

基于风险量化、数学规划与决策树模型的 中小微企业信贷策略制定

徐京伟¹, 金 霏², 李欣元³

(1.上海大学 计算机工程与科学学院, 上海 200444; 2.上海大学 经济学院, 上海 200444; 3.上海大学
社会学院, 上海 200444)

摘要: 本文针对解决银行对中小微企业信贷策略制定问题, 分别对有信贷记录企业和无信贷记录企业的信贷风险进行量化分析, 以及在突发事件下(如疫情)基于股市数据对不同行业的企业的信贷风险量化分析, 研究出合理的数学规划和决策树模型, 由此制定出使银行利润最大化的信贷策略。并以具体数据为例进行量化分析建立模型, 以此风险指标为基准制定出对应的具体信贷策略。

关键词: 数学规划; 决策树; 中小微企业; 信贷; 银行

中图分类号: F83

文献标志码: A

Credit Strategy Formulation Based on Risk Quantification, Mathematical Programming and Decision Tree Model

XU Jing-wei¹, JIN Fei², LI Xin-yuan³

(1.School of Computer Engineering and Science, Shanghai University, Shanghai 200444, China

2. School of Economics, Shanghai University, Shanghai 200444, China

3. School of Sociology and Political Science, Shanghai University, Shanghai 200444, China)

Abstract: The paper aims at solving the problem of bank for small and medium-sized enterprise credit policy. Respectively to have credit record and no credit record enterprises a quantitative analysis of credit risk, and under the emergency (such as the outbreak) based on stock data of different industry enterprise's credit risk quantitative analysis, study the reasonable mathematical programming and decision tree model, thus making up some credit strategy of bank profit maximization. Taking the specific data as an example, quantitative analysis is carried out to establish the model, and the corresponding specific credit strategy is formulated based on the risk index.

Key words: mathematical programming; decision tree; small and medium enterprises; credit; bank

收稿日期: 2020-11-13

通信作者: 徐京伟 (2000~), 男, 本科生. E-mail: ruojixjw@shu.edu.cn

在我国，中小微企业在经济与就业等领域中占有举足轻重的地位，因此，它对于中国经济的发展有着深刻而长远的意义。然而，中小微企业却常会因融资困难而遇到困境，其发展过程中出现的成本增加、人才匮乏等问题常使企业难以运转^[1]。银行信贷在其中便能发挥极大的作用，适合的银行信贷策略可以使中小微企业融资更为顺畅，同时，也能使银行自身拥有更长久的美好发展^[2]。所以，鉴于对国家经济发展的帮助以及银行自身的发展需要，制定适合的银行信贷策略尤为重要^[3]。

今年年初，突如其来的一场新冠疫情使中小微企业的生存环境恶化，不同行业、不同类型的中小微企业的信贷风险也随之产生变化。为使银行能提前做好应变措施以减少不必要损失，针对突发事件下的信贷策略调整也极其紧要。

为避免银行的过多损失，并尽量保证充分利用信贷资金，本文皆采用保守型借贷策略。将信誉过低、风险过高的企业排除在银行信贷对象之外，并对信誉高、风险低的企业设计出更为合适、优质的信贷策略。此外，本文将企业的基本信息、交易票据和银行客户流失率默认为已知信息。

1 不同信誉评级企业的风险量化模型

1.1 解决思路

由于针对有信贷记录企业的信贷策略需要量化企业信誉对于银行信贷风险的影响，本文根据已知的 A、B、C、D 四类信誉评级和企业的交易票据进行信贷风险估测。为了量化评估有信贷记录企业的风险指标，本文从已有数据中能够分析出风险与企业的业绩与银行给予的信誉评级有关。因此本文建立风险指标函数选取企业年化投入产出比和信誉评级作为自变量。同时信誉评级影响因子能够对 A、B、C 类企业的风险指标作出小幅度调整，并且使 D 类企业的风险指标达到一个很大的值。

