，进行用户的标签化分类工作。建立合理的客户类型初分体系是精准营销的基础。

（4）客户筛选。客户筛选是指对客户质量进行筛选把控，普遍基于大数据平台的黑白灰名单技术，对客户进行判定。原则上白名单客户属于精准营销判定的推荐客户；黑名单客户则是不能服务的客户；灰名单客户为风险提示类的客户。

（5）业务统一工作平台部署。作为精准营销的前台，通过 API 接口打通数据存储层、数据处理层、算法层以及高级业务层，通过门户网站、APP、APP 接口等方式提供银行产品推荐、客户准入、客户跟踪管理等高级营销策略。

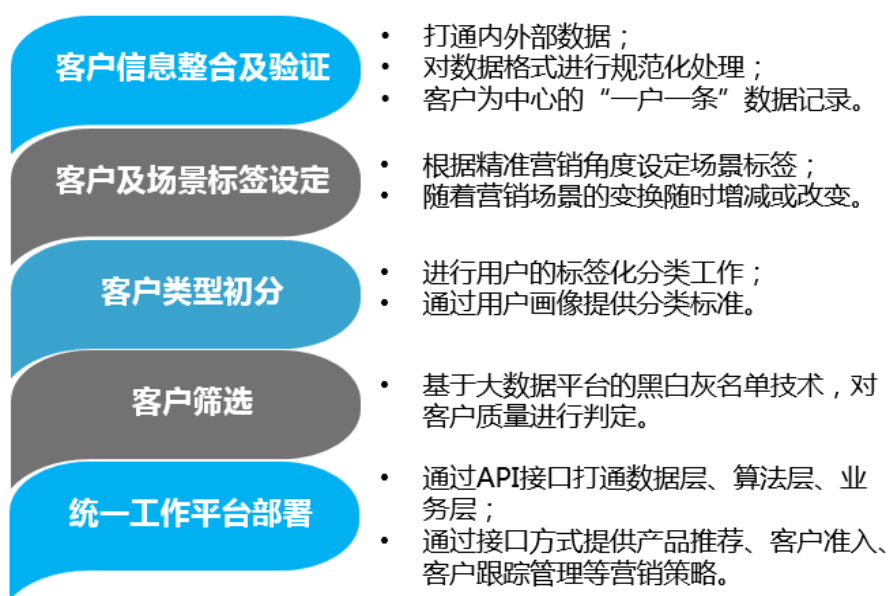


图 2 精准营销的技术流程

2. 黑产防范。互联网金融企业追求服务体验，强调便捷

高效，简化手续。而这一特点也易被不法分子利用，虚假注册、利用网络购买的身份信息与银行卡进行套现，“多头借贷”乃至开发电脑程序骗取贷款等已经形成了一条“黑色”产业链，对于互联网金融行业而言，欺诈风险高于信用风险。

大数据能够帮助企业掌握互联网金融黑产的行为特点，从业人员规模、团伙地域化分布以及专业化工具等情况，并制定针对性的策略。黑产特征如下：借款手机归属地与真实城市 IP 不匹配；设备上相邻两次借款（含跨平台）时间间隔极短；用户手机长期处于同一位置未移动过等。通过黑产识别和预警减少损失。

**3. 消费信贷。**消费信贷和传统企业信贷截然不同。它拥有小额、分散、高频、无抵押和利息跨度极大的特点。在贷款额度上可以小到 100 元人民币；一家机构一天放贷数量可能达到数万到数十万笔；90%以上是纯信用贷，只能依靠数据进行审批；年化利率从 4%到 500%的都有。

客户特点是年轻、消费观念超前、无信用记录。消费信贷客户大多数年龄都在 35 岁以下；月光族，愿意透支未来。某些人群甚至对借钱消费形成习惯性依赖；80%-90%的客群无人行借贷信用记录，导致拒绝率极高。

大数据需要贯穿到客户全生命周期的始末。基于大数据的自动评分模型、自动审批系统和催收系统是消费信贷的基础。利用大量行为数据分析弥补信贷数据的缺失。下面会列举一些趋势上的分析方法：1) 随着手机号使用年数的增加，

客户稳定性增加，违约风险逐步降低。2) 过去 12 个月内所有类日本地生活消费等级越高，违约风险越低。3) 最近 12 个月网络游戏消费金额越多，违约风险越高。4) 最近 12 个月内财经媒体访问天数越多，违约率风险越低。

