

ICS 03.060

CCS A 11

# 团体标准

T/NIFA 31-2024

## 数字信用评分指引

Guidance for digitalized credit scoring

2024-11-28 发布

2024-11-28 实施

中国互联网金融协会 发布



## 目 次

前言 .....	II
引言 .....	III
1 范围 .....	1
2 规范性引用文件 .....	1
3 术语与定义 .....	1
4 基本原则 .....	2
4.1 合规性 .....	2
4.2 公平性 .....	2
4.3 客观性 .....	2
4.4 一致性 .....	2
4.5 稳健性 .....	2
4.6 可靠性 .....	2
4.7 可持续性 .....	3
4.8 可解释性 .....	3
5 数字信用评分的数据基础 .....	3
5.1 数据分类 .....	3
5.2 数据处理要求 .....	3
6 数字信用评分的研发与管理 .....	4
6.1 概述 .....	4
6.2 评分模型的设计 .....	4
6.3 数字评分模型的开发 .....	5
6.4 模型实施 .....	9
6.5 模型追踪与迭代 .....	10
7 数字信用评分的应用 .....	10
附录 A（资料性）信用评分应用示例 .....	12
参考文献 .....	14

## 前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》和GB/T 20004.1—2016《团体标准化 第1部分：良好行为指南》给出的规则起草。

本文件由中国互联网金融协会提出。

本文件由中国互联网金融协会归口。

本文件起草单位：中国互联网金融协会、建信金融科技有限责任公司、百行征信有限公司、中国人民大学、中金金融认证中心有限公司、中国光大银行股份有限公司、天翼征信有限公司、江苏省联合征信有限公司、重庆富民银行股份有限公司、通联支付网络服务股份有限公司、中国投融资担保股份有限公司、百融云创科技股份有限公司、上海华瑞银行股份有限公司、北京国家金融科技认证中心有限公司、京东科技控股股份有限公司、中电金信软件有限公司、北京微财科技有限公司、武汉众邦银行股份有限公司、阿里云计算有限公司、腾讯控股有限公司。

本文件主要起草人：单强、杨农、王新华、刘志玲、司劲萱、邓康、李相宽、田昆、黄殷、黄丹阳、吴宝民、顾学靖、苗森、杨心翻、叶海鹤、王仲翼、陈康、王艳耘、李钦、张家铭、陈少荣、邵峻澜、颜欣、吴明新、李倩、张海燕、高天游、孔庆龙、杨春、梁嘉文、李锦秋、朱俊、廖军、张威、李克鹏。

## 引 言

数字信用评分是以数据信息为基础，从历史信息中挖掘与未来信用表现之间的关系，运用数据挖掘分析、统计处理以及信用评分建设流程，对信息主体的信用或类信用数据进行统计分析的过程。数字信用评分旨在预测和评估信用主体未来的信用风险等情况，是金融机构进行量化决策的重要工具。征信作为数字金融信贷风控管理的重要部分，为金融业发展提供了动力，亦是金融业高质量发展的必然要求。近年来征信业和金融科技携手利用大数据、人工智能等技术推动征信系统升级，提供多样化信用评估服务，助力金融业数字化转型。编制本文件主要目的是为了加强各类相关机构的数字信用评分建设能力，指导各相关机构规范信用评分开发流程。



# 数字信用评级指引

## 1 范围

本文件提供了采用数字化手段构建数字信用评级的基本原则、数据基础、研发与管理、应用等方面的指引。

本文件适用于对法人和非法人组织（以下统称企业）、个人开展数字信用评级工作的机构。

## 2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 22117—2018 信用基本术语

GB/T 35273—2020 信息安全技术 个人信息安全规范

## 3 术语与定义

下列术语和定义适用于本文件。

### 3.1

#### 信用 credit

个人或组织履行承诺的意愿和能力。

注：在经济领域，信用的含义等同于交易信用，是指交易各方在信任基础上，不用立即付款或担保就可获得资金、物资或服务的能力。这种能力以在约定期限内履约为条件，并可以使用货币单位直接度量。