1.2 模型假设

所有 A、B、C、D 类企业的毁约、守约率与本文根据具体数据的 A、B、C、D 类企业毁约、守约率相同。所有企业的管理层心理不受影响，即不同信贷利率时 A、B、C 类企业的客户流失率不变。客户流失率导致信贷失败的金额不能在同一批次再次借贷。（注：后同）

1.3 企业信贷风险量化模型的搭建

定义企业变量为 E ，设风险指标为 $R(E)$ ，企业的信誉评级为 $\alpha(E)$ ， $\alpha \in \{A, B, C, D\}$ ，每个企业年化投入产出比为 $i(E)$ ，信贷风险的影响主要源于近三年收益率、信誉、信贷政策及上下游企业的影响力，其中，现行信贷政策大力鼓励银行等金融机构对中小微企业提供信贷支持，而企业的上下游企业若为国家现行政策大力推进发展且影响力大（如 5G 华为基站建设等），则会优先给予其信贷支持，但运行及收益良好的传统行业或暂时并未发展迅速的新兴行业也应给予相应信贷支持，故将以上两个对风险评估仅存在微小影响的因素排除，即将信贷政策和上下游企业的影响力排除^[4]，因此设其影响因子分别为信誉评级影响因子 $Q(\alpha)$ 和年化投入产出比影响因子 $P(i)$ 。

$$R(E) = Q(\alpha) \times P(i) \quad (1)$$

对于有过信贷记录的企业进行统计后能得到下图（注：该数据来自本文所自设的具体数据，数据仅用于模型搭建测试）

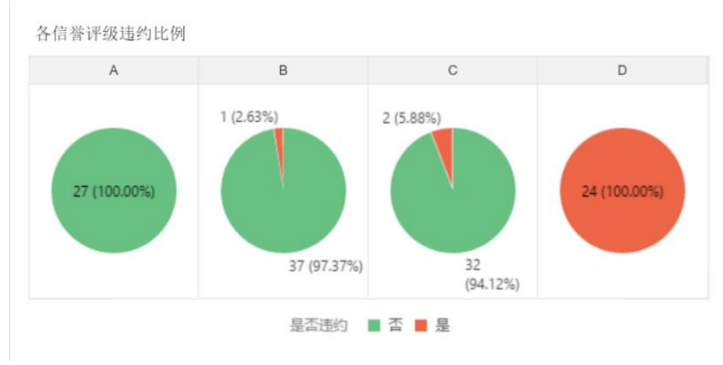


图 1 不同信贷评级公司的违约情况饼图

Fig.1 Pie chart of default of different credit rating companies

设守约率为 $\mu(\alpha)$ ，可得

$$\mu(\alpha) = \begin{cases} 1, \alpha = A \\ 0.9737, \alpha = B \\ 0.9412, \alpha = C \\ 0, \alpha = D \end{cases} \quad (2)$$

（注：该数据来自本文所自设的具体数据，数据仅用于模型搭建测试）

我们希望信誉评级影响因子 $Q(\alpha)$ 能够将 D 级企业的风险指标调整到极大，并且给予 A、B、C 类公司不同的风险指标调整，因此构造信誉评级影响因子为

$$Q(\alpha) = \frac{1}{e^{\mu(\alpha)} - 1 + 10^{-50}} \quad (3)$$

将（2）式代入（3）式后得到各个信誉评级的信誉评级影响因子数值

$$Q(\alpha) = \begin{cases} 0.582, \alpha = A \\ 0.607, \alpha = B \\ 0.640, \alpha = C \\ 10^{50}, \alpha = D \end{cases} \quad (4)$$

同时，我们希望年化投入产出比影响因子 $P(i)$ 能够将年化投入产出比高的企业的风险指标调整到较低数值，并且在 i 超过一定数值时减缓 $P(i)$ 的降低；同时能够将年化投入产出比低的企业风险指标调整到较高数值，并且在 $i < 0$ 时快速上升。因此构造年化投入产出比影响因子为

$$P(i) = e^{-i(E)} \quad (5)$$

将（3）、（5）式代入（1）式后得到

$$R(E) = \frac{1}{e^{\mu(\alpha)} - 1 + 10^{-50}} \cdot e^{-i(E)} \quad (6)$$

至此为企业信贷风险的量化模型搭建。

2 对有信贷记录企业的信贷策略模型

2.1 解决思路

通过已有数据之间的数量关系，本文通过分析利润函数内的局部表达式，来确定能使银行利润最大的利率值。以此将利润函数简化为一个只含三个自变量的函数，并结合所能提取的约束条件以及利用量化模型对各个企业的风险评估结果，求解能使银行利润最大化的策略。

