

案例使用说明：

大数据助力中小商业银行个人信贷风险管理

一、教学目的与用途：

1. 本案例主要适用于 MBA 学生学习数据挖掘、机器学习、数据模型与决策相关课程，也适用于风险管理相关问题的案例讨论。

2. 本案例的教学目的是帮助学生了解互联网环境下中小商业银行利用大数据进行个人信贷风险管理时遇到的实际问题，通过案例研讨与分析，使学员深刻理解中小商业银行所处的时代与行业背景，了解其进行线上个人信贷风险管理的必要性，以及有效进行风险管理的关键点；同时，通过本案例的研讨，学员不仅熟悉了利用大数据分析问题的典型的理论与方法，并接触到了中小商业银行在进行风险管理中的实际问题，这将帮助提升学员的专业知识水平以及综合分析问题的能力。

二、启发思考题

为了帮助学生进一步理解案例并启发学生思考，以下问题可以作为课前思考题布置给学生。通过回答下述问题，学生能够更清晰地了解中小商业银行在个人信贷风险管理方面的现状与瓶颈，基于此，能够更快速地识别出其运营过程中所面临的突出问题，并进一步思考解决该问题的可能方法，最后通过归纳法总结出广大中小商业银行在个人信贷风险管理方面所面临困境及解决方案的共性之处。

为了便于课堂讨论，教师可以在课前将学生分组，并根据授课重点挑选 2-3 个启发思考题，提前 1-2 周布置给学生，要求学生针对给出的思考题在组内进行多次讨论，并将成果制成 PPT 演示文稿，课堂上教师可以随机抽取 2-3 个小组进行演示，其他小组可以进行提问与补充。剩下的思考题可以作为在课堂上进行讨论的材料，并要求学生课后对思考题答案进行进一步整理与完善，以小组为单位形成书面报告并提交。

1. 请进一步查阅有关中小商业银行风险管理方面的资料，并结合案例思考，目前中小商业银行在个人信贷风险评估方面存在哪些突出问题？传统的专家评分系统有哪些不足？

2. 针对线上个人消费贷款新业务，评价中小商业银行在进行风险评估过程中可能存在的问题。并结合案例回答，项目组是如何从专家评分系统过渡到基于大数据的决策支持系统的？

3. 案例中的项目团队在开发风险评估系统的过程中遇到了什么关键难题，他们是如何解决的？

4. 线上个人消费贷款预授信系统开发出来之后，预授信额度如何确定，以及该系统在实际中应该如何运用？

三、分析思路

教师可以根据自己的教学目标来灵活使用本案例。这里提出本案例的分析思路，仅供参考。

本案例从中小商业银行面临的困境出发，针对中小商业银行有关线上个人消费贷款的风险评估现状，围绕“提出新形势下中小商业银行采用大数据手段进行风控管理的必要性→梳理出中小商业银行风控管理系统化过程中存在的难点→建立全面的线上个人消费贷款风控系统→测试风控系统实施效果”这一主线展开问题分析与讨论。

本案例的具体分析逻辑思路如图 1 所示：

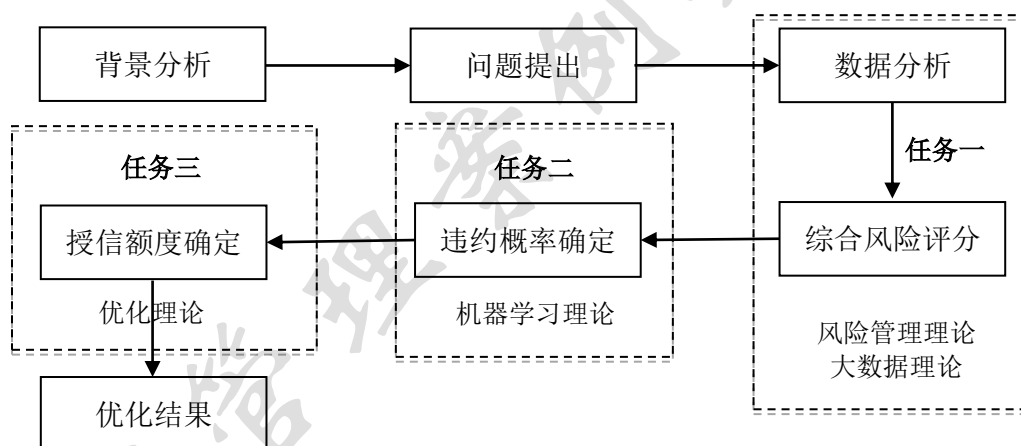


图 1 分析逻辑图

Figure 1 Logical map of analysis

首先，通过分析目前我国中小商业银行在风险管理方面所处的市场大环境，结合新形势下中小商业银行运营现状，站在后台优化提供者的角度，分析中小商业银行采用大数据手段进行风险控制的必要性。