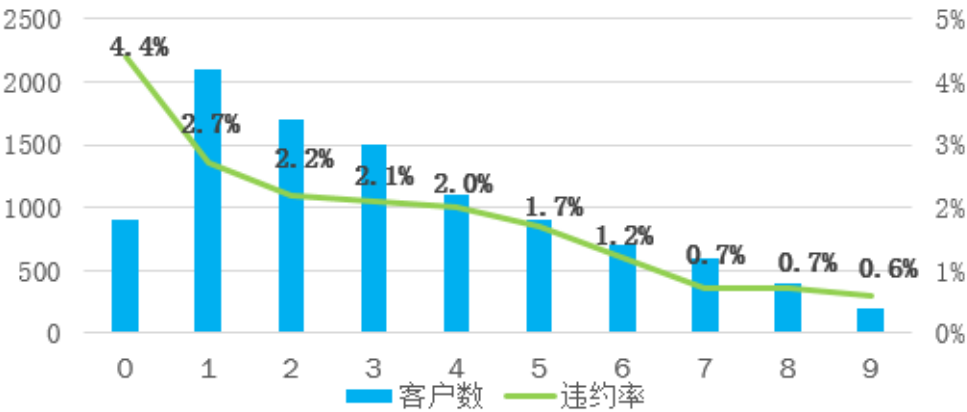


图 3 手机号使用年数与违约率

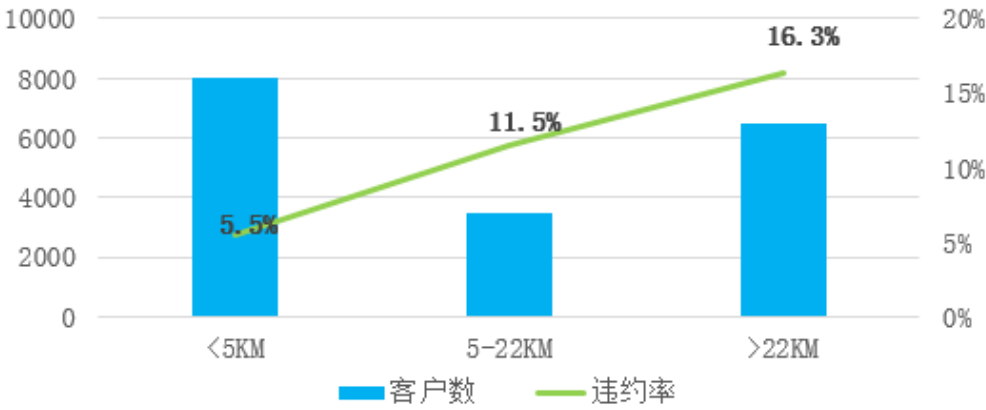


图 4 申请家庭住址距离差与违约率

三、金融大数据的典型案例分析

2017 年中国支付清算协会金融大数据应用研究组面向协会金融科技专业委员会成员单位征集大数据应用案例，共



征集到 40 多个案例。经专家组评审，从中评选出 24 个“金融大数据创新应用优秀案例”，进行深入分析。

### **（一）风险控制类应用**

#### **1. 中国银行“艾达”大数据风控平台**

**（1）项目背景。**近年来国内外金融形势愈加复杂，GDP 增速放缓，企业业绩下滑，负债率持续攀升。金融持续脱媒，企业经营呈“跨业、跨界、跨境”态势，对银行信贷依赖度降低。集团客户关联关系复杂，更加分散化、隐蔽化、多元化，越来越多的资金从实体经济转向虚拟经济，风险蔓延速度迅猛，并呈现“跨渠道、跨地域、跨产品”传播，与此同时，银行对资金流向的监控手段却非常有限。在新的经济环境下，原来被动的风险防控方式已经难以满足新常态下客户高效性和多样性的需求。

**（2）应用场景。**应用于企业用户画像、风险管理、精准营销等场景。

**（3）成果概述。**“艾达”大数据智能风控平台是面向全行前、中、后台业务人员，包括客户经理、风险经理、审计经理和管理层的大数据风控平台。通过对结构化、非结构化的整合，运用大数据、AI 等新技术重塑业务流程与风险管理模式，不断挖掘数据价值。将大数据应用作为提升风险管理能力的关键工具和重要途径，也是中国银行首次尝试用大数据建模进行风控管理。

**（4）项目收益。**打通了数据孤岛，挖掘并提升行内存

量数据价值。解放生产力、发展生产力，提升企业运营效率，“艾达”嵌入业务环节，节约的时间保守估计在 20 个工作日以上，特别是对于突发事件、隐藏风险知晓的及时性，可有效规避损失。企业实时预警监控，降低授信风险，挽回资金损失。“艾达”上线至今，触发风险预警标签 4303088 个，成本中心向利润中心转变。提升全行大数据应用能力，打造数据生态圈。

## 2. 中国交通银行信用卡中心电子渠道实时反欺诈监控交易系统

**(1) 项目背景。**交通银行需要实时接收电子渠道交易数据，整合系统其他业务数据，通过规则实现快速建模、实时告警与在线智能监控报表等功能。总体要求能实时接收官网业务数据，整合客户信息、设备画像、位置信息、官网交易日志、浏览记录等，通过规则实现快速建模、实时告警与在线智能监控等功能。

**(2) 应用场景。**应用于风险预警场景。

**(3) 成果概述。**明略数据通过为交通银行卡中心构建反作弊模型、实时计算、实时决策系统，帮助拥有数十 TB 历史数据，日均增逾两千万条日志流水的国有银行卡中心建立电子渠道实时反欺诈交易监控系统。利用分布式实时数据采集技术和实时决策引擎，帮助信用卡中心高效整合多系统业务数据，处理海量高并发线上行为数据，识别恶意用户和欺诈行为，并实时预警和处置，通过引入机器学习框架，对

海量数据进行分析、挖掘构建并周期性更新反欺诈规则和反欺诈模型。

**（4）项目收益。**明略数据反欺诈系统上线后，运转稳定、高效，迅速监控电子渠道产生的虚假账号、伪装账号、异常登录、频繁登录等新型风险和欺诈行为：系统 7\*24 小时稳定运行，日均处理逾两千万条日志流水、实时识别出近万笔风险行为并进行预警下发。数据接入、计算报警、案件调查的整体处理时间从数小时降低至秒级，监测时效提升近 3000 倍，上线 3 个月已帮助卡中心挽回数百万元的风险损失。

### **3. 光大银行大数据风控智能化数据产品：滤镜**

**（1）项目背景。**以大数据、移动互联网为代表的新型技术正在推动实体经济变革与创新，银行的风险管理决策方式也逐步转向智能化和移动化。近年来，光大银行科技创新机制催化大量创新项目支持行内业务智能化转型，其中大数据实验室在风险管理领域积极开展风险预警分析研究，成功孵化数据产品“滤镜”。该产品利用大数据技术对企业客户进行过滤，提示高信用违约倾向的企业名单，向总分行风险管理决策者提供更加科学精准的决策支持，同时依托首个移动数据应用平台“光速观察”提升风险决策的时效性。

**（2）应用场景。**应用于风险预警、风险分析、风险排查等场景。

**（3）成果概述。**滤镜数据产品依托光大银行首个移动应用平台“光速观察”，运用多项大数据分析技术，构建大

数据风险预警信号，采用名单式管理模式，向总分行风险管理决策者提示具有高信用违约风险的企业信息，有效提升风险决策时效性和精准度。

滤镜将社交网络分析技术充分应用到风险预警监控，综合运用图算法进行客户风险预测实现客户、群体、网络评分，精确锁定潜在风险客户。该数据产品对隐匿在客户行为和客户交往圈中的资金短缺、风险传播、群体违约等风险事件特征化，揭示客户潜在的信用违约风险。

**（4）项目收益。**“滤镜”自2016年5月底开始上线试运行，通过后评价分析，大数据预警信号过滤的企业在预警后6个月发生违约的平均概率27%，比传统基于专家规则模型有明显提升。滤镜目前以月度为单位产生预警客户名单，月均预警客户数约300个左右，对应预警客户授信余额达300亿，可以据此估算可能减少的损失金额可达数十亿元。

#### **4. 恒丰银行全面风险预警系统**

**（1）项目背景。**恒丰银行近年来陆续推出了信贷工厂、消费金融、供应链金融等一系列网贷、平台贷业务，为不同行业、不同规模的客户提供了丰富的信贷类产品。业务规模快速发展的同时，如何快速、全面识别、监测、防范客户信用风险，成为了全行风险管理领域最为重要的工作之一。对此，恒丰银行提出通过运用大数据技术构建信用风险预警系统，加强风险信息归集、监测、审查的准确性、及时性，强化风险预测能力。

**（2）应用场景。**应用于贷前、贷中、贷后风险检查、监测、预测和缓释控制等场景。

**（3）成果概述。**全面风险预警系统依托于星环大数据基础平台，整合行内外数据，包括行内授用信、不良、逾期、征信，行外司法、舆情、风险信号、工商、关联关系等，以多种方式为用户提供客户的风险提示和风险发现功能，为及时处置风险争取时间。数据主要来自于行内交易、业务数据沉淀、外部数据采购等方式接入，结构化、非结构化数据共治，整合了共计数千万家企业的相关信息，通过批处理和数据加工处理，能够在 1s 之内为用户提供需要的服务。