2.2 数学规划模型的搭建与应用

由于银行对信誉评级为 D 的企业原则上不予放贷，故不作分析。

设银行的年度信贷总额为 w 万元，给 A 类企业信贷的总金额为 x 万元，给 B 类企业信贷的总金额为 y 万元，给 C 类企业信贷的总金额为 z 万元。可得

$$x + y + z = w \quad (7)$$

设 A 类企业共有 a 家，B 类企业共有 b 家，C 类企业共有 c 家，由于银行对确定要放贷企业的贷款额度为 100 万元，因此有约束条件

$$\begin{cases} x \in [0, 100a] \\ y \in [0, 100b] \\ z \in [0, 100c] \end{cases} \quad (8)$$

记银行信贷给 A 类企业的统一利率为 k_A ，信贷给 B 类企业的统一利率为 k_B ，信贷给 C 类企业的统一利率为 k_C ，同时记 A 类企业对应信贷利率的客户流失率为 $m_A(k_A)$ （注：此处客户流失率为本文自设的具体数据），B 类企业对应信贷利率的客户流失率为 $m_B(k_B)$ ，C 类企业对应信贷利率的客户流失率为 $m_C(k_C)$ 。对 A、B、C 类企业的客户流失率与统一利率进行散点的曲线拟合，可得

$$\begin{cases} m_A(k_A) = 0.6690 \ln k_A + 2.2386 \\ m_B(k_B) = 0.6506 \ln k_B + 2.1576 \\ m_C(k_C) = 0.6586 \ln k_C + 2.1680 \end{cases} \quad (9)$$

此外同时有

$$\begin{cases} k_A \in [0.04, 0.15] \\ k_B \in [0.04, 0.15] \\ k_C \in [0.04, 0.15] \end{cases} \quad (10)$$

其拟合曲线与散点图如下

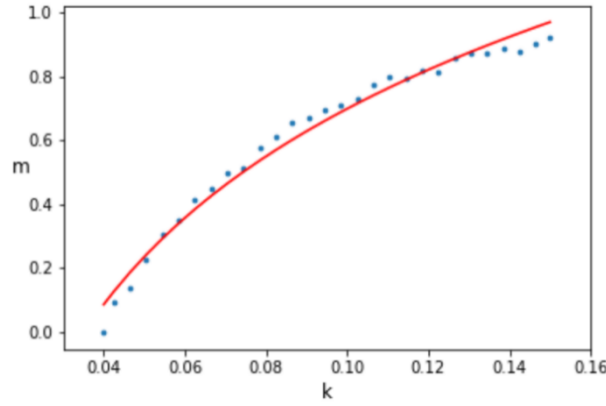


图 2 A 类企业客户流失率与信贷利率关系图

Fig.2 Relationship between customer churn rate and credit interest rate of Class A Enterprises

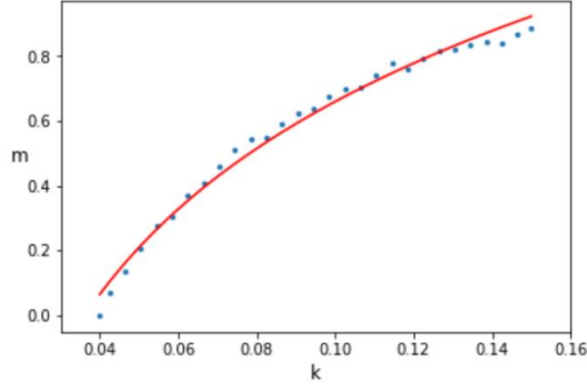


图 3 B 类企业客户流失率与信贷利率关系图

Fig.3 Relationship between customer churn rate and credit interest rate of Class B Enterprises

