

基于 ^{PID} 自动控制的暖气房屋控温模型

摘要

摘要

关键词: 元胞自动机 边缘检测 形状匹配

一、问题重述

1.1 问题背景

为了给予中小企业现金支持，银行针对中小企业的特点，推出一系列不同的信用贷款。对于实力较强信用较好的企业，银行倾向于提供更加优惠的利率。我们需要针对小微企业的开票情况，建立数学模型来给出银行的信贷策略。

1.2 问题重述

经过分析整理，我们需要解决以下问题：

1. 对附件 1 给出有信贷记录的 123 家企业的信贷风险进行量化分析，给出该银行在年度信贷总额固定时对这些企业的信贷策略。
2. 在问题 1 的基础上，对附件 2 中没有信贷记录的 302 家企业的信贷风险进行量化分析，并给出该银行在年度信贷总额为 1 亿元时对这些企业的信贷策略。
3. 综合考虑附件 2 中各企业的信贷风险和可能的突发因素（例如：新冠病毒疫情）对各企业的影响，给出该银行在年度信贷总额为 1 亿元时的信贷调整策略。

二、问题分析

2.1 问题一的分析

为了解决问题 1，需要利用附件中的小微企业发票数据和信用评级来构建数学模型，以给出银行在年信贷额固定时的信贷策略。因此，我们利用企业发票数据估计营业额，资金缺口和稳定情况，且结合信用评级来给出信贷策略。因此，我们需要构建模型以解决贷款给谁、贷款多少以及利率多少的问题。

2.2 问题二的分析

2.3 问题三的分析

三、模型假设

1. 假设企业发票明细完整无误，没有瞒报漏报。

原因：依据企业的发票开具情况，可以直观显示出一个企业的收入支出。因此通过完整的发票记录，可以得到企业的经营特点。

四、名词解释与符号说明

4.1 名词解释

1. **dada**

dsadw

2. dsadc

dasdsas

4.2 符号说明

以下是本文使用的符号以及含义：

符号	说明	单位
L_0	仓库长度	m

五、模型的建立与求解

以下将对提出的三个问题进行建模求解。

5.1 基于熵权法的银行信贷模型

银行信贷的过程主要是解决三个问题：贷给谁，贷多少钱以及贷款利率多少。针对借贷对象而言，需要评估其还款能力来做出贷款选择，我们将在 5.1.1 一节中详细说明一种基于多元线性回归的方法来界定。针对贷款数量而言，我们利用利润率，有效交易占比，供应链丰富度，信誉等级，平均单价，资金缺口六个要素，使用熵权法确定了每个企业的贷款数目，具体内容在 5.1.2 节中体现。针对贷款利率而言，我们在 5.1.3 节中提出一种综合贷款利率和用户流失的多目标规划模型。一批用户经由还款能力评估，贷款数量界定，贷款利率确定后，便可以完成借贷工作。

5.1.1 企业还款能力评估

银行放出贷款后是否能够获得收益，取决于贷款是否能够连本带息如数收回。银行信贷盈利的前提是对企业的还款能力进行有效评估，只有向还款能力较好的企业投放信贷，才会有较低的风险，保证收益来源。

为了衡量企业的还款能力，我们提出以下指标：

1. 利润率

利润率^[2] 在经济学中被解释为总所得和总成本的差额同总成本之间的比值。在题目所给的条件中，结合假设条件，我们认为进项发票的票值代表购买产品的成本，而销项发票的票值代表卖出商品的销售额。利用这一指标，可以判断企业的经营情况，根据统计局的相关数据^[2] 显示，2019 年中小企业营业收入利润率为 5.6%，对于高于这一水平的企业，可视为其有较高的利润水平。我们给出利润率 x_1 的计算公式

$$x_1 = \frac{\sum_{t \in T \text{ 销项}} t - \sum_{t \in T \text{ 进项}} t}{\sum_{t \in T \text{ 进项}} t} \quad (1)$$