**（4）项目收益。**恒丰银行信用卡风险预警系统自投产上线以来，风控能力逐步提升，在客户风险识别效率、准确率、成本控制等方面较传统风控手段有了大幅提高。新增信贷资产质量大幅提升。以某平台贷为例，自风控系统启用以来，其新增授信业务逾欠率控制在 1%以内，且呈逐渐降低态势；新增的网贷、平台贷授信业务发放效率显著提升。基于大数据风控技术的航信票贷、恒信快贷等业务产品实现了 24 小时、8 小时放款；新增业务的客户贷前调查成本大幅降低。经对某平台贷的数据统计，风险预审过程可综合节约近 80% 的人力成本，同时基于该统计数据调整的业务发展规划更为科学、符合实际。

## **5. 腾讯云‘天御’大数据反欺诈平台**

**（1）项目背景。**在金融领域互联网化的背景下，金融

机构特别是互联网金融企业，更加追求便捷高效，简化手续，强调服务体验，而这一特点也易被不法分子利用，虚假注册、利用网络购买的身份信息与银行卡进行套现，“羊毛党”通过低成本甚至零成本取得互联网平台奖励，“多头借贷”乃至开发电脑程序骗取贷款等等已经形成了一条“黑色”产业链，互联网金融行业面临着严禁的挑战，对于互联网金融行业而言，欺诈风险高于信用风险。

**（2）应用场景。**应用于互联网金融行业贷前审核风控及贷后监控、支付行业防盗刷、互联网行业线上营销风控、网站及 App 安全风险防控等场景。

**（3）成果概述。**腾讯云“天御”大数据反欺诈平台（AF）是腾讯首次在云端输出反诈骗技术能力，依托 19 年安全积累、亿级体量的黑产数据，腾讯从计算力、算法、数据等三方面能力为反诈骗 AI 创新提供条件。腾讯云反欺诈产品包含有反薅羊毛、反骗贷、反洗钱、反骗保（保险）、移动银行 APP 保护、防盗刷等众多应用程序接口（API），无需改动企业 IT 系统。“天御”系统数据来源包括支付画像、群组画像、社交画像、设备画像、行为画像等几大类别。主要应用于银行、证券、保险、P2P 等行业客户，准确识别恶意用户与行为，解决客户在支付、借贷、理财、风控等业务环节遇到的欺诈威胁，降低企业的损失。

**（4）项目收益。**贷前审核与贷后监控方面，微众银行微粒贷产品逾期率低于 0.3%。活动防刷方面，一是注册环节，

识别虚假注册。二是在登录场景，登录环节通过验证码、短信验证码等手段来降低自动机登录的效率。三是在活动环节，通过短信、语音验证码降低黑产刷单的效率。黑产情报方面，全面掌握互联网金融黑产的行为特点、从业人员规模、团伙地域化分布以及专业化工具等情况，并制定针对性的打击策略。黑产风险防控方面，在黑产情报方面，基于腾讯的生态系统，其积累了丰富的丰富情报收集和自动学习能力。

## 6. 百度“磐石”金融科技产品

**(1) 项目背景。**百度的技术基因正在全面注入百度金融，百度金融拥有百度独有的数据特征和算法，百度的梯度增强决策树可以聚合大数据高维特征，可以实现高维数据降维、增加风险区分度；百度的深度学习，将特征嵌入，利用关联挖掘等解决数据稀疏问题。

**(2) 应用场景。**应用于身份识别、反欺诈、信息核验、信用分等场景。

**(3) 成果概述。**百度金融以人工智能、大数据、云计算为代表的科技能力为基础搭建金融科技产品——磐石，旨在为银行、互联网金融机构等提供身份识别、反欺诈、信息核验、信用分等系列产品能力及一体化解决方式。

磐石依托百度实时、海量、多维数据，100 亿+次搜索行为(日)，8.6 亿+账号用户画像，单日数十亿级别的百度数据的增加，如搜索、贴吧、文库、知道等，以及通过多维度的深度特征加工与深度学习算法，在分布式集群上，进行同步

的事实并行集成运算，可实时或准实时产出相关核心产品。

**（4）项目收益。**截止目前，磐石金融科技产品累计为百度内部信贷业务拦截数十万欺诈用户，拦截数十亿不良资产、减少数百万人力成本。累计合作近 500 家外部金融机构，客户类型主要有：银行类、持牌类消费金融机构、持牌类小贷机构等，如：中国农业银行、招商银行卡中心、马上消费金融、维信金科、真融宝等机构，目前，外部机构日查询次数达数百万次，帮助机构降低资产损失，有效提升了金融行业整体风险防控水平。

## **7. 京东金融基于大数据的行为分析系统**

**（1）项目背景。**近年来，随着移动互联网的兴起，移动支付已经为人们的生活带来诸多便利。而近两年，人脸支付、指纹支付、声纹支付等支付方式也慢慢被人们所接受，除了支付领域，人脸打卡、人脸识别登录、指纹登录等也有较为广泛的应用，但是目前这些技术手段仍然存在较大的安全风险，比如用户利用面具、假人、假手指等方式欺骗识别系统，为了提高系统的安全性，我们采用大数据的方式对当前用户的行为进行前期建模，构建用户特征模型，通过特征分析和比较来提高系统的安全性，一旦发现非本人操作，系统将给出预警或者禁止本次交易的发生，从而最大限度的保护用户的数据安全。

**（2）应用场景。**应用于行为分析、支付安全场景。

**（3）成果概述。**每个用户在使用手机 APP、PC 时，用



户的行为习惯是有明显的差别的，即使两个很相似的人，他们的行为习惯例如左右手、点击速度、点击位置、体态特征等也是有细微的差别，而且这些大多数细微的行为习惯是不被用户所感知的，用户在潜移默化中表现出了自己独有的特征。基于此，我们利用京东大数据平台，设计用户行为分析算法模型，用来学习用户的行为特征，并对用户的行为特征最终做出综合评价，从而得出当前用户的行为是否与机主的行为匹配，一旦发现风险，则会立即给出风险预警，进一步进行身份确认，例如转账、付款等行为。

**（4）项目收益。**京东金融严格遵守相关规定，在合规前提下利用大数据平台对用户平时信息进一步分析排查，识别认证用户的身份，有效地防止非法用户假冒合法用户身份占用资源、删除或者篡改用户存储的数据。该算法有较为广泛的使用场景，例如人脸识别、支付、APP、安全等级较高的商业系统等领域均可使用。针对不同的场景我们可以进行系统定制，系统的模块化划分较细，可以针对不同的使用场景进行积木式组合，以达到最大的可复用性。

## **8. 壹账通微表情面审辅助系统**

**（1）项目背景。**当前以人工智能等新技术为驱动力的新一轮产业变革，正深刻改变人类生产生活和经济发展模式。世界主要发达国家纷纷布局新科技，抢占新一轮国际竞争制高点。在金融领域，通过 AI 新技术打造贷款风控新模式，已然成为广大金融机构转型创新的重要方向。

**（2）应用场景。**应用于风险管理场景。

**（3）成果概述。**微表情面审辅助系统，整合人工智能与大数据领域前沿科技，在贷款面审环节引入业界领先的高达 53 种的微表情识别技术，可识别 90%以上表情变化。通过实时抓取客户的微表情，智能判断并提示欺诈风险。实践中该技术可实现与人工判断 80%的吻合率，50%的人工面审替代比例。同时，该产品还整合了基于知识图谱的智能问答引擎，整合了 10 个行业的 1000 个问题库，通过大数据交叉校验挖掘潜在风险。