[来源：GB/T 22117—2018]

### 3.2

#### 信用主体 subject of credit

参与信用活动的个人或组织。

[来源：GB/T 22117—2018]

### 3.3

#### 公共信用信息 public credit information

依法行使公共职能的部门履职过程产生的有关各类市场主体的信用信息。

注：依法行使公共职能的部门包括行政机关、司法机关以及依法行使公共管理（服务）职能的企、事业单位和社会组织。

[来源：GB/T 22117—2018]

### 3.4

#### **征信 credit reporting**

采集、整理、保存、加工个人或组织的信用信息，并向有合法需求的信息使用者提供信用信息服务，帮助市场主体判断控制风险、进行信用管理的活动。

[来源：GB/T 22117—2018]

### 3.5

#### **征信机构 credit reporting agency**

经营征信业务的机构。

[来源：GB/T 22117—2018]

### 3.6

#### **信用评分 credit scoring**

根据信息主体的信用信息，运用统计和其他方法，建立信用评分模型，对信用主体的信用进行评价，并用分数的形式表现出来的活动。

注：信用评分一般包括通用评分和定制化评分。

[来源：GB/T 22117—2008]

### 3.7

#### **个人信息 personal information**

以电子或者其他方式记录的能够单独或者与其他信息结合识别特定自然人身份或者反映特定自然人活动的各种信息。

[来源：GB/T 35273—2020]

## 4 基本原则

### 4.1 合规性

信用评分在数据的采集、处理、加工、模型的开发及评分结果输出等过程中，都需严格遵守相关法律法规的要求。

### 4.2 公平性

信用评分旨在对信用主体提供公正的评价标准。

### 4.3 客观性

信用评分通过统计模型避免受外界主观因素影响来客观地反映信用主体的信用风险水平。

### 4.4 一致性

信用评分对于信用信息相同的主体给出相同的预测结果。

### 4.5 稳健性

信用评分在不同的时间和样本上都能表现出较好的排序能力和区分能力。

### 4.6 可靠性

信用评分使用的数据经过信息提供者审查以提高来源和质量的可靠。

#### 4.7 可持续性

信用评分从数据的获取到评分结果输出的过程中是稳定且可持续的。

#### 4.8 可解释性

信用评分结果及入模变量符合业务逻辑，可提供清晰可靠的解释。

### 5 数字信用评分的数据基础

#### 5.1 数据分类

##### 5.1.1 概述

所用数据信息通常分为信用类数据和可选数据。

##### 5.1.2 信用类数据

信用类数据指与信用消费相关的数据，通常产生于银行、消费金融公司和小额贷款公司等机构。信用类数据包括但不限于：

- a) 贷款还款记录；
- b) 未结清余额；
- c) 还款期数；
- d) 负债比率；
- e) 信贷组合。

##### 5.1.3 可选数据

可选数据指除信用类数据外其他可以用于评估、预测借款人信用状况的信息。对于没有传统信用记录的人群，可选数据也可以成为判断其信用风险的依据，将其用于构建数字信用评分，可以进一步扩大金融服务的范围。可选数据来源于多个维度，包括政府的公共信用信息以及市场化机构在开展业务过程中产生的信息等。可选数据包括但不限于：

- a) 支付信息；
- b) 位置信息；
- c) 社交媒体信息；
- d) 运营商信息；
- e) 网络平台信息；
- f) 社保信息；
- g) 税务信息；
- h) 司法信息；
- i) 公积金信息；
- j) 水、电、气缴费信息；
- k) 知识产权信息。