其中 T 代表发票集合, T 进项, T 销项 分别代表进项和销项发票, t 代表某一发票的数额, 下同。

2. 有效交易占比

在购买商品时, 如果出现质量问题, 双方无法协商一致时可以选择退货, 这是消费者的权益之一。一家店铺的退货数量较多, 可以从侧面反映出其存在问题, 无论是商品质量, 还是服务是否周全, 都可在这一指标中体现, 所以我们认为开票金额为正的有效发票是有效交易, 计算有效交易的占比以判断经营状况的好坏, 给出下面的计算公式:

$$x_2 = \frac{\text{card}(T_{\text{有效}, t>0})}{\text{card}(T)} \quad (2)$$

其中分子分母中 T 都来自于同一公司的发票记录。

3. 供应链丰富度

一个企业不能脱离于其他的企业而孤立存在, 每个企业都或多或少向其他企业购买产品或者服务, 并向其他企业出售。当企业具有丰富的上下游关系时, 其抗风险的能力较高, 同样反映出企业的组织管理水平较好。供应链丰富度这个指标, 我们定义为与企业发生资金往来的企业数目, 可以使用单位代号进行标识统计。给出供应链丰富度 x_3 的计算方法:

$$x_3 = \text{card}(c_{\text{购方}}) + \text{card}(c_{\text{销方}}) \quad (3)$$

其中 c 代表企业的集合。

4. 信用等级

信用评级^[2] (信用等级) 的目的是显示受评对象信贷违约风险的大小, 一般由某些专门信用评估机构进行。对于已经有信贷记录的企业而言, 银行已经具有信用评级, 可以作为参考, 由于使用 A、B、C、D 四个字母代表不同的信用评级, 为此将其量化为:

$$x_4 = \begin{cases} 1, A \\ 0.75, B \\ 0.5, C \\ \text{不予放贷}, D \end{cases} \quad (4)$$

其中信用等级为 D 的不予放贷, 仅为了完整性罗列于此。

5. 平均单价

平均单价反映了企业流水的规模, 银行更偏向于向流水规模更高的公司提供信贷。在所给条件下, 我们可以使用进销项的平均金额来计算其平均单价。给出计算式:

$$x_5 = \frac{\sum_{t \in T_{\text{销项}}} t + \sum_{t \in T_{\text{进项}}} t}{\text{card}(T_{\text{进项}}) + \text{card}(T_{\text{销项}})} \quad (5)$$

企业的还款能力需要综合上述的五个因素来看，因此需要为五个指标赋予权重。为此，我们借鉴 Chesser 模型^[7]的思想，提出了求解权重的方法。首先基于五个自变量列出多元线性回归^[2]判别法的一般公式：

$$Y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 \quad (6)$$

其中 α 代表常系数， β 代表比例系数， Y 代表最终输出结果，即企业是否还款。各项 x 的数值已在前文说明，总结如下：

$$\begin{cases} x_1 = \frac{\sum_{t \in T_{\text{销项}}} t - \sum_{t \in T_{\text{进项}}} t}{\sum_{t \in T_{\text{进项}}} t} \\ x_2 = \frac{\text{card}(T_{\text{有效}, d > 0})}{\text{card}(T)} \\ x_3 = \text{card}(c_{\text{购方}}) + \text{card}(c_{\text{销方}}) \\ x_4 = \begin{cases} 1, A \\ 0.75, B \\ 0.5, C \\ \text{不予放贷}, D \end{cases} \\ x_5 = \frac{\sum_{t \in T_{\text{销项}}} t + \sum_{t \in T_{\text{进项}}} t}{\text{card}(T_{\text{进项}}) + \text{card}(T_{\text{销项}})} \end{cases}$$

我们输出的结果应当是企业是否还贷款，只有“是”和“否”两个选项。是一个典型的二分类问题，认为 0 为不还贷款，1 为还贷。这样待求出的系数可以使用多元线性回归的工具求出。

在得到多元线性回归方程 (6) 后，为了进行预测还贷情况，需要将 Y 映射到 $[0, 1]$ 区间内，以符合数理统计的规律。所以在多元线性回归的基础上引入 Logit 变换，其步骤如下：