其次，梳理中小商业银行运用大数据手段针对线上个人消费贷款进行风险评估过程中存在的关键问题与难点，并结合案例中涉及银行的实际情形进行进一步的分析和说明。

再次，以案例中涉及的中小商业银行为主体，明确其数据库中所现有数据

维度与数据结构，并基于此开展后续分析。

接着，站在后台优化提供者的角度，提出建立全面的线上个人消费贷款风控系统的三个关键步骤，包括综合风险评分、违约概率确定及授信额度确定，并针对每个步骤进行详细的阐述与解释。

最后，采用某银行实际的样本数据带入新建的风险评估系统进行测算发现，相对于传统的专家评分系统，新建的系统能够更好地评估风险，尤其是能够识别一些传统专家评分系统不能够识别的高风险人群，从而得出新建系统的优势。

了解了上述整体分析逻辑思路之后，教师可以结合启发思考题，一步一步地引导学生进行课堂讨论与思考，问题的答案并不唯一，分析的思路也不仅仅局限于以下内容，教师可以引导学生进行发散思考，帮助学生运用所学理论知识进行问题分析并提出解决方案即可。

四、理论依据与分析

1. 请进一步查阅有关中小商业银行风险管理方面的资料，并结合案例思考，目前中小商业银行在个人信贷风险评估方面存在哪些突出问题？传统的专家评分系统有哪些不足？

本题共包含两个子问题，呈递进关系，需分步作答。第一个问题需要学生搜集相关资料，并进行归纳总结，第二个问题需要学生结合案例材料思考后回答。学生通过了解目前中小商业银行针对风险评估方面存在的问题，结合行业现状和市场需求，分析如何从“经验决策”过渡到“数据决策”，因此学生应掌握风险管理的相关理论，包括：

- 风险管理的相关定义

风险管理：目的是最大化关系人的利益，但风险管理无法完全消除风险，只是通过系统的手段和策略来对风险做相应的应对，如：避免、减少、转移、持有。

信用风险：从狭义上将是指借款人到期不能或不愿履行借款协议，无法按期还本付息导致银行蒙受损失的可能性，也就是信贷违约风险；从广义上说，商业银行信用风险是指商业银行在经营过程中，由于各种不可测或不确定因素的影响，使银行在信贷经营与管理过程中，经营实际结果与预期收益目标发生背离，从而导致银行蒙受资金损失或收益下降的一种可能性或不确定性。

信用风险评估：指运用统一的方法和标准，根据客户的个人属性信息、交易行为、不良信贷情况、家庭基本情况、偿还能力、贷款抵质押情况、人行征信记录等信息，对贷款及时、足额还款的可能性进行评估，确定贷款的违约概率、违约损失率和违约风险暴露情况，从而进一步确定客户授信额度的风险计量方法。

- 风险类别

银行面临主要的风险包括以下几类：

信用风险又称违约风险，是指由于债务人或交易对手违约或信用评级、履约能力降低而造成损失的风险。

操作风险是银行在个贷业务办理时所产生的各种风险，譬如人员操作失误，管理制度的弊端等，都将给银行贷款带来一定的风险。在《巴塞尔新资本协议》定义的标准的操作风险主要包括四大类：由于人员因素发生的风险，例如员工操作失误、银行违反用工业务条例、关键人员流失等情况；由于操作风险流程因素引起的，例如操作流程设计不合理或者流程执行不严格引发的风险；由于操作流程因素引起风险，例如操作流程设计不合理或者执行不严格等引发的风险。

市场风险主要表现在三个方面：一是由于借款人收入发生的变动而引发的风险，二是借款人提供的抵押物价值的变化（一般为抵押物贬值）引发的风险，三是由于金融市场不规范引发的风险。