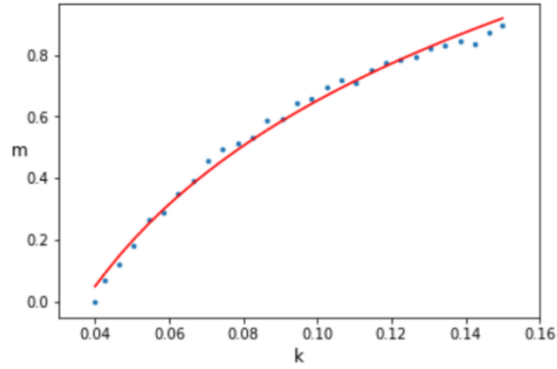


图 4 C 类企业客户流失率与信贷利率关系图

Fig.4 Relationship between customer churn rate and credit interest rate of Class C Enterprises

由于以上三幅图中当利率为 0.04 时与拟合曲线的误差较大，因此单独讨论。

①当利率不等于 0.04 时

设银行在本次信贷的总盈利额为 σ ，可得

$$\sigma = x \cdot k_A \cdot (1 - m_A) \cdot \mu(A) + y \cdot k_B \cdot (1 - m_B) \cdot \mu(B) + z \cdot k_C \cdot (1 - m_C) \cdot \mu(C) \quad (11)$$

以 A 类企业为例，定义

$$\gamma_A(k_A) = k_A \cdot [1 - m_A(k_A)] \quad (12)$$

为 A 类企业的利率影响因子。

将 (9) 式代入 (12) 式中可得，

$$\gamma_A(k_A) = k_A \cdot [1 - (0.6690 \ln k_A + 2.2386)] \quad (13)$$

对 (13) 式求关于 k_A 的导数，可得

$$\frac{d\gamma_A}{dk_A} = -(1.9076 + 0.6690 \ln k_A) \quad (14)$$

可得当

$$k_A = 0.0578 \quad (15)$$

$\frac{d\gamma_A}{dk_A} = 0$ ，且当 $k_A \in [0.04, 0.0578]$ 时， $\frac{d\gamma_A}{dk_A} > 0$ ；当 $k_A \in [0.0578, 0.15]$ 时， $\frac{d\gamma_A}{dk_A} < 0$

故对于 A 类企业，采取 $k_A = 0.0578$ 可得到使得银行收益最大的利率影响因子 γ_A 。

同理，对于 B、C 类企业可以采取同样的处理方式，最终得到当

$$\begin{cases} k_B = 0.0621 \\ k_C = 0.0625 \end{cases} \quad (16)$$

时 B、C 类企业的利率影响因子 γ_B 和 γ_C 可取得最大值。

$$\begin{cases} \gamma_A = 0.0386 \\ \gamma_B = 0.0404 \\ \gamma_C = 0.0411 \end{cases} \quad (17)$$

将 (2)、(17) 式代入 (11) 式中可得

$$\sigma = 0.0386x + 0.0393y + 0.0387z \quad (18)$$

②当利率为 0.04 时

可得

$$\sigma = x \cdot k_A \cdot \mu(A) + y \cdot k_B \cdot \mu(B) + z \cdot k_C \cdot \mu(C) \quad (19)$$

将 (2) 代入 (19) 可得

$$\sigma = 0.0400x + 0.0389y + 0.0376z \quad (20)$$

综上可知对于 A 类企业，最好的信贷策略是以 4.00% 的利率贷款。而对于 B、C 类企业，最好的信贷策略分别是以 6.21% 和 6.25% 的利率贷款。

结合 (7)、(8)、(18)、(20) 式可得

$$\begin{cases} x + y + z = w \\ \sigma = 0.0400x + 0.0393y + 0.0387z \\ x \in [0, 100a] \\ y \in [0, 100b] \\ z \in [0, 100c] \end{cases} \quad (21)$$