1. 首先引入比例数 (Odds) 的概念，企业可以还款的概率为 P ，则其比例数定义为下式：

$$\text{Odds} = \frac{P}{1 - P} \quad (7)$$

2. 对其取对数，得到 θ 。

$$\theta = \ln \text{Odds} = \ln \frac{P}{1 - P} \quad (8)$$

3. 对式 (8) 变形后，得到概率 P 与 θ 之间的关系：

$$P = \frac{1}{1 - e^{-\theta}} \quad (9)$$

Logit 模型事实上就是将线性回归的输出 Y 视作与 θ 等同，这样最终输出结果在 0-1 之间，符合要求。因此估计一家企业的获得贷款的概率 P 由下式计算：

$$P = \frac{1}{1 - e^{-Y}} \quad (10)$$

$$Y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5$$

综上所述，在针对一名用户判断是否放贷时，首先判断其信用评级是否为 D 若是，则不予放贷，若不是，则利用式 (10) 来求解，得到概率 P 若大于 0.5 则认为可以贷款，若反之，则不予放贷。流程如下图所示：

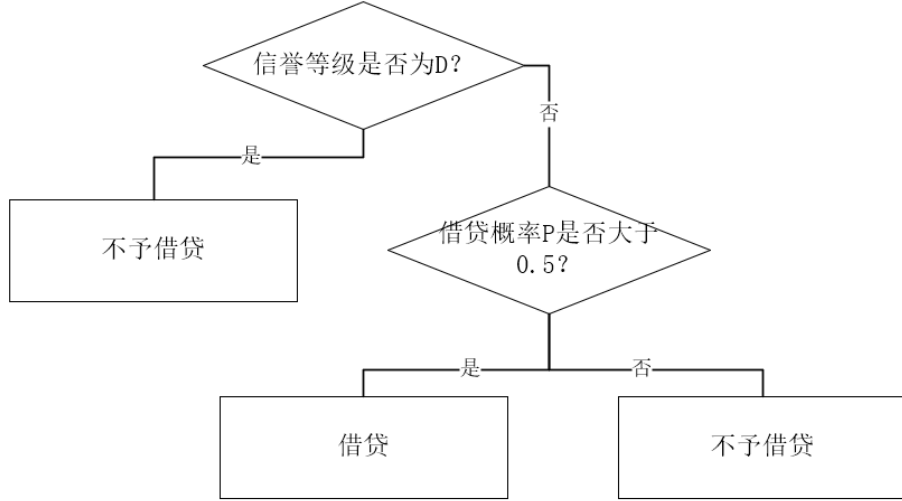


图 1: 银行信贷判断图

5.1.2 贷款数量界定

中小企业由于其经营策略的不同，需要周转的资金数量也有差异。界定中小企业的贷款需求，精准投放有限的资金储备，是银行需要解决的问题。在完成发放贷款与否的界定后，我们将使用熵权法来给出企业贷款数额的界定。

为了更好的判断借款数额，不仅需要利用式 (1-5) 所定义的数值，还需要引入企业的资金缺口 x_6 作为附加的变量。企业在经营的过程中，遵循着先投入后产出的规律，只有首先购进材料产品等，随后售卖盈利，在这一过程中，存在一段时间企业投入资金大于收入所得。在这段时间内产生的支出与获利之差我们定义为资金缺口。当企业获得贷款能够补齐资金缺口时，便可以顺利保证经营的正常进行。我们记录资金缺口为：

$$x_6 = \sum_{t \in T \text{ 缺口进项}} t - \sum_{t \in T \text{ 缺口销项}} t \quad (11)$$

对于许多小企业而言，需要使用同一套标准来计算其贷款额度。计算的依据是 x_1, \dots, x_6 这六个指标，分别代表利润率，有效交易占比，供应链丰富度，信誉等级，平均单价，资金缺口六个要素。这六个要素量纲不同，数值不同，需要利用熵权法来给定各个因素在贷款分配策略中所占有的权重。其具体步骤如下：

1. 利用附件中的数据，计算 i 用户 $(i = 1, 2, 3, \dots, m)$ 的 j 项指标 x_{ij} $(j = 1, 2, 3, \dots, 6)$ 数