法律风险主要是指当借款人的主体资格、抵押物价值等受到法律的因素影响而引发的风险，与个人信贷相关的法律风险主要的三种形式：一是由于借款人的主体资格因素所导致的个人贷款风险，不同的借款主体符合不同的贷款品种，在法律处置的过程中，当借款人的主体资格失效后也会引发风险，二是由于借款人贷款资金投向所引发的风险，在国家及银行规定中，个人贷款不能投向禁止的领域，贷款发放后必须做好贷后跟踪检查，以防贷款资金被挪用；三是由于诉讼时效过期而引发的风险，在个人贷款后管理中，由于对诉讼时效把握不准，在通过法律途径处置不良贷款时，诉讼时效已经时效，导致贷款处置时带来不确定因素。

利率风险是指当市场利率出现不利于商业银行的财务状况时所引发的风险，不仅是银行的盈利水平，银行的资产、负债以及其他表外金融工具的经济价值也会受到利率所影响，利率风险的主要四个表现形式：银行收益曲线，银行重新定义风险、基准风险、和期权风险。

系统风险又称作不可分散风险，是指因为某种全局性的共同因素引起的投资收益的可能变动，这种因素以同样的方式对所有的收益产生影响。

（1）请进一步查阅有关中小商业银行风险管理方面的资料，并结合案例思考，目前中小商业银行在个人信贷风险评估方面存在哪些突出问题？

目前，国内商业银行都有各自的信贷风险管理办，一般按照信贷流程的贷前、贷中和贷后三个阶段进行风险管理和控制（如图2所示）。这种传统的风险管理方式基于专家法，普遍依赖有限的管理人员，利用人工进行贷前调查，识别信贷业务中可能存在的潜在风险，贷中阶段进行资料审核、还款人资质审核，贷

后阶段在海量的贷款中盲目抽查、寻找风险点的传统风险管理模式。因此，传统信用风险管理表现出缺乏科学的定量分析的手段而更多地倚重定性分析和管理者主观经验和判断的艺术性的管理模式。

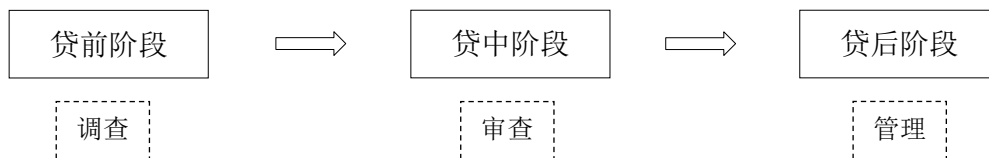


图 2 风险管控的三个阶段

Figure 2 Three stages of risk management and control

通过传统的方式方法完成了风险识别和评估后，传统商业银行一般采取风险回避、风险抑制、风险转嫁、风险自留的方式对各种风险进行处置，确保个人信贷资产按照预定目的用途使用并按时归还。部分外资的商业银行也允许个人信贷的客户进行提前还款，适当归还本金，或是通过担保或保险的方式转嫁风险，而风险自留则是根据贷款人的信用级别，匹配贷款拨备，然后提取风险准备金，按照规定完成会计处理。总之，商业银行在个人信贷风险控制方面通过系统的、成体系的、规范的方式对个人消费信贷风险因素识别、评价、控制和处理，将包含借款人本身、贷款银行和外部环境的影响等的一般的风险因素综合考虑，传统上基本利用专家预测法等来对这些因素进行识别。

个人信贷业务面对客户众多，而个人客户面对的经济业务纷繁复杂，加上我国目前诚信制度还待完善中，信息的不对称使金融机构难以对每个信贷客户进行深入的调查了解，因此存在借款人不符合贷款资格、借款人不具备还款能力甚至违背借款人真实意图的贷款申请，主体资格的缺失意味着贷款发放存在巨大风险。

随着风险控制要求在持续提升，人力资源又往往很有限，这两者之间矛盾不断加深，严重破坏了风险防控效果，从而导致银行在重大风险彻底暴露后才能发现，错失风险化解良机，银行的信贷风险管理模式迫切需要转型，即从“经验决策”迈向“数据决策”。