**（4）项目收益。**传统面核面审，主要靠业务人员人工审核，水平参差不齐，风险大。而智能微表情面审辅助系统，可以实时智能判断并提示风险；同时，覆盖 10 个行业 1000 个问题库，可大幅提升面审问题的随机性和质量，让面审更加科学高效，降低风险。微表情面签系统已在与 17 家银行进行的联合贷款合作中应用，实现面审流程智能化、规范化，减少 40%以上的人工干预。

## **9. 壹账通信贷风险态势感知平台**

**（1）项目背景。**互联网大数据时代，金融信贷业务也纷纷转战线上市场。对金融机构来说，掌握风险控制、在线资质审查、审批时效等能力，是持久稳定开展智能化、自动化线上信贷业务的核心命脉。面对流量大、变化快、捕捉难的信贷风险困境，信贷风险态势感知平台应用先进算法技术，自动实时甄别并可视化展现信贷风险点，有效帮助合作机构

监控、识别、捕捉风险信息。

**(2) 应用场景。**应用于风险管理场景。

**(3) 成果概述。**利用用户行为数据构建了行为指纹识别算法，并以行为指纹和用户画像相结合拼接数据碎片；在客户风险量化方面利用隐马尔科夫、GBDT、逻辑回归、XGBOOST 等机器学习算法识别和预测客户资产财富水平、消费能力、公开信用、强金融违约概率及违约迁移概率等关键信息。利用迁移学习、GAN 等技术有效扩充了团伙及个人的欺诈特征。帮助金融机构监控客群风险变动及在各分级金融机构的迁徙，监控不同机构的客户风险及资产质量，监控产品的舆情及产品漏洞攻击来源，监控各地域客群信用及欺诈风险。搭建一个立体的、多维度、多展现方式的金融行业信息平台。

**(4) 项目收益。**某互联网金融机构在具体业务方面经常出现客户多头借贷、拆东墙补西墙的情况。在资产组合层面出现客群与资产分布配置不合理等情况。在应用信贷风险地图后，不仅有效管控了多头借贷客户的风控策略，同时结合客户迁徙轨迹和概率有效管控了客户额度与定价策略。在产品层面，不仅优化了自身金融产品特性，还结合客群用户画像的共性特征研发了“结婚酒席礼金贷”等金融产品。在组合层面，有效利用客群及资产组合的违约相关性与迁徙概率，实现了整体资本投资收益风险最优化配置和精细化管理目标，其资本回报率提升 2.35%。

## 10. 平安壹钱包智能风控系统

**(1) 项目背景。**近几年来，第三方移动支付市场发展十分迅速，移动支付的应用场景不断拓展，交易规模快速增长。但是，在为客户提供便捷的支付服务的同时，支付风险也在不断增加，羊毛党规模日益扩大，各类黑客攻击及电信诈骗案件频发，作案手段更加隐蔽和多样化。传统风险控制中仅通过单一纬度的风控规则进行风险交易识别已经不能满足新形势下风险控制要求。如何完善风控系统来胜任多场景、多维度的风险管控要求，增强实时对抗和自学习能力，已成为风险管理领域的主要研究方向。

**(2) 应用场景。**应用于支付交易风险管理场景。

**(3) 成果概述。**为了保障支付的安全性，壹钱包结合自身积累的海量数据及平安集团内部风险名单库等信息，加入图像识别、指纹识别、人脸识别等技术搭建了以大数据为核心的智能风控系统，突破了传统风险控制中仅通过单一纬度进行风险识别的局限，根据海量交易及行为数据，运用深度挖掘、机器学习、用户画像等方法建立精准的风控模型和策略来识别风险特征，实现风险的事先预测和实时处置。同时采用大数据流式计算技术建设系统资损实时监控引擎，从交易的完整性、一致性、幂等性等角度进行实时监控，秒级发现资金异常情况。

**(4) 项目收益。**根据 17 年的实际效果来看，壹钱包智能风控系统帮助实现欺诈资金损失率控制在百万分之级别，

在各类营销活动水军识别覆盖率超过 90%。在达成风险控制的同时，极大程度降低了对正常用户的干扰，干扰率仅在十万分之级别。系统资金异常监控时效从日提高到秒，极大的提升了发现异常的时效性，帮助公司众多重点项目监测交易异常，为客户资金安全，公司业务发展保驾护航。

## **11. 银联商务大数据普惠金融**

**(1) 项目背景。**对于大部分中小微企业来说，其在人民银行征信中心没有征信记录，使得金融机构缺乏真实、可靠的数据对其进行评估并授信放款。为了能够帮助这部分中小微企业享受到普惠金融服务，需要依托外部数据源建立模型进行授信评估，同时通过系统对接的方式实现在线评估和放款，帮助金融机构降低成本和控制风险，最终达到企业、金融机构双赢的局面。

**(2) 应用场景。**应用于中小微企业融资服务场景。

**(3) 成果概述。**银联商务推出银杏大数据服务平台。依托自有数据包括银联商务自有业务合法采集和积累的 700 多万商户，8 年累计的百亿级消费数据，各类融资理财数据等。外部数据包括工商、法院、失信被执行、税务、互联网媒体等合法公开或用户主动授权的信息，为中小微企业提供无抵押无担保的普惠金融服务。企业信息来源于银联商务自有数据和外部采集数据。

**(4) 项目收益。**银杏系列产品已面向银联商务体系内互联网金融平台、商业保理公司，以及银行、消费金融、小

额贷款公司等上百家机构输出服务，累计服务中小企业几十万家，个人消费者上百万人。2017 年累计提供服务上亿次，实现直接经济收入上千万元，间接帮助金融机构实现无抵押纯在线融资业务发生额千亿元，实现互联网金融业务收入同比增长约 80%。

## **12. 天翼电子商务甜橙欺诈盾**

**（1）项目背景。**近年随着移动互联网的普及，基于互联网技术的新型诈骗、犯罪现象屡见不鲜。在互联网金融的业务场景下，客户不再面对面和业务人员交流，而是通过网络与服务器交换数据就可以自动化的完成绝大部分的业务。这些业务不再需要服务提供商与客户面对面打交道，而互联网通讯协议本身又是匿名的。这就给欺诈者带来了伪装身份的便利。欺诈者通过简单的技术手段就可以完成年龄、性别、身份等伪装，通过网络实施各种欺诈活动。目前从事网络欺诈的不法分子，已经形成庞大的黑色产业链，相关从业人员逾 160 万，网络诈骗“黑色产业”市场规模高达千亿，网络反欺诈形势非常严峻。

**（2）应用场景。**应用于风险管理场景。

**（3）成果概述。**天翼电子商务有限公司依托中国电信集团及翼支付大数据能力和科技创新优势，结合多年来在通信、支付、消费金融等领域积累的反欺诈经验，倾力打造“中国电信甜橙欺诈盾”产品。旨在识别消费金融、P2P、支付交易等场景下潜在的欺诈风险。甜橙欺诈盾作为电信集团的