#### 5.2 数据处理要求

数据的处理包括采集、存储、使用、加工、传输、提供、公开、删除等，数据处理中注意事项包括

但不限于：

- a) 不得欺骗、胁迫、诱导，不得向信息主体收费，不得通过非法渠道或者其他侵害信息主体合法权益的方式采集信息；
- b) 征信机构与信息提供者在开办业务及合作中应当通过协议等形式明确信息采集的原则以及各自在获得客户同意、信息采集、加工处理、信息更正、异议处理、信息安全等方面的权利义务和责任；
- c) 征信机构采集个人信用信息应当经信息主体本人同意，并且明确告知信息主体采集信用信息的目的，依照法律法规公开的信息除外；
- d) 征信机构通过信息提供者取得个人同意的，信息提供者应当向信息主体履行告知义务；
- e) 采集企业信用信息，应当基于合法的目的，不得侵犯商业秘密；
- f) 处理个人信息应当遵循合法、正当、必要和诚信原则，不得通过误导、欺诈、胁迫等方式处理个人信息；
- g) 处理个人信息应当具有明确、合理的目的，并应当与处理目的直接相关，采取对个人权益影响最小的方式；
- h) 采集个人信息，应当限于实现处理目的的最小范围，不得过度采集个人信息；
- i) 处理个人信息应当保证个人信息的质量，避免因个人信息不准确、不完整对个人权益造成不利影响；
- j) 个人信息处理者应当对其个人信息处理活动负责，并采取必要措施保障所处理个人信息的安全；
- k) 个人不良信息保存期限5年届满后，征信机构应当将个人不良信息在对外服务和应用中删除；作为样本数据的，应当进行匿名化处理。

## 6 数字信用评分的研发与管理

### 6.1 概述

信用评分是研发和管理流程包括评分模型的设计、评分模型的开发、评分模型的实施以及评分模型的监控与迭代等步骤。

注：评分模型是一种广泛用于风险管理和信用评估的工具，主要通过对一系列相关变量进行加权求和或通过其他数学运算，最终生成一个数值或分数来代表某个实体（如个人或企业）的总体风险水平。这种模型是一种定性分析和定量分析相结合的综合评价模型。

### 6.2 评分模型的设计

#### 6.2.1 建模目的

评分模型设计阶段首先需要明确建模的目的及对应的模型类别（括号内为建模目的对应的模型类别），建模目的包括但不限于预测以下情况：

- a) 按期履约还款（风险模型）；
- b) 产生的实质性收入（收入模型）；
- c) 有较高概率追回当前欠款（催收模型）；
- d) 偿还较高比例的核销款（回收模型）；
- e) 阻止客户流失（账户流失模型）；
- f) 恶意骗贷（欺诈模型）；
- g) 欺诈交易的实施（欺诈模型）；
- h) 产品偏好（偏好模型）；

- i) 消费者对市场营销活动的响应（市场反应模型）；
- j) 宣布破产（破产模型）。

### 6.2.2 建模考虑因素

模型设计需要考虑的因素包括但不限于：

- a) 模型性质，例如属于风险预测、收入预测或其他类型；
- b) 数据来源，例如征信机构数据和其他数据；
- c) 应用领域，例如贷款发放审批业务、账户管理等；
- d) 数据日期，例如表现、观察和表现窗口的时长等。观察窗口用于获取样本行为特征。观察点通常是一个时间点，从这一点后的样本表现被称之为“未来行为”。在这一点或这一点之前，每个观察样本的信息用于产生模型开发用的预测变量，或称自变量。在样本数量较少的情况下需要一个滚动窗口而非一个观察点以保证充分数量的样本可用于评分开发。表现窗口指可以观察到预先定义好坏的期间，在表现时点客户被分类为好/坏/不确定；
- e) “好”和“坏”的定义，例如放款后12个月内发生90天及以上逾期定义为“坏”；
- f) 最低评分条件，又称“排除规则”，例如已死亡人群；
- g) 关于法律遵从及接受度的考虑，例如不使用可能涉及歧视的变量；
- h) 分组，例如根据账龄、逾期等进行分组；
- i) 建模技术，例如逻辑回归，XGBOOST等；
- j) 评分最终分数的延展方式，例如0-1000分。

## 6.3 数字评分模型的开发

### 6.3.1 概述

评分模型开发流程具体包括：

- a) 取样；
- b) 数据抽取与准备；
- c) 数据分析与处理；
- d) 表现、预测及控制变量的生成；
- e) 构建模型开发数据库并准备建模；
- f) 分组分析；
- g) 模型构建；
- h) 评分延展与原因码生成；
- i) 模型验证。