（2）传统的专家评分系统有哪些不足？

如案例中所述，在风险评分中专家对于影响因子的判断，具有一定的权威，但是得分和违约概率并非是线性影响关系，例如较高的年龄未必意味着违约概率越大。

其次，专家评分系统过度依赖于专家个人决策能力，难以保证各个专家决策的一致性和准确性，应用具有局限性。

最后，由于专家法普遍基于经验判断，缺乏定量分析技术，导致客户信用风险评估缺乏客观性，可能会带来操作风险。

2. 针对线上个人消费贷款新业务，评价中小商业银行在进行风险评估过程中可能存在的问题。并结合案例回答，项目组是如何从专家评分系统过渡到基于大数据的决策支持系统的？

本题分为两个子问题，第一个问题考察学生对行业的了解程度以及提炼观点的能力，第二个问题考察学生运用理论知识以及综合分析问题的能力。要回答本题，学生应了解大数据的相关理论与应用，包括：

- 大数据的定义

大数据技术的发展历程已经超过三十年，BIG DATA 这个概念是由学者多弗雷于 20 世纪八十年代提出来，推出时并无过大影响，但后来无数实践证明是其推动了第三次科技浪潮。随后在业界著名科学期刊等借鉴了这个大数据名词的说法，并预测大数据很可能是下个颠覆社会生产力的技术，更慢慢沿用至今。大数据之所以称之为大数据，不仅仅意味着数据的体量庞大，数据的结构复杂、类型众多，数据的集合采集、存储和计算量大，也意味着在合理时间内，对数据集合的管理、处理、并整理成为能帮助政府机构和企业进行管理、决策信息的能力的要求高，既从无序的、杂乱的数据和信息中筛选出相关性数据的困难性颇高，还有就是数据要永恒线上，随时保持被调用状态。

- 大数据的发展阶段

20 世纪末，是大数据的萌芽期，处于数据挖掘技术阶段。随着数据挖掘理论和数据库技术的成熟，一些商业智能工具和知识管理技术开始被应用。2003 年-2006 年是大数据发展的突破期，社交网络的流行导致大量非结构化数据出现，传统处理方法难以应对，数据处理系统、数据库架构开始重新思考。2006 年-2009 年，大数据形成并行计算和分布式系统，为大数据发展的成熟期。2010 年以来，随着智能手机应用，数据碎片化、分布式、流媒体特征更加明显，移动数据急剧增长。2011 年麦肯锡全球研究院发布《大数据：下一个创新、竞争和生产力的前沿》，2012 年维克托·舍恩伯格《大数据时代：生活、工作与思维的大变革》宣传推广，大数据概念开始风靡全球。2013 年 5 月，麦肯锡全球研究所发布了一份名为《颠覆性技术：技术改进生活、商业和全球经济》的研究报告，报告确认了未来 12 种新兴技术，而大数据是这其中需求技术的基石。

目前大数据技术发展可以分为六个方向：

1) 大数据采集与预处理方向。这个方向最常见的问题是数据的多源和多样

性，导致数据的质量存在差异，影响到数据的可用性。

2) 大数据存储与管理方向。这个方向最常见的挑战是存储规模大，存储管理复杂，需要兼顾结构化、非结构化和半结构化的数据。分布式文件系统和分布式数据库相关技术的发展正在有效的解决这些问题。其中大数据索引和查询技术、实时及流式大数据存储与处理的发展。

3) 大数据计算模式方向。如今出现了多种典型的计算模式，包括大数据查询分析计算、批处理计算，流式计算、迭代计算、图计算、内存计算。

4) 大数据分析挖掘方向。在数据类迅速膨胀的同时，还要进行深度的数据分析和挖掘，因此越来越多的大数据分析工具和产品应运而生。

5) 大数据可视化分析方向。通过可视化方式来帮助人们探索和解释复杂的数据，有利于决策者挖掘数据的商业价值，进而有助于大数据的发展。

6) 大数据安全方向。当我们在用大数据分析和数据挖掘获取商业价值的时候，黑客很可能在向我们攻击，收集有用的信息。通过文件访问控制来限制呈现对数据的操作、基础设备加密、匿名化保护技术和加密保护等技术正在最大程度的保护数据安全。