核心大数据风控产品，目前已被广泛应用于翼支付消费金融、风控等部门相关业务场景，效果卓著。有效解决了虚假注册、薅羊毛、营销套利等目前市场急需解决的问题。随着互联网的发展，甜橙欺诈盾这样的产品会越来越重要，其功能也会随着技术的发展越来越完善，届时将更加广阔的市场并应用到更多的场景，将全流程覆盖行业痛点，为翼支付与电信集团保驾护航。

**（4）项目收益。**从欺诈盾产品推出开始，目前已经在内部风控系统与外部商户调用方面进行了大量的验证，已经为企业带来实际的经济价值，降低了企业资损率；外部商户的调用也为其业务开展带来了明显的提升。某 top3 互联网公司通过使用甜橙欺诈盾账户风险评估模块中的二次放号风险预警功能点。二次放号识别率达到 2%左右，将这一部分历史黑名单用户成功洗白。甜橙金融使用甜橙欺诈盾对营销套利用户评估的准确率达到 93.48%，召回率达到 72.24%。经评估帮公司有效降低约 2000 万元左右的风险损失。

### 13. 享宇金服移动手机贷

**（1）项目背景。**数据是用技术手段衡量风险的基础。电信运营商具有本地化渠道覆盖、用户规模、终端入口、落地运营等多方面的差异化优势，金融科技公司的核心能力在于对银行业务的深刻理解、对数据的标准化应用、对风控模型的精密设计和模块化输出。双方联手，可以打出进军互联网金融的“通信大数据+金融”的组合拳，为客户提供免费、

高性价比、便捷的金融增值服务。

**(2) 应用场景。**应用于风险控制场景。

**(3) 成果概述。**移动手机贷深耕消费金融。针对银行、持牌消费金融机构等金融机构的风控难题和线上获取客户难题，移动手机贷依托运营商数据，以及享宇金服自有知识产权的“蜂巢”大数据风控模型，一旦获得客户授权，便能运用大数据数据分析能力，合规合法使用运营商脱敏数据，自动对用户进行身份识别，3-10 秒即可为客户完成“画像评分”，输出客户标识、用户主手机号、消费情况、流量使用情况、来电稳定性、去电稳定性、蜂巢模型反欺诈等 8 大类共计 85 个细项内容，从不同维度对用户画像进行详尽描述，并可以进行统一内容的多维度交叉核验，解决金融机构长期对反欺诈客户的痛点问题。

**(4) 项目收益。**截至 2017 年 6 月末，“移动白条”总申请人数达到 500 万人次，日均注册 15000 人，日均申请 13000 人；用信额人均 1.02 万；审批通过率为 13.8%；不良率仅为 0.89%；从申请到放款，平均用时 1 分钟。目前，移动手机贷每月新增放款近亿元，每天新增注册用户超过 15000 人，注册申请率为 85%，注册转化率达 10.2%。

## 14. 人人贷风控体系

**(1) 项目背景。**随着移动互联网的迅速发展和技术进步，小微企业主和个体经营户的互联网普及程度不断提升，大量数据随之产生，机构可以通过标准化和规模化的方式为



其提供服务，并通过对个人信用的评估来满足其经营性融资需求。大数据、金融科技为业务发展提供了强大的技术支持，能够使网贷平台更好地识别用户风险，为其提供更有针对性的服务。

**(2) 应用场景。**应用于风险管理场景。

**(3) 成果概述。**人人贷创立线上线下结合的数字化风控体系，覆盖从信贷产品设计环节的客户门槛、产品定价、流程设计、风险埋点，到贷款审批环节的资料核实、大数据反欺诈、客户分类、数据化风险排序、人工复核，再到贷中管理环节的客户提醒、数据追踪，到风险预警，到最后催收环节的电话提醒、失联后数据查找、催收作业等。

**(4) 项目收益。**通过运用智能反欺诈体系、信用风险评估、人工+机器三重贷前的风控机制，人人贷将欺诈风险遏制在源头，严格筛选优质借款人。同时，通过贷后专业的管理控制资产风险，大幅降低欺诈水平与逾期率。并提升信审效率，将信审时效控制在1小时之内。截止2018年3月，人人贷累计撮合借款金额超520亿元，累计服务借款人数超66万，成交笔数超70万。据2017年年度报告显示，截至2017年末，人人贷金额逾期率为0.27%，项目逾期率为1.11%。

## **15. 通付盾基于大数据和设备指纹技术的风险决策与用户行为分析系统**

**(1) 项目背景。**传统反欺诈手段，基本上是以黑名单与业务名单的比对，经过人工经验的判断，最后发现欺诈行

为。通过外部渠道发现欺诈行为，耗时将近 300 多天，而通过内部渠道，也要 60 天左右才能发现欺诈。而多个系统之间账号体系、操作数据很难统一，尤其是跨机构、跨系统情况，直接导致金融机构用户分析的有效性降低，无法达到预期目标。

**（2）应用场景。**应用于设备画像、用户行为分析、用户画像、反欺诈风险管理、精准营销等场景。

**（3）成果概述。**通付盾提出基于大数据和设备指纹技术的风险决策和用户行为分析系统。用户在同一移动设备（手机、平板等）上使用不同 APP，不同账号，使用设备 ID 作为唯一标识，同时不强制上传用户数据，简化跨系统的接入成本，设备 ID 作为隐形的账号体系，打通各机构和各业务系统。通付盾设备指纹通过获取上网设备的多重属性为该设备生成跨浏览器跨应用的唯一的智能设备 ID，基于快速识别在线设备的各项属性，判断交易的可信度，从而达成风控及反欺诈的目的。同时，通付盾设备指纹加入了防篡改、防劫持等黑客防范手段，并根据行业风险特性建模提取风险设备，全行业黑名单共享，群防群控，联合打击业内欺诈行为。

**（4）项目收益。**在反欺诈风险预警方面，通过阻断及时交易，冻结银行卡，短信、邮件、电话提醒客户等方式，已累计为客户避免损失超过百亿元。在精准营销方面，通付盾通过用户行为分析、用户画像、用户分群、渠道效果分析、渠道作弊防护、风险分析、UGC 自动检测、精准推送等手段，

将客户转换率提高 10%，新增用户留存率增长 5%。在提高工作效率方面，通付盾通过各类大数据工具如设备指纹、用户行为分析系统、反欺诈系统等，为金融机构客户减轻人工审批，人工风险监控，人工逾期管理等工作内容，工作效率提高 80%以上。

## **16. 数尊信息众邦银行线上风控一体化实施项目**

**（1）项目背景。**近年来，伴随互联网金融、消费金融等业务形态的崛起，个人金融信贷业务范围逐步扩展至新场景、新客群。黑产盛行，客群趋向年轻化、长尾化，不喜等待、重体验轻费用等特点使得传统的风控手段不再满足当前的业务需求。同时，数据分析、人工智能、机器学习等新技术也不断推动着银行整体的能力升级。