### 6.3.2 抽样

在建模数据样本量足够大的情况下，根据大样本的统计定理和中心极限理论，可以使用部分样本代替全体数据。抽样可以节省数据成本和大型数据处理成本，减少程序执行的时间，从而延长模型开发的时间。

抽样后需进行权重计算并赋值，权重表示每个信贷个体的样本比例，以保证样本真实反映原始数据。取样前每个样本权重设定为1。

### 6.3.3 数据抽取与准备

充分了解相关数据系统，数据信息来源、信息量及数据的可获取性，设计数据抽取模板并进行数据抽取。包括但不限于：

- a) 抽取样本中所有可获取记录;
- b) 抽取所有数据项;
- c) 数据项标准化;
- d) 所有数据项的分布, 并参照数据词典进行检查;
- e) 对数据项的分布和均值作质量检查;
- f) 对在数据字典中不存在的值进行调查。

#### 6.3.4 数据分析与处理

数据分析与处理通过对数据深入分析和研究, 分析其是否满足建模需求, 并对数据进行汇总整合, 包括但不限于:

- a) 充分理解数据: 梳理数据表结构和字段, 充分理解每个数据表内的每个变量, 并结合实际, 理解每个变量的格式、含义、定义的逻辑、标准及每个变量内各个符号和代码表示的意义, 并为每个变量添加数据标签;
- b) 原始数据检测: 通过对现有数据分析了解当前数据状况。变量可分为数值型、字符型、日期型以及关联键四类变量, 针对不同类型的变量采取不同分析的方法, 比如:
  - 原始数据时间段检查;
  - 主键唯一性检查;
  - 缺失值分析;
  - 原始数据统计量分析等。
- c) 数据关联分析: 梳理表间逻辑关系, 通过关联键将各表进行关联, 观察数据匹配程度和关联情况, 加强对数据系统化和结构化的理解。

#### 6.3.5 表现、预测及控制变量的生成

进行变量开发应对前期数据准备结果进行复审确保数据加工逻辑和准确性。变量开发包括开发表现变量、预测变量和控制变量。

##### 6.3.5.1 表现变量的生成

开发表现变量需明确表现定义:

- “好”样本为表现期内守信且有较好的还款记录, 放贷机构愿意放贷的账户;
- “坏”样本为表现期内失信且逾期较严重, 放贷机构了解已报告账户后可能拒绝放贷的账户;
- 不确定样本为介于“好”和“坏”的中间状态;
- 时间不足与排除: 包括符合“良好”资格, 却因在档时间不足而无法归入“良好”的已报告账户、处于争议状态的账户和报告中无账户的样本等。

表现定义可以结合滚动率确定。

注: 滚动率分析是指观察不同逾期状态的账户其逾期状态降低、增多或回归正常的比例, 以此作为好坏定义的参考依据。如C-M1滚动率, 指逾期1个月内的账户余额占正常状态余额的比率。

##### 6.3.5.2 预测变量的生成

预测变量指所有可能用于预测样本表现的变量, 信用相关的常用预测变量类型包括但不限于:

- 付款历史, 关键因素包括最近一次逾期距今时间、最差逾期天数等;
- 未偿债务, 关键因素包括欠债总额、额度使用率、未偿债务占总贷款金额百分比等;
- 信贷时长, 关键因素包括开户时长、各账户开立距今平均月份等;
- 近期信贷需求, 关键因素包括最近问询次数、近期新开账户数等;

- 信贷组合，关键因素包括信用卡数目、分期类贷款占总信贷产品比例等；

### 6.3.5.3 预测变量的生成

控制变量指在控制所研究的解释变量和被解释变量之间的关系时，需要控制的可能影响的变量。包括但不限于：

- 数据匹配值（关联变量）；
- 样本权重变量；
- 用于排除逻辑的变量，如死亡标记变量；
- 用于验证目的地，如地理标记变量；
- 用于数据和匹配质量评估的变量；
- 用于区分建模数据和测试数据的标记变量等。