(1) 针对线上个人消费贷款新业务，评价中小商业银行在进行风险评估过程中可能存在的问题。

首先，线上个人消费贷款业务面临客户规模陡增，可能导致银行风险防控方面面临人手不足的困境，从而影响贷款效率与风险防控效果。

其次，“非面对面”的业务开展方式使得零售业务风险更加复杂，传统的客户身份识别方式、信用风险评估方法都已无法满足新形势下的风险防控需求，可能导致不良率的大幅上升。

(2) 并结合案例回答，项目组是如何从专家评分系统过渡到基于大数据的决策支持系统的？

基于传统专家系统在处理线上个人消费贷款新业务时存在的潜在问题，项目组利用大数据建立数据驱动评分系统，通过该系统与专家评分系统的综合运用，帮助识别出目前由专家评分系统无法识别的高风险客户，并且对专家评分有一定的纠偏作用。通过综合专家经验与大数据知识，从传统专家评分系统的“经验决策”过渡到基于大数据的决策支持系统的“数据决策”，如图3所示。

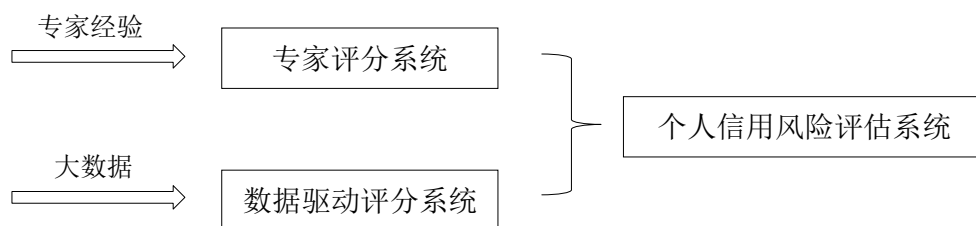


图 3 从“经验决策”到“数据决策”的过渡

Figure 3 Transition from “experience-based” to “data-based” decision

3. 案例中的项目团队在开发风险评估系统的过程中遇到了什么关键难题？他们是如何解决的？

本题共包含两个子问题，呈递进关系，需分步作答。第一个问题主要考察学生分析材料与思辨的能力，学生需要根据案例正文归纳总结出项目团队在实施项目的过程中遇到的关键问题；第二个问题考察学生的归纳总结能力。回答本题，学生需要了解机器学习的相关理论，包括：

朴素贝叶斯：是一种构建分类器的简单方法。该分类器模型会给问题实例分配用特征值表示的类标签，类标签取自有限集合。在本案例中用来针对样本构造标签，用以表征不同的特征。

K-means 聚类：是一种简单的迭代型聚类算法，采用距离作为相似性指标，从而发现给定数据集中的 K 个类，且每个类的中心是根据类中所有值的均值得到，每个类用聚类中心来述，聚类目标是使得各类的类间距离和最小。在本案例中用来找出异常样本，即违约样本。

稀疏自编码：是一种无监督机器学习算法，通过计算自编码的输出与原输入的误差，不断调节自编码器的参数，最终训练出模型。在本案例中可以用来构建标签数据与违约概率之间关系的模型。

全监督学习：利用一组已知类别的样本调整分类器的参数，使其达到所要求性能的过程，也称为监督训练或有教师学习。在本案例中用以构建标签数据与违约概率之间关系的模型。

半监督学习：(Semi-Supervised Learning, SSL)是模式识别和机器学习领域研究的重点问题，是监督学习与无监督学习相结合的一种学习方法。半监督学习使用大量的未标记数据，以及同时使用标记数据，来进行模式识别工作。在本案例中用以构建标签数据与违约概率之间关系的模型。

逻辑（logistic）回归：又称 logistic 回归分析，是一种广义的线性回归分析模型，常用于数据挖掘，疾病自动诊断，经济预测等领域。在本案例中用来对违

约概率进行预测。

(1) 案例中的项目团队在开发风险评估系统的过程中遇到了什么关键难题？

项目团队遇到的关键难题是数据样本不足，特别是违约样本不足。一方面，单家银行零售业务规模有限，业务开展周期较短，积累的业务数据量不足，导致了样本数据特别是违约样本数量不足，对以‘大数据量’训练为基础的深度学习技术提出了挑战。另一方面，商业银行普遍缺少系统性数据维护和管理机制，数据架构不完善，存在数据瓶颈制约，这些问题导致其部分数据项存在缺失，对后续数据积累和使用造成影响。