**（2）应用场景。**应用于风险管理场景。

**（3）成果概述。**依循“金融科技”和“智能风控”的核心脉络，众邦银行携手数尊科技搭建了一站式风控决策大脑——“麒麟盾”平台。在支持线上信贷场景全生命周期智能风控策略承载的基础上，该平台创新式构建了数据智能引擎、灵活的风控嵌入（业务）机制、高便捷操作平台、智能节费策略、风控图谱应用等，为风控与业务的平衡、用户体验与机构收益的平衡、长远规划与当前需求的平衡奠定了坚实的基础。

**（4）项目收益。**截止目前，该平台每天承载业务量超过数万笔。自动化线上审批流程的实现，大大降低了人工审

核的成本，将一般银行业务中三天到一周才能完成的审批流程，在毫秒级就实现了自动化的审批和放款。同时精准的风险模型和策略，极大的对好坏客户进行了有效区分，通过率从之前的 30%达到了 55%，业务增长明显，与此同时，风险逾期率也有效控制在 3%以内，达到了风险和收益的极大平衡。

## **（二） 精准营销类应用**

### **1. 陆金所投资者适当性管理体系**

**（1）项目背景。**互联网财富管理平台利用高效开放的互联网工具，在一定程度上降低了传统资产管理机构服务门槛，提升了价格发现效率，带给投资者更开放、普惠的投资体验。但同时由于监管环境较为宽松，行业准入门槛较低，致使行业参与者良莠不齐，一些平台的跑路和倒闭，更是给广大投资者造成了伤害与损失。

**（2）应用场景。**应用于投资者适当性管理、用户画像、精准营销、风险管理等场景。

**（3）成果概述。**陆金所设计并建立了投资者适当性管理体系。该体系创新地利用大数据和机器学习等新技术对投资者风险承受能力进行精准评估，同时为了让客户直观了解自己风险承受能力高低，对其进行了量化处理，输出“坚果财智分”。通过准确对金融资产进行风险分类，对投资者进行风险承受能力分类，按量化模型对两者进行匹配，辅之以清晰、完整、全流程的信息披露以及投资者教育体系，实现资产和资金的精准匹配和产品全生命周期风险管理，将合适

的产品卖给合适的投资人，保护投资者的合法权益。

**(4) 项目收益。**自 2016 年以来，实现了超 3 万亿产品与投资者销售匹配；累计对 127 万人、236 万笔交易进行风险超配提示、拦截，金额 3829 亿元；拦截不匹配的投资者人数和交易笔数处于下降趋势且趋于稳定，说明该体系有效引导投资者购买适合自己的产品。

## **2. 恒丰银行基于大数据的客户关系管理系统**

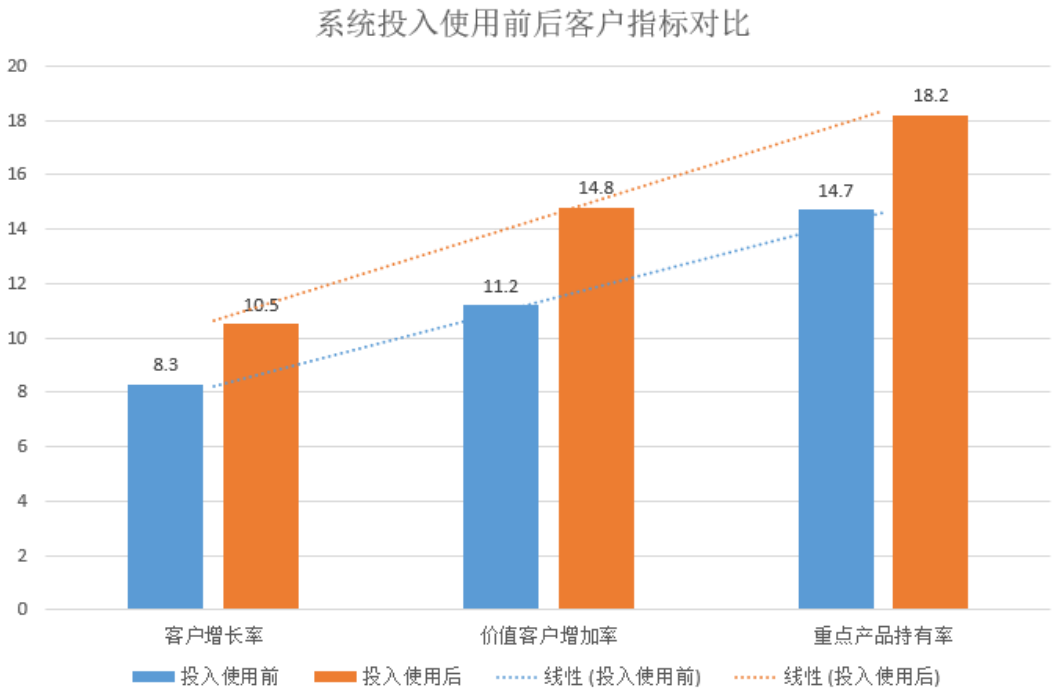
**(1) 项目背景。**在互联网金融迅速发展的背景下，差异化营销和个性化服务越来越成为银行长期客户维系的重要方面。传统银行 CRM 主要关注内部数据，关注如何把银行内部各个业务环节中零散的客户信息搜集、汇聚起来。而在大数据时代，伴随社交和移动化的盛行，外部数据越来越丰富，促使银行不仅要关注内部数据，更要想办法把外部数据整合利用起来。通过多种渠道获取大量中、高价值潜在客户信息，获取更多的销售商机和线索，充分了解客户的个性需求并提供差异化的服务和解决方案；拓展传统销售渠道，利用新媒体、新渠道开展精准营销，提高营销环节的投入产出比。

**(2) 应用场景。**应用于客户关系管理、精准营销场景。

**(3) 成果概述。**CRM 系统依托恒丰银行自主研发的企业级大数据技术平台，采用微服务软件架构、实时流处理技术和人工智能技术，通过对内外部数据的深度整合和价值提炼，提供客户 360 视图、工作提醒、智能客户推荐、营销机会发

掘、产品货架与优化组合方案、行业资讯、客户风险预警、移动信贷业务、团队协同管理、业绩看板等业务功能，为业务团队掌握市场动态、识别客户价值、预见客户风险、实现精准营销和团队协作提供信息技术支撑，有力提升商业银行的客户服务水平和市场竞争。

**（4）项目收益。**客户经理通过产品分析生成的流失客户预警进行客户挽留，降低客户流失率。同时通过产品推荐和智能获客，提高了新客户增长率和产品持有率，新客户增长率、价值客户增加率和重点产品持有率明显提升，以下是客户指标在2016年2月至2017年4月之间的使用前后对比。