值，构成用户指标矩阵 $X_{m \times n}$ ，每项指标的计算方法如下：

$$\begin{cases} x_1 = \frac{\sum_{t \in T \text{ 销项}} t - \sum_{t \in T \text{ 进项}} t}{\sum_{t \in T \text{ 进项}} t} \\ x_2 = \frac{\text{card}(T_{\text{有效}, t > 0})}{\text{card}(T)} \\ x_3 = \text{card}(c_{\text{购方}}) + \text{card}(c_{\text{销方}}) \\ x_4 = \begin{cases} 1, A \\ 0.75, B \\ 0.5, C \\ \text{不予放贷}, D \end{cases} \\ x_5 = \frac{\sum_{t \in T \text{ 销项}} t + \sum_{t \in T \text{ 进项}} t}{\text{card}(T_{\text{进项}}) + \text{card}(T_{\text{销项}})} \\ x_6 = \sum_{t \in T \text{ 缺口进项}} t - \sum_{t \in T \text{ 缺口销项}} t \end{cases}$$

2. 利用 X ，按照熵权法的步骤进行分配权重的计算。

$$p_{ij} = \frac{x_{ij}}{\sum_{i=1}^n x_{ij}} \quad (12)$$

$$e_j = -\frac{1}{\ln n} \sum_{i=1}^n p_{ij} \quad (13)$$

$$g_j = 1 - e_j \quad (14)$$

$$w_j = \frac{g_j}{\sum_{j=1}^m g_j} \quad (15)$$

式 (13) 利用各个指标出现的概率 p 计算熵值，并且在式 (15) 中计算各项指标的权重 w_{ij} 。该方法由数据自身给出结果，较为客观有效。

3. 利用得出的权重 $w_j, j = 1, 2, \dots, 6$ ，可以计算每个企业的贷款评分数值 s_i 。

$$s_i = \sum_{j=1}^m w_j \cdot x_{ij} \quad (16)$$

4. 在计算出每一企业的贷款评分数值 s_i 后，根据数值在当年贷款企业总评分 S 之间的占比来分配银行贷款总额 C_0 。

$$\begin{aligned} S &= \sum_{i=1}^n s_i \\ C_i &= C_0 \cdot \frac{s_i}{S} \end{aligned} \quad (17)$$

至此，我们便计算出在有能力偿还贷款的企业中，各企业分配的贷款金额 C_i 。

5.1.3 贷款利率确定

银行的贷款利率会影响客户的续借情况，过高的贷款利率给企业带来负担，最终导致企业不再继续向该行借款。这种不再续借的现象称为用户流失。银行要做到收益最大化，需要平衡利率和流失率之间的关系。