(注：为使各企业都能得到信贷，故设每个企业最多投 100 万元，数据仅用于模型搭建测试)

为了评估信贷各个企业的风险是否过高，需要定义风险评估的基准值。由于当一个企业的 $\alpha = A$ 且 $i = 0.04$ 时，该企业必定能够偿还利率为 4% 时的债务，因此选此值作为基准值 R_0 。

代入 (6) 式可得

$$R_0 = 0.559 \quad (22)$$

经过统计可得符合条件的 A、B、C 类企业

$$\begin{cases} a = 22 \\ b = 28 \\ c = 30 \end{cases} \quad (23)$$

(注：此处所计企业数为本文自设的具体数据，数据仅用于模型搭建测试)

将 (23) 式代入 (21) 式可得

$$\begin{cases} x + y + z = w \\ \sigma = 0.0400x + 0.0393y + 0.0387z \\ x \in [0, 2200] \\ y \in [0, 2800] \\ z \in [0, 3000] \end{cases} \quad (24)$$

当 $w \in [0, 2200]$ 时，银行优先给信誉评级为 A 并且按照风险指标由低至高给企业放贷；当 $w \in [2200, 5000]$ 时，银行在给所有 A 类企业放贷 100 万元后，再按照风险指数由低至高给 B 类企业放贷；当 $w \in [5000, 8000]$ 时，银行在给所有 A、B 类企业放贷 100 万元后，再按照风险指数由低至高给 C 类企业放贷。

3 对无信贷记录企业的信贷策略模型

3.1 解决思路

为了对无信贷记录的企业做出信贷策略, 本文将数据中统计得到的各个有信贷记录企业的进项价税合计和销项金额合计作为特征属性值, 将数据中银行给出的信誉评级作为分类值, 训练决策树分类模型。随后对数据中的所有无信贷记录企业进行信誉评级预测, 并根据预测结果, 使用量化模型风险评估和数学规划模型进行投资策略制定。

3.2 决策树模型的训练与应用

令年度信贷总额为 1 亿元, 有

$$w = 10000 \quad (25)$$

以 123 家企业 (注: 该多家企业的数据来自本文自设的具体数据, 数据仅用于模型搭建测试) 统计所得的总进项价税合计以及总销项金额作为特征属性值, 同时以该 123 家企业的信誉评级作为分类值, 来训练决策树分类器。并以训练集的数据作为测试对象, 输出该分类器的性能报告和混淆矩阵。

表 1 决策树性能报告和混淆矩阵

Tab.1 Decision tree performance report and confusion matrix

	precision	recall	f1-score	support
0.0	1.00	1.00	1.00	24
1.0	1.00	1.00	1.00	34
2.0	1.00	1.00	1.00	38
3.0	1.00	1.00	1.00	27
avg / total	1.00	1.00	1.00	123
confusion matrix:				
[[24 0 0 0]				
[0 34 0 0]				
[0 0 38 0]				
[0 0 0 27]]				

由此可知该分类器的稳定性较好。

随后将 302 家企业 (注: 该多家企业的数据来自本文自设的具体数据, 与之前的 123 家企业数据完全不相同, 数据仅用于模型搭建测试) 的总进项价税合计以及总销项金额作为输入的特征值, 并输出各企业信誉评级的预估值。

随后根据风险指标量化模型可以算出适合信贷的 A、B、C 类企业, 最终由统计可得

$$\begin{cases} a = 41 \\ b = 121 \\ c = 48 \end{cases} \quad (26)$$

将 (25)、(26) 式代入 (21) 式可得

$$\begin{cases} x + y + z = 10000 \\ \sigma = 0.0400x + 0.0393y + 0.0387z \\ x \in [0, 4100] \\ y \in [0, 12100] \\ z \in [0, 4800] \end{cases} \quad (27)$$

对其求解可得当

$$\begin{cases} x = 4100 \\ y = 5900 \\ z = 0 \end{cases} \quad (28)$$

时, σ 可取到最大值 395.87 万元。

具体每种信誉评级的各个企业信贷分配情况的部分结果如下

表 2 企业信贷金额表 (部分)

