

# 贷后评分卡模型解析

## 一、现状

评分卡模型我们对于申请评分卡（A卡）、行为评分卡（B卡）了解的比较多，但是对于催收评分卡（C卡）介绍很少，很多公司对C卡也没有那么重视。C卡因为它处于风险管理的末尾-贷后流程中，一方面，是因为贷前风控是整个风控的重点，只有贷前风控做好了，贷中贷后工作就相对轻松。另一方面，是因为很多公司在过了M1就委外处理了。（银行一般是M3）

随着行业发展，以及国家政策的影响，各家公司开始重视自己的贷后催收团队，贷后催收也越来越会向精细化的方法发展。

## 二、贷后评分卡模型分类

催收评分卡和申请评分卡和行为评分卡不太一样，一般申请评分卡和行为评分卡使用一个模型就可以了。催收评分卡通常根据账龄周期以及风险高低分为还款率模型（预测客户是否会迁徙正常还款）、账龄滚动模型（预测客户从一个状态迁徙到另一种状态的概率）、失联预测模型（预测客户是否会迁徙正常还款）等。从多个维度对客户进行分类，进一步提高管理的精细化程度，从而可大大降低运营成本，降低风险。

### 三、催收评分卡模型开发方法

催收评分的构建主要应用数据挖掘和统计学方法，建模方法一般包括如下三种。

#### 1、决策树模型

以树的形式表示目标变量与预测变量的关系的模型，树中每个节点表示预测变量，每个节点的分叉表示预测变量的不同取值，每个树叶代表不同路径对应的预测值，这种算法比较直观且容易理解，并且对数据质量的要求相对较低。

Logistic 回归模型公式为

$$y = \frac{1}{1 + e^{-(\alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}}$$

式中  $y$  是需要预测的目标变量， $X$  为已知变量， $\alpha$  和  $\beta$  是用极大似然法估计得到的模型参数。这时模型计算出的  $Y$  值为  $0 \sim 1$ ，表示事件发生的概率，在应用该模型时需要注意相关性和非线性等问题。

#### 2、神经网络模型

模拟神经细胞的运作模式，建立由输入层、输出层和隐含层的神经元组成的神经网络，经过反复学习来调整权重，使得预测的误差达到预期要求的模型。在应用时应注意根据输入数据的情况来设定学习算法与节点个数。

#### 3、评分结果展现

根据模型计算得到每个客户的分数，分数不同的客户未来表现也不同，据此设定客户类别，达到区分客户的目的。表 1 是模型评分结果展现的一个例子。

表 1 催收评分模型展现

评分结果分段	分数段对应的总体客户 (个)	分数段对应的预测违约客户 (个)	预测违约客户比例 (%)
1 ~ 100	1000	500	50.00
101 ~ 200	1000	250	25.00
201 ~ 300	1000	100	10.00
301 ~ 400	1000	50	5.00
401 ~ 500	1000	10	1.00

#### 四、催收评分模型搭建流程

- 1、逾期客服画像与应用，通过统计建模深度挖掘客户标签，实现逾期客服分群、智能催收策略设计。
- 2、针对不同阶段客户，采用不同模型和策略，采用模型融合方式。贷后评分卡模型的搭建流程的 A 卡 B 卡搭建的流程都是大同小异的。业务理解、数据准备、数据分析、模型设计、样本衍生、模型搭建、策略设计、测试上线好监控优化。

#### 五、催收评分卡模型常见数据变量

- 1、常见变量分类
  - 1) 客户基本信息

性别、年龄、学历、婚姻、注册记录等。

## 2) 银行信息

信用卡额度、信用卡余额、信用卡账单、消费记录等。

## 3) 通讯信息

设备信息、运营商基本信息、账单记录、通话记录、短信记录、联网记录等。

## 4) 借款行为信息

登录记录、授信额度、借款笔数、借款金额、借款时间、借款用途等。

## 5) 客户信息变动

学历变动、婚姻变动、工作变动、一致性验证等。

## 6) 还款行为信息

提前还款类、自动扣款类、逾期还款类、还款预测类、账户还款阶段等。

## 7) 客户偏好

登录时间偏好、产品使用偏好、借款时间偏好等。

## 2、模型常用变量

### 1) 还款率模型（预测客户是否会迁徙正常还款）

逾期天数 (DPD)、历史还款率信息、个人信息、DBR (debt burden ratio)、联系人关系信息等。

### 2) 账龄滚动模型（预测客户从一个状态迁徙到另一种状态的概率）

逾期天数 (DPD)、历史还款率信息、个人信息、DBR (debt burden ratio) 等。

### 3) 失联预测模型（预测客户是否会迁徙正常还款）

逾期天数 (DPD)、逾期金额占比 (outstanding/limit)、个人信息、联系人关系信息、运营商信息等。