### 6.3.6 建模样本准备

将表现变量、预测变量及控制变量进行合并，通过变量分布和多变量交叉分布等统计分析对表现变量及预测变量进行重审并评估匹配质量，基于宽表对变量进行转换，转换方式包括连续型转换和分箱转换。

分箱转换根据样本表现对特征进行分组，可以较好地拟合建模样本从而提高模型表现，同时通过分箱可以处理趋势不单调的变量。分箱转换根据编码方式不同又可以分为证据权重（Weight of Evidence，下称“WOE”）转换和哑变量转换。连续型转换通过对原始变量进行平方、平方根、立方根、倒数和自然对数函数等数学转换拟合特征与预测变量间的关系。

#### a) WOE转换：

对每个变量，按照其各个分组的 WOE 值进行转换，WOE 计算公式如下所示：

$$WOE = \ln \left( \frac{f_G(j)}{f_B(j)} \right), f_G(j) = \text{Good\% that fall into bin } j, f_B(j) = \text{Bad\% that fall into bin } j$$

注： $f_G(j)$  为分箱内好样本占全部好样本的比例， $f_B(j)$  为分箱内坏样本占全部坏样本的比例

对于变量取值为连续型的变量，其细分组基本方法为：将变量按照其取值排序，细分为不超过 20 组，每组取值约占 5% 的样本数，分析每个分组的好、坏、不确定账户数及占比，计算每个分组的 WOE。在获得备选变量的细分组后，检验变量 WOE 在细分组情况下的变化趋势是否单调，从而确保该变量具有良好的区分能力，可通过先细分，再粗分合并来实现，基本方法为将 WOE 接近的相邻多个细分组进行合并以消除波动趋势。

#### b) 哑变量转换：

哑变量又称虚拟变量或虚设变量，在统计学和计量经济学中，是一种特殊类型的变量，它用于代表分类数据或者定性特征，并将其转化为数值形式以便在定量分析模型（如回归分析）中使用。哑变量转换需要对各变量建立二元的哑变量，通常取值为 1（是）和 0（否）。建立二元哑变量又分为两种情况，一类是对分类型变量生成哑变量，另一类是对连续型变量生成哑变量。对于分类型变量，例如婚姻关系包含多个变量属性：未婚、已婚、离异、丧偶等，可直接将每个属性作为一个分析，相对应生成一个二元哑变量。对于连续型变量，若变量的不同取值较少，可直接按照不同取值进行分组，比如“最近 3 个月的最大逾期月数”取值为 0、1、2，则可以直接按照不同取值进行分箱，否则将借助卡方检验的统计方法来确定分箱，并生成二元哑变量。

#### c) 连续型转换：

连续型转换使用各种对数和幂函数对变量进行处理，具体步骤包括：

- 1) 缺失值处理：连续型变量需要先进行缺失值处理。缺失分成“虚假缺失”和“真实缺失”。“虚假缺失”是指该账户不应有信息缺失，但特殊的业务含义需要设定一个特殊值代表缺失。然而，一些缺失是由于数据本身导致的真实缺失。对于真实缺失可应用下列的推导方

法：首先检查和识别有缺失值的变量，然后分配一个与真实值呈现相同的坏帐率的新值。为保证模型的更准确性，应在每个子模型进行缺失值推论：

- 2) 上下限处理：通过应用设定上下限去除极端的异常值，比如数字型变量的上限定在 99%分位数，下限定在 1%分位数；
- 3) 使用各种对数和幂函数进行变量的变形，确定每个连续型变量的最佳变形。对每个给定变量的每种变形对应变量进行相关性计算。变形顺序如下：原始变量、对数、平方根、平方、立方根、倒数。算法首先基于皮尔逊相关系数（Pearson Correlation Coefficient）选择与因变量关系最紧密的变形。上述描述仅适用于最初的变量降维，在建模阶段也需评估使用未变形变量 K-S 值（Kolmogorov-Smirnov）的差异以及不同变形形式的模型稳定性。如果备选变量中变形变量和未变形变量之间区别不大，且如果使用未变形变量没有反作用（如：损失稳定性或 KS），则通常使用未变形变量。