Tab.2 Part of enterprise credit amount table

企业代号	风险指标	是否低于风险基准值	预期信誉评级	分配金额
E124	0.905351396	0	C	0
E125	0.966695223	0	C	0
E126	0.124062543	1	A	100
E127	0.000430476	1	A	100
E128	0.034293569	1	A	100
E129	0.156171003	1	B	100
E130	0.303313072	1	B	0
E131	0.269635267	1	C	0

4 考虑突发因素的企业信贷策略模型修正

4.1 解决思路

为了探索在突发情况下银行针对不同行业的企业所调整的信贷策略, 本文以 2020 年的新冠疫情为例, 根据股市各个板块盘从 2020 年 1 月 20 日到 3 月 23 日的盈亏情况, 将数据中的不同企业分为利好、利空、无影响企业 (注: 股市数据为真实数据), 并分别选取盈利板块盈利率和亏损板块亏损率的中位数来处理利好企业和利空企业的销项总金额, 并使用处理好之后的数据重新评估数据中每个企业的风险指标, 通过数学规划模型重新决定哪些企业是值得银行的信贷。

4.2 企业信贷量化风险模型的修正

以新冠疫情为例。在新冠疫情初期, 依据股市的数据直观表明, 受疫情影响, 除去生物、医疗、通信等行业有小幅上涨以外, 如机场航运、采掘服务等行业的企业受到重创, 尤其对大部分中小微企业来说, 因为规模小, 因此疫情甚至对其有致命影响。但因未来存在发展的不确定因素, 故银行需以当时的行情为标准进行放贷以减少损失, 疫情初期的涨跌幅数据越高, 则越优先放贷, 且按当时企业的需求进行放贷。并且, 根据疫情中后期的股市涨跌幅数据, 随着疫情常态化的影响, 可以发现初期股市上涨的行业在中后期几乎依旧涨幅巨大且稳定, 故对其行业的中小微企业进行优先放贷是可行且可靠的。

为了评估新冠疫情对于不同产业的影响, 我们根据 2020 年 1 月 20 号到 3 月 23 号不同板块的股市涨跌幅, 对无信贷记录的 302 家企业进行了分类, 依次将各企业分为利好产业、利空产业和无影响产业。我们将涨幅>3%的板块内企业归为利好产业, 将跌幅>3%的板块内企业归为利空产业, 将涨跌幅位于[-3%, 3%]区间的板块内企业归为无影响企业。

记利好产业涨幅中位数为 φ , 利空产业跌幅中位数为 ϕ , 对进行股票市场的统计可以得出

$$\begin{cases} \varphi = 4.19\% \\ \phi = -11.96\% \end{cases} \quad (29)$$

记利好产业的调整后年化投入产出比为 $i_1(E)$, 利空产业的调整后年化投入产出比为 $i_2(E)$, 疫情发生后各企业的风险指标为 $R'(E)$ 。

令

$$\begin{cases} i_1(E) = i(E) + \varphi \\ i_2(E) = i(E) + \phi \end{cases} \quad (30)$$

作为利好利空产业在新冠疫情发生后的年化投入产出比,经统计后可得到新冠疫情发生后无信贷记录的 302 家企业的风险指标 $R'(E)$

表 3 调整后的企业风险指数表 (部分)

Tab.3 Part of Enterprise risk index table after adjustment

企业代号	调整后年化投入 产出比	预期信誉评级	调整后风险指标	是否低于风险基 准值
E124	-0.342384789	C	0.900993168	0
E125	-0.407944906	C	0.962041695	0
E126	1.545644605	A	0.124062543	1
E127	7.209294967	A	0.000430476	1
E128	2.831472566	A	0.034293569	1
E129	1.359523319	B	0.15584122	1
E130	0.695709386	B	0.302672572	1
E131	0.868868177	C	0.268337283	1