### 3. 万丈金数保险行业智慧电销解决方案

**（1）项目背景。**保险行业传统电销陷入模式同质化困境，在监管趋严的态势下，产能压力和运营压力逐渐提升，

表现在：一是行业增速放缓，保险电销增速总体趋于平缓，2014 年增速 19.6%，2015 年增速 16%，2016 年增速依旧为 16%，业务红利期逐渐过去；二是数据痛点，企业有数据，但字段简单，形不成标签；数据沉睡难激活；数据不流动，没有实时性；三是业绩转化难题。行业数据显示，目前电销渠道的成单率有 1.5%--3%。这个数据一方面反映了社会大众对电话销售的认同度极低，另一方面也反映了电话销售对客户数据的需求之大。四是人力难题。由于工作内容同质化、秒挂、投诉、业绩压力，使得传统电销坐席的工作价值感极差，再加上业绩不佳导致的收入困难，流动过于频繁。以上种种痛点和难点加速推动了险企的转型突破需求。

**（2）应用场景。**应用于潜在客户识别、数据运营、交易促成等场景。

**（3）成果概述。**万丈金数保险行业智慧电销大数据平台依托用户大数据中心，为传统电销中心提供一揽子“技术+服务”解决方案。方案的核心思路是：整合电销中心存量数据、拓新数据、渠道数据、场景数据等，建设用户大数据中心，然后通过用户画像、用户评分体系等模型，筛选出不同价值潜在客户，再依托智能电销系统、语音识别、人工智能等基础设施和技术支持，针对不同类型的潜在客户制定个性化营销策略，包括智能话术、个性化内容推送、AI 互动、事件营销等。

**（4）项目收益。**项目收益主要表现在电销中心的效率

提升和业绩提升两个方面。初步计算，目前电销行业的人均产能 1.62 万，经过解决方案的系统化提升，初步能够实现约 15~30%的提升。

### **（三） 其他类型应用**

#### **1. 监管科技—支付宝反洗钱智能分析产品**

**（1）项目背景。**识别违法犯罪活动能力的提升需要长期的培训和‘实战’经验沉淀。而这在一定程度上阻碍了金融机构提升反洗钱和打击犯罪能力，导致漏过和无法识别出违法犯罪活动。蚂蚁金服需要有一种自动沉淀分析经验进行可疑交易分析的产品。

**（2）应用场景。**应用于监控和分析可疑交易场景。

**（3）成果概述。**反洗钱智能分析产品主要具备两个功能，一个是反洗钱智能关系网络，它通过后台大数据计算后将与客户有关的资金链路和数据分析结果直观展示给可疑交易分析人员，用于快速定位和识别洗钱风险；另一个是智能审理，它能够智能学习可疑交易审理人员的经验，对可疑交易案件进行自动分析，并将分析结果提供给审理人员参考。

**（4）项目收益。**通过反洗钱智能分析产品，蚂蚁金服反洗钱中心将可疑交易人工分析的效率提升了 30%。对疑似涉嫌赌博、非法集资、传销等犯罪类型的识别时间从平均 30 分钟缩短到 5 分钟，并且识别的准确性始终保持在 90%以上。在识别可疑资金的流向以及可疑交易主体之间是否具有团伙关系等上做到了直观可见的，完全不需要反洗钱中心工作



人员手工绘制资金流向图，对业务效率的提升起到了显著的提升作用。

## **2. 监管科技—金信网银大数据监测预警金融风险平台**

**（1）项目背景。**近年来，随着国家对金融市场管控政策的不断调整以及互联网金融的快速发展，非法集资的犯罪手段和形势发生了很大变化。仅在 2016 年，全国检察机关公诉部门受理非法集资案件 9500 余件，其中，非法吸收公众存款案 8200 余件、集资诈骗案 1200 余件。受理金融犯罪案件 30900 余件 48700 余人，起诉 23700 余件 36300 余人。从以上数据来看全国打非形势依然非常严峻。同时，伴随着互联网大环境下网络非法集资的规模大比例增加，作案手段更趋隐蔽化、多样化给金融监管工作带来了更大的困难。

**（2）应用场景。**应用于舆情监测、风险预警场景。

**（3）成果概述。**大数据监测预警金融风险平台以大数据云计算为技术支撑，构建针对五大领域十七个行业的分析模型。通过机构合规程度、网络投诉率、收益率、特征词命中率、传播力等 5 个维度多项数据，综合分析计算监控对象非法集资风险相关度（即冒烟指数）。冒烟指数测评分数越高，则非法集资风险就越高。预警平台融合文本挖掘技术、社会网络分析方法、LBS 分析技术等大数据前沿技术，通过对企业、人、资金三方面监测，不断优化升级“冒烟指数”监测预警模型，提升监测预警准确性。

**（4）项目收益。**截至 2017 年 10 月底，大数据监测预

警金融风险平台根据预警模型选定的 100 余项指标进行分类筛选并分析计算出最终结果。据监测，大数据监测预警金融风险平台应用已对全国 100 多万家类金融企业常态监测，及时预警高风险企业，已主动识别高风险企业 800 余家。其中高风险企业（冒烟指数 $\geq 60$  分）共计 46 家，存在涉嫌违规经营、承诺高收益、涉嫌传销、涉嫌诈骗、涉嫌虚假宣传、存在涉诉信息、存在经营异常、存在行政处罚、存在失信被执行人信息、未在基金业协会备案、信息披露不透明等问题。在监测预警、风险排查方面发挥重要作用。

### **3. 智能投顾—通联浙商大数据智选消费基金**

**（1）项目背景。**传统投资研究对于宏观环境、行业研究、公司分析的数据大多存在时间滞后、样本偏差、统计失真、费时费力等问题。另外，随着上市公司的数量增加，涉及行业范围不断拓展，传统投研方法也会遭遇瓶颈。

**（2）应用场景。**应用于投资管理场景。

**（3）成果概述。**通联支付通过对自有的消费类支付相关数据，其中包括支付的细分行业、销售额、笔数、刷卡数，以及实际的销量和价格等进行分析，可以实时了解行业（尤其是消费行业）销售需求的情况，按行业汇总各商户的刷卡支付情况，获得行业最新的景气边际变化，进而将资金更多的配置在景气向好的行业上，然后利用经典量化模型，精选相应行业内的上市公司，并基于此发行了一支名为“浙商大数据智选消费”的偏股混合型基金。

**(4) 项目收益。**该基金于 2017 年 1 月 11 日成立，截至 2017 年 11 月 13 日，累计收益率 9.6%，年化收益率 11.5%，夏普比率 1.35。

#### **4. 智能投顾—中诚信资产交易智能扫描平台 AXIS**

**(1) 项目背景。**近年来，互联网的广泛普及，大数据和人工智能等新技术的发展，给信用产业的革新提供了新的契机。如何将新技术成功应用于信用产业，使金融机构能够更加准确地进行对信用风险的识别、计量和定价，从而能够更加有效地管理信用风险？信息不对称是制约消费金融 ABS 市场发展重要的瓶颈，这个问题在资产方主体信用较低时尤为突出。破解消费金融 ABS 市场的信息不对称困局，实现风险可控且高效的资产流转，是 AXIS 系统在消费金融 ABS 领域要解决的核心问题。