### 6.3.7 分组分析

不同人群之间的行为差异导致预测变量的选择可能不同，因此一个群组的给定变量可能比另一群组有更强的预测力。对不同的组群，宜采用不同的变量组建评分模型，从而挖掘各个群组预测规律的异同，进而提高针对各个子群而建的评分模型的预测力。

分组分析的目的是为了找出最佳分组方案从而使整个系统的预测力最大化。与单模型相比，分组将选择不同的方案，并且每个方案都可能提高评分模型系统的总体表现。模型分组分析包括生成分组方案和分组方案的评估与确立。分组的生成既需要基于过去经验的判断，也需要基于数据统计研究。

分组方案生成后需要通过主模型/子模型（Master/Niche）对各分组方案进行评估和选择，具体步骤包括：

- a) 开发基于总体的主模型；
- b) 评估母体模型在各个子群上的预测能力；
- c) 分别开发基于每个子群的模型；
- d) 评估各个子群在该子群上的预测能力；
- e) 在各个子群上，比较子群模型对于母体模型预测力的改进程度；
- f) 将子模型生成的评分校准为同样的几率-评分比，并且将校准后的评分汇合；
- g) 比较汇总后的子群模型评分相对于母体模型评分对于总体样本预测能力的改进程度。

分组方案评估需要考虑的因素包括但不限于：

- a) 分组数目；
- b) 样本大小；
- c) 各子群坏账率的差异。

### 6.3.8 模型构建

模型构建时，重点考虑因素包括但不限于：

- a) 对变量趋势和转换方式进行研究；
- b) 在考虑变量相关性的基础上分析找出重要变量，除了变量重要性外，应对不同类别变量进行平衡与考量。变量降维过程决定这些变量的哪些子集最适合预测样本表现。变量降维常用步骤：
  - 1) 缺失值和单一值筛选：特征发散性低的变量，包括缺失率过高或者单一值占比较高的变量的预测能力较弱，可以设定阈值进行先行剔除；
  - 2) 信息值（Information Value，下文简称“IV”）筛选：通常信息值大于0.02则具有统计意义，IV值小于0.02的变量可以先行剔除，需要注意的是，这个0.02的阈值不是绝对的，对于一些原本区分能力就不会太强的模型，阈值可考虑适当下调；

- 3) 趋势筛选：在进行单变量分析时，在细分组和粗分组之后，再结合业务理解进行变量趋势筛选，确保变量各分箱坏账率趋势应和业务理解一致；
  - 4) 逐步判别：目的是从变量池内所有可能的备选变量中确定最具预测能力的特征变量组合；从模型中没有变量开始，逐个引入变量。每增加一个变量后都要检查，去除那些无助于模型预测能力的变量。
- c) 确认评分卡中变量得到合理的权重，包括确认特殊值的权重以确保模型稳定耐久性和合理原因码的生成；
  - d) 建模是一个反复迭代的过程，一个优质的评分卡在迭代建模产生；
  - e) 评分建设需分析师积极参与，包括审查、解释与调整；
  - f) 分析师应富有行业经验。

### 6.3.9 评分延展与原因码生成

信用评分模型最终以评分的形式输出，应该把基于不同样本建立的评分卡分数进行校准，从而使分数具有相同的意义，并使评分落在一定的范围，如0~1000分。

征信机构通常返回至少四条原因码，可根据与最大分值的差距选择和排序而生成原因码，或其他可行的方式产生。

### 6.3.10 模型验证

模型的验证是要基于独立的样本检测验证已开发模型的各项评估标准，包括入模变量的稳定性、入模变量的业务含义和趋势、评分分布的稳定性、评分的排序能力、评分的单调性等。包括但不限于：