分析附件三中的数据，并做可视化，得到图（2）：

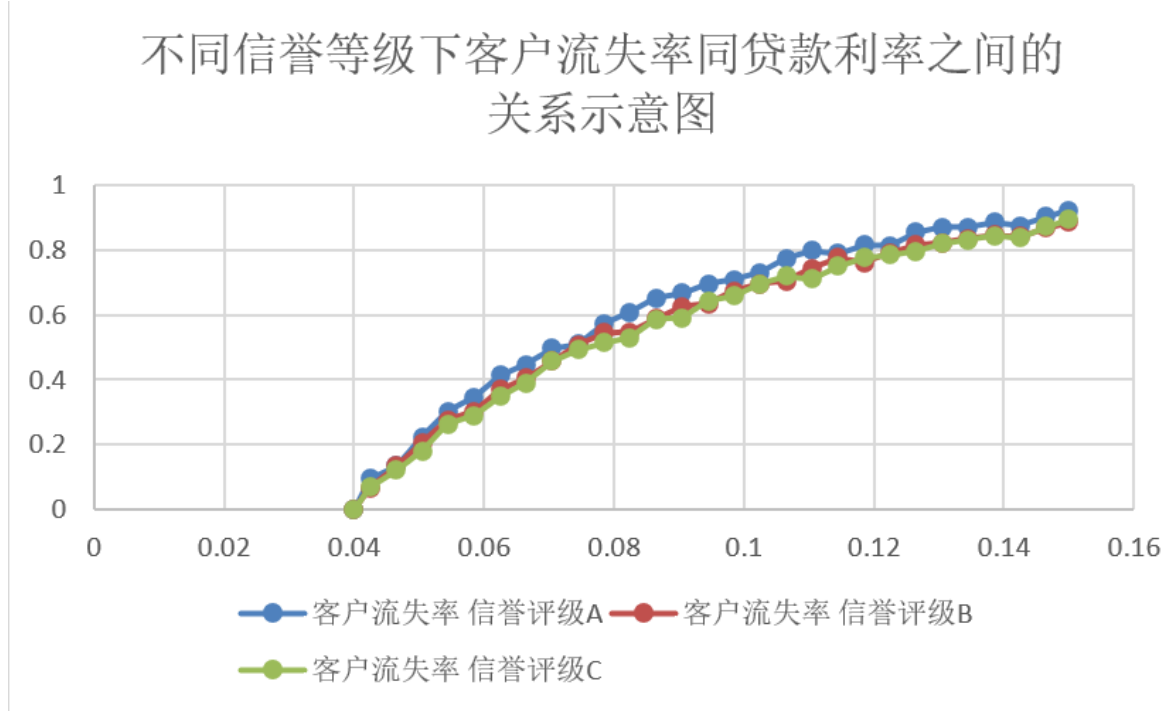


图 2: 不同信誉等级下客户流失率同贷款利率之间的关系示意图

可见当贷款利率上升时，用户流失率也随之上升，基本反映了客观规律。且横向对比不同信誉评级的客户，具有较高信誉评级的客户，面对同一借款利率是较容易流失的。这是由于信誉较好的企业能够在同等条件下获得更佳的利率优惠，银行承担的违约风险较小，因而面对较高的贷款利率时可以做出其他选择。

银行借贷的策略，不仅需要保证当期的贷款能够获得较高利润，还要确保未来还有稳定的客户来源，因此需要保持较低的流失率。为此，我们提出了一个兼顾二者的方法，将这一问题转换为多目标规划的问题。银行考虑取得最佳收益，这要求对每个用户都给出最佳的利率 I_i ，以兼顾收益和发展。

给出两个指标：信贷收益 R 和流失率 L 。

- 信贷收益 R

即银行借款，按照利率 I 收回本息后，可以获得的利润。对于用户 i 而言，若其利率为 I_i ，银行贷款额度（5.1.2 节）为 C_i ，则银行在该用户上的信贷收益 R 计算如下：

$$R_i = C_i \cdot I_i \quad (18)$$

- 流失率 L

流失率定义为次年用户不再续借人数同当年借款人数的比例，其关于贷款利率和信誉等级的关系已经由附件三给出，可以使用拟合的方法给出关系式 $f(I, grade)$ 来求解。

$$L_i = f(I_i, grade) \quad (19)$$

为了综合考虑这两个指标，还需要对得到的数据做进一步处理，在计算出每个用户的信贷收益 R 和流失率 L 后，计算两个指标的最大值最小值，做归一化处理。

$$\begin{aligned} R_i^* &= \frac{R_i - R_{min}}{R_{max} - R_{min}} \\ L_i^* &= \frac{L_i - L_{min}}{L_{max} - L_{min}} \end{aligned} \quad (20)$$

利用归一化后的指标，分别赋权，作为目标函数 F ：

$$\operatorname{argmax}_{I_i} F = \omega_1 \cdot R_i^* + \omega_2 \cdot L_i^* \quad (21)$$

，其中 R_i^* 和 L_i^* 分别由式 (18) 式 (19) 以及式 (20) 得出，如下所示：

$$\begin{aligned} R_i^* &= \frac{C_i \cdot I_i - R_{min}}{R_{max} - R_{min}} \\ L_i^* &= \frac{f(I_i, grade) - L_{min}}{L_{max} - L_{min}} \end{aligned} \quad (22)$$

至此便得出针对每一用户的最佳利率。

5.1.4 模型求解

5.2 未知

六、敏感性分

七、模型的评价

7.1 模型的优点

1. 采用

7.2 模型的缺点

- 利用较

附件

附件清单：

- xxx 代码

sobel 边缘检测代码

function GAdsa