(2) 他们是如何解决的？

实践表明，大部分申请客户具有较低的违约概率，他们的解决方法是通过聚类分析，识别与大多数申请者不同的用户来进行网贷违约风险控制。

4. 线上个人消费贷款预授信系统开发出来之后，预授信额度如何确定，以及该系统在实际中应该如何运用？

本题为开放性问题，主要考察学生对行业的了解程度以及综合分析问题的能力。回答本题，学生需要了解优化相关的理论，包括：

线性规划 (Linear programming, 简称 LP)：研究线性约束条件下线性目标函数的极值问题的数学理论和方法。它是运筹学的一个重要分支，广泛应用于军事作战、经济分析、经营管理和工程技术等方面。为合理地利用有限的人力、物力、财力等资源作出的最优决策，提供科学的依据。

一般来说，系统要投入使用，需要由银行高层决定违约率与预期收益率，然后作为参数输入系统，即可求得相关的授信额度。使用到的具体模型如下：

首先根据据贷率拒绝若干评分最低的客户。然后在剩下的非拒贷客户中，假设第*i*个客户工资为 s_i ，违约概率为 p_i ，贷款年利率为 r_i ，响应概率为 re_i ，决策变量 x_i 为其工资倍数 ($L \leq x_i \leq U$)，且工资倍数为整数。假设银行总信贷规模为 T ，个人贷款上限为 A ，决策的目标为最小化期望损失，即有：

$$\begin{aligned} \max_{x_1, \dots, x_n} \quad & \sum_{i=1}^n [re_i * (1 - p_i) * s_i * x_i * (1 + r_i) - re_i * p_i * x_i * r_i] \\ \text{s.t.} \quad & \sum_{i=1}^n re_i * s_i * x_i \leq T \\ & L \leq x_i \leq U \quad \forall i = 1, \dots, n \\ & x_i \in \mathbb{Z} \quad \forall i = 1, \dots, n \\ & s_i * x_i \leq A. \quad \forall i = 1, \dots, n \end{aligned}$$

该模型从理论上阐述了不确定环境下的风险控制和决策。其中，第一条约束表示总的期望授信规模不超过总信贷规模；第二条约束表示最高和最低的工资倍数；第三条约束表示工资倍数为整数；第四条约束表示个人贷款额度不超过个人贷款上限。

在实践中，由于每个样本的响应概率难以获得，对此，我们在一定假设条件下，提出了风险控制更加严格的鲁棒预授信额度分配模型。根据客户的违约概率，把非拒贷客户从低到高按照升序排序，根据 L 和 U ，把样本均匀分为 K 类。其中， $K = U - L + 1$ ，集合 K_i 表示第 i 类客户样本集合，则第一类人群是风险最低人群。

假设每一类客户的响应率为 re_i ， $i = 1, \dots, K$ ，有

$$\begin{aligned}
 & \max_x \sum_{i=1}^K \tau_i \\
 \text{s.t. } & \tau_i = \min_{v_i} \sum_{j \in v_i} (1 - p_j) * s_j * x_j * (1 + r_j) - p_j * x_j * r_j \\
 & |v_i| \geq re_i * |K_i| \quad \forall i = 1, \dots, n \\
 & \sum_{i=1}^K \sum_j^{v_i} s_j * x_j \leq T \\
 & L \leq x_i \leq U \quad \forall i = 1, \dots, n \\
 & x_i \in \mathbb{Z} \quad \forall i = 1, \dots, n \\
 & s_i * x_i \leq A. \quad \forall i = 1, \dots, n
 \end{aligned}$$

其中，第一条约束表示每一类客户在最坏情况下的利润；第二条约束保证了每一类的响应人数；第三条约束最坏收益情况下的授信规模不超过总信贷规模；第四条约束表示最高和最低的工资倍数；第五条表示工资倍数为整数；第六条约束表示个人贷款额度不超过个人贷款上限

五、背景信息

1. 行业背景

当前，我国经济已由高速增长阶段转向高质量发展阶段，并且正处在转变发展方式、优化经济结构、转换增长动力的攻关期。金融市场深化改革、利率市场化加速推进都对中小商业银行的发展带来了严峻的挑战。除宏观因素外，监管层对系统性金融风险的从严防控也在客观上限制了中小商业银行过快的规模扩张，中小商业银行想要通过做大体量实现弯道超车的可能性逐步减小。为顺应新时代经济发展阶段转向，中小商业银行纷纷推出了金融服务提质增效、优化结构的“大零售”战略转型。随着中小商业银行在理财、贷款、支付等多个领域进行产品创