**(2) 应用场景。**应用于智能投顾、风险控制等场景。

**(3) 成果概述。**中诚信征信自主研发的针对消费金融资产证券化信息服务解决方案——资产交易智能扫描系统 (AXIS)，在资产端，可以为主体信用不足但资产信用良好的资产方做到间接增信；在资金端，可以帮助投资者减少逆向选择的风险，增加优质的资产标的。对于整个市场，可以打通从借款人到 ABS 投资人的信息流和资金流，降低信息成本和资金成本。

**(4) 项目收益。**自 2017 年 7 月产品发布以来，基于 AXIS 系统在执行和洽谈的项目超过 20 个，由 AXIS 系统处理的资

产超过 100 亿。对于投资人，AXIS 系统可以帮助投资人从源头上把控资产信用情况，降低逆向选择的风险，增加优质的资产标的，实现更加精确的信用风险管理。对于资产方，AXIS 可以帮助资产方打通风控端和资金端，从而降低资金成本。基于已完成的项目数据的估算，中诚信征信基于 AXIS 系统的解决方案可以帮助资产生成方节省 0.3%-0.5% 的资金成本。

## 5. 基础数据—中国移动乾坤大数据

**(1) 项目背景。**在金融科技迅速发展的时代大背景下，金融行业面临互联网大数据应用的冲击，在多个业务领域对跨界数据有十分迫切的需求。中国移动拥有超过 9 亿的个人用户，日均产生海量的业务数据。借助运营商的数据资源，金融机构可以对用户的身份特征、消费能力、行为偏好、信用情况等进行分析，从而为用户画像、精准营销、用户身份验证、征信、风控、运营优化等工作提供支撑。

**(2) 应用场景。**应用于金融机构用户画像、精准营销、征信、风险管理、客户认证、运营优化等业务场景。

**(3) 成果概述。**中国移动乾坤大数据产品是依托中国移动的海量数据资源，整合网络能力，打造的一款面向应用场景的多种数据产品和服务。产品面向银行等金融机构，提供基于大数据的数据标签、商业咨询、智能定位等服务，用于金融行业征信评估、风险管控、精准营销等业务领域，助力客户完善风险防控手段，为精细化运营提供保障。

**(4) 项目收益。**基于该成果，中国移动已与交通银行、

民生银行、浦发银行开展了大数据合作，用于金融机构日常在风险防控、运营分析等领域。预计未来随着金融机构大数据应用范围的扩大，全国性商业银行的年均查询量预计将达到亿级。借助运营商的数据资源，金融机构将可以对用户的身份特征、消费能力、行为偏好、信用情况进行多维度、综合分析，从而为用户画像、精准营销、用户身份验证、征信评估、风控、运营优化等工作提供支撑。

## **四、金融大数据应用面临的挑战**

### **（一）金融行业的数据资产管理应用水平仍待提高**

金融行业的数据资产管理仍存在数据质量不足、数据获取方式单一、数据系统分散等一系列问题。一是金融数据质量不足，主要体现为数据缺失、数据重复、数据错误和数据格式不统一等多个方面。二是金融行业数据来源相对单一，对于外部数据的引入和应用仍需加强。三是金融行业的数据标准化程度低，分散在多个数据系统中，现有的数据采集和应用分析能力难以满足当前大规模的数据分析要求，数据应用需求的响应速度仍不足。

### **（二）金融大数据应用技术与业务探索仍需突破**

金融机构原有的数据系统架构相对复杂，涉及的系统平台和供应商相对较多，实现大数据应用的技术改造难度较大，而且系统改造的同时必须保障业务系统的安全可靠运行。同时，金融行业的大数据分析应用模型仍处于探索阶段，成熟

案例和解决方案仍相对较少，金融机构应用大数据需要投入大量的时间和成本进行调研和试错，一定程度上制约了金融机构大数据应用的积极性。而且，目前的应用实践反映出大数据分析的误判率还比较高，机器判断后的结果仍需要人工核查，资源利用效率和客户体验均有待提升。

### **（三）金融大数据的行业标准与安全规范仍待完善**

当前，金融大数据的相关标准仍处于探索期，金融大数据缺乏统一的存储管理标准和互通共享平台，涉及金融行业大数据的安全规范还存在较多空白。相对于其他行业而言，金融大数据涉及更多的用户个人隐私，在用户数据安全和信息保护方面要求更加严格。随着大数据在多个金融行业细分领域的价值应用，在缺乏行业统一安全标准和规范的情况下，单纯依靠金融机构自身管控，会带来较大的安全风险。

### **（四）金融大数据发展的顶层设计和扶持政策还需强化**

在发展规划方面，金融大数据发展的顶层设计仍需强化。一方面，金融机构间的数据壁垒仍较为明显，数据应用仍是各自为战，缺乏有效的整合协同，跨领域和跨企业的数据应用相对较少。另一方面，金融行业数据应用缺乏整体性规划，当前仍存在较多分散性、临时性和应激性的数据应用，数据资产的应用价值没有得多充分发挥，业务支撑作用仍待加强，迫切需要通过行业整体性的产业规划和扶持政策，明确发展重点，加强方向引导。

## **五、促进金融大数据发展应用的相关建议**

### **（一）出台促进金融大数据发展的产业规划和扶持政策**

建议针对产业发展需求和政策空白领域，出台促进金融行业大数据发展应用的指导性政策意见，明确产业发展的目标、方向、路径和要求，完善产业发展的配套保障体系和发展能力评估建设体系。指导和支持金融大数据在产业标准、安全和商业化等多个领域的相关研究。逐步加快发布和形成金融大数据产业应用标准体系和行业规范，以标准促进产业合作，创造更加良好的产业发展环境，增强产业界发展积极性。

### **（二）分阶段推动金融数据开放、共享和统一平台建设**

针对金融机构数据分散和隔离问题，建议监管机构牵头，分阶段推进金融行业安全可控的数据开放共享。首先从制定统一数据目录，明确最低开放标准着手，逐步鼓励金融机构创新合作模式，搭建金融行业统一数据平台，克服跨组织数据流通障碍。未来可鼓励金融机构探索混合所有制，建立独立运营主体，负责金融行业大数据的统一管理和运营，开展跨行业、跨领域应用合作，促进金融大数据在社会经济各领域的价值实现。

### **（三）强化金融大数据行业标准和安全规范建设**

建议组织金融行业各方主体，协同制定统一的金融行业大数据交易规范，明确交易各方的数据安全责任，保障金融

大数据市场的健康、有序发展；制定明确的数据安全使用标准，对金融大数据的使用权限、使用范围、使用方式和安全机制等，进行严格的规范化、标准化管理；建立有效的投诉机制和惩罚机制，实施全程全网的数据安全使用管控与源头追诉。

#### **（四）依托行业平台推进金融大数据应用成果共享合作**

积极发挥以“中国支付清算协会金融大数据应用研究组”为代表的行业组织的平台作用，打造具有品牌影响力的金融大数据交流分享平台，建立金融大数据行业的长效沟通机制，促进金融大数据应用成果的经验分享和互动交流。同时，积极推动金融行业和电信、电商、旅游等跨行业的沟通和合作，通过专题活动宣传和推广，展示金融大数据在各个行业领域的应用成果，增加金融大数据应用的社会关注度。