- a) 验证维度：模型验证需要证实从多个特定的角度而言都是有效的，比如：
  - 1) 样本内（in-time）/样本外（out-time）；
  - 2) 从各个行业的角度；
  - 3) 从各个应用领域，如新客户、现有客户等；
  - 4) 对特定的放贷机构客户等。
- b) 验证数据：除基于开发数据的验证外，还应采集不同于开发数据（包括观察和表现）的新数据，并用这些更新、更独立的样本来验证已开发的模型。对于特定的放贷机构的账户，通过合并观察日征信机构数据与信贷机构客户的表现数据，来验证已开发的模型对于该机构样本是否有效；
- c) 模型稳定性：可通过群体稳定性（Population Stability Index, PSI）验证，一般认为PSI小于等于0.1时，评分没有变化或很少变化；PSI大于0.1且小于等于0.25时，评分有变化且应继续监控后续变化；PSI大于0.25时评分发生较大变化，应进行特征项分析；
- d) 模型预测力度：可以用过K-S和ROC等指标验证。K-S测量好坏累计分布间的最大距离点，应用较广泛且易于视觉理解。局限点在于当K-S远离操作点时最大化，则K-S检验可能毫无意义，且当数据不足时测量可能不准确。ROC的优势在于，跨总体权衡曲线的比较很有用处，如果权衡曲线彼此交叉，则有助于说明哪个模型整体上更加有效，局限点在于在特定的操作门槛可能无法显示哪个模型更好。

## 6.4 模型实施

模型实施阶段需要考虑输入、评分模块的构建、速度以及输出。包括但不限于：

- a) 输入：所有实施环境（在线，批次）中的生产数据结构与开发数据结构相同，并将所有的数据，甚至包括尚未使用的数据都作为输入导入，因为未来更新时可能用到所有这些数据；
- b) 评分模块的构建：所有实施环境（在线，批次）中评分计算的处理应在一个评分模块中完成，以便日后更新在这个模块内全部完成；

- c) 速度：速度是模型实施的关键，宜在一秒之内生成评分并返回信用报告；
- d) 输出：应开发两种不同的输出，即交付给评分用户的生产输出以及包括内部评分的特征值和组件权值的测试布局，应注意测试布局不向评分用户敞开。

## 6.5 模型追踪与迭代

模型上线投入使用后，应对模型定期进行追踪，检测总体或某个子群体的表现是否和预期一致。需要进行模型追踪的原因包括合规检查、模型风险管理，以及发现评分优化机会，提高管理准确度等。

- a) 关键的前端追踪报告包括：
  - 1) 总体稳定性报告；
  - 2) 特征分析报告；
  - 3) 最终评分报告。
- b) 关键的后端追踪报告包括：
  - 1) 详细逾期报告-最长逾期；
  - 2) 详细逾期报告-目前逾期；
  - 3) 账龄分析 (Vintage Analysis)；
  - 4) 逾期变更矩阵等。
- c) 通过跟踪和监测评分模型的表现，分析已经发生的变化可以判断是否需要评分模型进行迭代优化，需要进行模型迭代的原因包括但不限于：
  - 1) 信用主体行为发生变化；
  - 2) 信用信息或特征改进；
  - 3) 信用评价技术提高。