新，其借助互联网“传播效应”短期内获得大量零售客户，但“非面对面”的业务开展方式使得零售业务风险更加复杂，中小商业银行传统的客户身份识别方式、信用风险评估方法都已无法满足新形势下的风险防控需求。

在上述形势下，依托大数据、人工智能、云计算等技术发展起来的金融科技，为中小商业银行信用风险数据的收集、度量、评估提供了崭新的视角和方法，为中小商业银行业务创新乃至实施“大零售”战略转型提供了巨大的发展契机与可能性。

2. 中小商业银行零售客户信用风险评估的重要性

商业银行作为现代社会经济运转的枢纽之一，关系整个经济市场的稳定发展。同时，商业银行也是风险的聚集地，其作为经营和管理风险的企业，信用风险特别是零售客户信用风险评估是其风险管理的重要内容。

习近平总书记在中共中央政治局第十四次集体学习时强调：“切实把维护金融安全作为治国理政的一件大事。坚决守住不发生系统性金融风险底线。”银行业监管机构也愈发重视大数据技术在金融机构风险防控中的应用，《中国银行业信息科技“十三五”发展规划监管指导意见》中要求：“推进大数据应用，全面提升数据治理与数据服务能力”；《关于强化银行卡磁条交易安全管理的通知》（银办发【2017】120号文）要求：要“建立基于大数据技术的风险防控机制”、“完成基于大数据技术的银行卡风险防控系统建设”；在《关于印发银行业金融机构数据治理指引的通知》（银保监发【2018】22号）中，更是明确要求：“银行业金融机构应当加强数据应用，持续改善风险管理方法，有效识别、计量、评估、监测、报告和控制各类风险。”

随着近年来去产能、去杠杆、去库存任务的不断进行，加之利率市场化基本落地及互联网金融公司的拓展，“金融脱媒”现象日益明显；众多中小商业银行纷纷借助金融科技，在理财、贷款、支付等多个领域与互联网金融展开正面竞争。客户规模陡增、“非面对面”的业务开展方式，使得零售业务风险更加复杂，上述形势均对中小商业银行的风险防控能力出了更大的挑战，根据2018年中小商业银行披露的年报数据显示，全国多省份的中小商业银行不良率呈现上升趋势。同时，国内中小商业银行虽然单家资产规模较小，但由于中小商业银行数量众多，总体规模巨大。以山东省城商行联盟为例，其运行的线上资产规模已突破1.7万亿。由于中小银行区域化特征明显，同一区域内的中小银行容易产生连锁反映，导致系统性风险。因此，加强中小商业银行零售客户信用风险评估，不仅是中小商业银行作为金融主体对银行自身内部信用风险管理的自主行为，从宏观层面也

是防范银行业信用风险导致银行体系崩溃、产生系统风险、引发金融危机的需要。

3. 中小商业银行零售客户信用风险评估面临的问题

近年来，中小商业银行对信用风险管理的重视程度大大加强，遵循银行业监管机构要求，在信贷政策制定、信用风险评估、授权审批、信贷责任等方面投入了大量工作，但中小商业银行由于起步较晚，在风险管理方面与国有商业银行、股份制商业银行相比依然有较大差距，目前面临的问题也比较多，主要体现在以下几个方面：

（1）信用风险评估创新不足，“同质化”严重

长期以来，中小商业银行出于政策和经济目标考虑，信贷额度集中向资产优良的大中型对公客户倾斜，导致其面向零售客户的信用风险控制力度明显不足。近几年，中小商业银行响应国家“普惠金融”政策号召，逐步发力零售客户，面对“互联网+”时代大量的零售客户，中小商业银行信用风险评估方法已无法满足新时期业务发展的需求。由于信用风险评估管理体系建设过程复杂，中小商业银行普遍采用咨询项目、外包建设等方式引入。而目前业界主流信用风险评估解决方案仍为基于“规则引擎”的“评分卡”模式，中小商业银行由于业务周期较短、缺乏经验积累等原因，评分规则多为依靠专家经验部署或业界基础规则，“同质化”现象严重，最终得出的客户信用评级缺乏客观性，无法满足众多中小商业银行的差异化需求。随着大数据、深度学习技术的发展，业界已开始重视基于大数据技术的信用风险评估方法的研究，但现阶段普遍以理论性学术研究为主。由于缺乏数据基础支撑，无法保证模型精准性，目前已落地的研究成果较少，模型有效性无法得到充分验证。