统计后可得到风险值低于 $R'(E)$ 风险基准值的各类企业数量

$$\begin{cases} a = 41 \\ b = 120 \\ c = 48 \end{cases} \quad (31)$$

(注:该多家企业的数据来自本文所自设的具体数据,为前文的 302 家企业)
将 (25)、(31) 代入 (21) 式中可得

$$\begin{cases} x + y + z = 10000 \\ \sigma = 0.0400x + 0.0393y + 0.0387z \\ x \in [0,4100] \\ y \in [0,12100] \\ z \in [0,4800] \end{cases} \quad (30)$$

经计算可得到最优方案为 A 类企业共信贷 4100 万元, B 类企业共信贷 5900 万元,由于对于同一类企业采取的信贷策略是由评估的风险指标由低到高依次进行金额发放,因此针对 B 类企业中的信贷策略相较于疫情发生前的信贷策略有发生变化。

表 4 调整后的企业信贷金额表 (部分)

Tab.4 Part of Enterprise credit amount table after adjustment

企业代 号	调整后年化投入 产出比	预期信誉评级	调整后风险 指标	是否低于风 险基准值	信贷金额
E124	-0.342384789	C	0.900993168	0	0
E125	-0.407944906	C	0.962041695	0	0
E126	1.545644605	A	0.124062543	1	100
E127	7.209294967	A	0.000430476	1	100
E128	2.831472566	A	0.034293569	1	100
E129	1.359523319	B	0.15584122	1	100
E130	0.695709386	B	0.302672572	1	100
E131	0.868868177	C	0.268337283	1	0

5 模型的稳定性测试

将 B、C 类企业的违约率分别上升和下降 5%，并重新计算 B、C 类企业的风险指标，并与基准值比较来进行风险评估，能够得到如下处于风险基准值以下的 A、B、C 类企业数量

表 5 模型敏感度测试结果表

Tab.5 Model sensitivity test results			
违约率变化率	a	b	c
0%	41	121	48
5%	41	121	48
-5%	41	121	48

表 6 毁约率加 5%时的企业风险指标表（部分）

Tab.6 Part of Enterprise risk index table when breaking contract rate plus 5%

企业代号	风险指标	预期信誉评级	是否低于风险基准值
E124	0.905351396	C	
E125	0.966695223	C	0
E126	0.124062543	A	0
E127	0.000430476	A	1
E128	0.034293569	A	1
E129	0.156171003	B	1
E130	0.303313072	B	1
E131	0.269635267	C	1

表 7 毁约率减 5%时的企业风险指标表（部分）

Tab.7 Part of Enterprise risk index table when breaking contract rate minus 5%

企业代号	风险指标	预期信誉评级	是否低于风险基准值
E124	0.896664051	C	0
E125	0.957419251	C	0
E126	0.124062543	A	1
E127	0.000430476	A	1
E128	0.034293569	A	1
E129	0.155512395	B	1
E130	0.302033932	B	1
E131	0.267047969	C	1

由于该模型在敏感度测试中，对于各企业的预期信誉评级没有发生变化，因此可知该模型的稳定性较好。

6 结论

6.1 模型优点

可以量化比较同一信誉评级不同企业的信贷风险情况；可以制定出针对同一信贷评级所有企业的信贷具体金额；可以结合针对同一信贷评级所有企业的信贷具体金额和同一信誉评级内的信贷风险排名，将信贷具体金额细化到每一个企业。

6.2 模型缺点

不同信誉评级企业之间信贷风险的量化比较不可靠;无法评估决策树模型对无信贷记录企业进行信誉评级时的误差。

致谢 在此感谢经济学院周思老师对本文的悉心指导,感谢学校支持学生跨学院跨专业进行研究。

参考文献:

- [1] 张效英. 我国商业银行中小企业信贷策略研究[D]. 山东大学, 2012.
- [2] 李金峰. 商业银行中小企业信贷策略研究[D]. 山东大学, 2011.
- [3] 何金兰. 基于中小企业融资业务的 ZS 银行信贷策略研究[D]. 兰州大学, 2013.
- [4] 夏志琼, 沈国金. 国外中小企业信贷扶持政策及启示[J]. 中国房地产金融, 2012, 000(002):64-65.