## 7 数字信用评分的应用

通用信用评分，可以应用在信贷生命全周期的各个阶段，结合金融机构内部的自有评分进行决策。

- a) 信用评分可应用的决策阶段包括但不限于：
  - 1) 客户筛选：通过评分进行客户分层（参见表A.1），筛选出目标和潜在客户，可以降低未来逾期率和营销进件成本，通过提高响应率进而提高盈利能力；
  - 2) 申请审批：根据评分结果预测客户风险，采取通过/拒绝/人工审核等决策，可以管理进件成本，简化决策流程，提高账户质量并控制逾期；
  - 3) 账户管理：根据内部和外部风险情况优化客户服务模式，寻找业务拓展机会，依据评分结果可进行客户分层，优化管理策略，促进对最优客户的积极处理，如增加额度等，提高优质客户的挽留率，同时对未来可能发生的逾期进行早期预警以减少未来损失；
  - 4) 催收排序：对于已逾期的客户通过信用评分预测未来还款概率，对具有不同还款概率的客户使用不同的催收手段，最大化催收效益；
  - 5) 交叉销售：参考评分结果对客户还款能力和意愿进行评估，根据客户的不同风险水平交叉销售不同的产品，增强客户粘性，提高收益。
- b) 信用评分的应用方式包括但不限于：
  - 1) 冷启动：对于尚处在新业务产品或数据积累阶段的机构，由于其客群风险不明确，且尚不满足分析或量化评分体系建设的基本要求，在获取用户授权同意的基础上，可直接使用信用评分作为业务决策依据进行客户分层，或参考专家以及行业经验，做出判断规则和处理逻辑；
  - 2) 决策变量：作为机构内决策变量所需的“外部来源”，金融机构可以将信用评分作为决策变量，建立二维交叉矩阵（参见表A.2），或者与内部评分结合建立决策树（参见图A.1），提升决

策的精准度；

3) 融合模型特征（参见图A.2）：作为机构内融合评分建设所需特征属性的“外部来源”，作为变量参与机构内部评分建设，提升内部评分区分能力。

全国团体标准信息平台

附 录 A  
(资料性)  
信用评分应用示例

客户分层示例如表A.1所示。

表 A.1 客户分层示例

信用评分	[0, 200]	(200, 400]	(400, 600]	(600, 800]	(800, 1000]
等级	A	B	C	D	E
风险级别	高风险	中高风险	中风险	中低风险	低风险

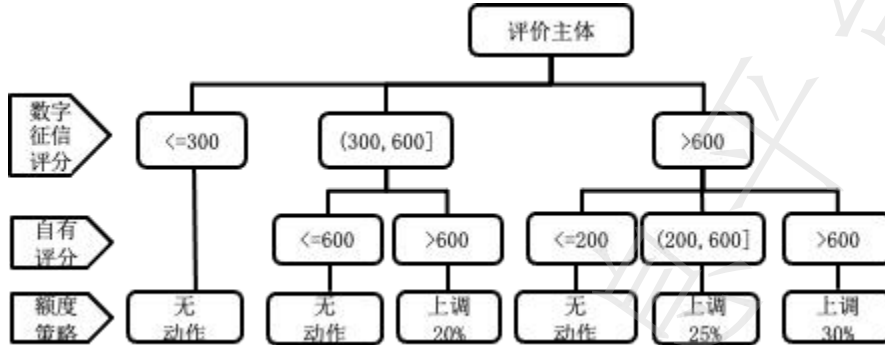
二维交叉矩阵示例如表A.2所示。

表 A.2 二维矩阵示例

	自有评分			
征信 评分	[0, 250]	(250, 500]	(500, 750]	(750, 1000]
[0, 250]	建议拒绝	建议拒绝	人工审核	人工审核
(250, 500]	建议拒绝	人工审核	人工审核	建议通过
(500, 750]	人工审核	人工审核	建议通过	建议通过
(750, 1000]	人工审核	人工审核	建议通过	建议通过

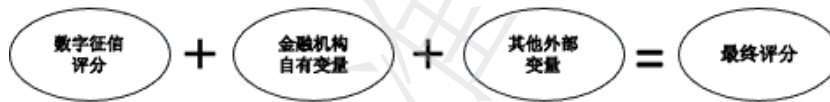
决策树示例如图A.1所示。

图 A.1 决策树示例



融合评分示例如图A.2所示。

图 A.2 融合评分示例



### 参 考 文 献

- [1] GB/T 19001—2008 质量管理体系 要求
  - [2] GB/T 22117-2018信用 基本术语
  - [3] GB/T 35273—2020 信息安全技术 个人信息安全规范
  - [4] 中华人民共和国个人信息保护法（2021年8月20日第十三届全国人民代表大会常务委员会第三十次会议通过）
  - [5] 中国人民银行令（2021）第4号（征信业务管理办法）
-

全国团体标准信息平台