（2）数据样本不足，特别是违约样本不足

中小商业银行体量和数据运营机制决定了其数据样本量特别是违约样本量不足。一方面，单家中小商业银行零售业务规模有限，业务开展周期较短，积累的业务数据量不足，导致了样本数据特别是违约样本数量不足，对以“大数据量”训练为基础的深度学习技术出了挑战。另一方面，中小商业银行普遍缺少系统性数据维护和管理机制，数据架构不完善，存在数据瓶颈制约，这些问题导致其部分数据项存在缺失，对后续数据积累和使用造成影响。

（3）专业人才储备不足

信用风险管理需要大量的专业人员参与数据的收集、整理、模型建立、运用、结果输出、验证等工作，而中小商业银行员工人数较为精简，人员设置无法涵盖经济学、数学、金融学以及统计学等多个学科人才。中小商业银行由于所处城市

产业布局、发展速度、制度体系建设等方面制约了人才流入，同时受限于自身规模、体制等原因，专业人才吸引能力不足，短期内难以完成风险评价体系人才的梯队建设。

六、关键点

本案例描述了一个真实的中小商业银行构建基于大数据的个人信用风险评估系统的整个过程，重点涉及在面临违约样本缺失的情形下，无法提取违约标签时，项目组成员如何巧妙地利用大数据进行违约标签的提取，以及如何充分利用专家评分系统与数据驱动评分系统的优势，构建最小化风险的预授信模型。

本案例涉及的关键知识要点和应培养的能力点包括：

1. 中小商业银行进行信用风险评估时传统的专家评分法与数据驱动方法的核心思想。重点明确专家评分法与数据驱动方法的不同之处，以及如何进行优势互补，从而构建出专家经验与大数据相结合的个人信用风险评估系统。培养学生从案例反映的问题中发现单纯采用“经验决策”的不足之处，了解大数据的作用，以及采用大数据进行决策分析的必要性及可行性。

2. 利用大数据进行个人信用风险评估时的关键难点。从案例的现状描述中，总结“数据决策”实施过程中的难点和应注意的问题，认识到基于大数据的决策支持系统对数据样本的要求，以及当数据足够多时，系统将会得到更好的效果。

3. 基于大数据的个人信用风险评估所涉及的关键过程。引导学生分析当缺少违约样本时，如何通过大数据分析构建违约标签，并进一步了解如何从违约标签的构建到违约概率的预测与授信额度的确定等一系列过程。培养学生利用大数据进行风险评估的基本思路与方法路线。

4. 大数据相关理论的应用，主要包括机器学习的相关理论。引导学生了解稀疏自编码、朴素贝叶斯、K-means 聚类、半监督学习、全监督学习等相关方法的适用情景与应用。

七、建议课堂计划

本案例可以作为专门的案例讨论课来进行。如下是按照时间进度提供的课堂计划建议，仅供参考。

整个案例课的课堂时间控制在 90 分钟内。

课前计划：

提前发放案例，请学员自行完成阅读并根据案例的启发思考题进行归纳与思

考，拟出存疑点与需要进一步探讨的地方。

课中计划：

1、根据案例使用说明中的背景知识介绍目前中小商业银行在风险评估方面的行业现状，为学员更深入地理解案例做好铺垫（5 分钟）

2、通过提问与讲解相结合的方式引导学生梳理案例中存在的决策点（20 分钟）

3、针对案例中提出的解决方案进行分组讨论，每组提出存疑点与需要进一步探讨的问题（15 分钟）

4、教师提问并邀请小组发言，过程中教师从旁点评并针对案例中涉及的关键理论知识进行串联与讲解（30 分钟）

5、教师对案例中的关键点进行回顾与总结，在此过程中可适当结合提问与讨论（20 分钟）

课后计划：

根据课堂讨论，进一步完善小组发言稿，以小组为单位提交案例分析报告。

八、相关附件

本案例包含如下附件：

附件 1. “百篇优秀管理案例”评选公函

附件 2. 案例授权书的说明及《企业证明书》

附件 3. 作者授权书

（案例使用说明字数：11320）