

武汉理工大学

数学建模暑期培训论文

第 3 题

2020 年 C 题真题解析

中小微企业最优信贷策略模型的分析与改进

第 020 组

姓 名	方 向
陶梓	编程、写作
李佳荟	建模、写作

2024 年 2 月 25 日

摘要

本文对于 2020 年 C 题进行真题解析,通过解析 C109、C002、C227 三篇优秀论文,为优化模型算法的简洁度,对问题的求解进行分析与改进。问题一在 C002 的基础上进行改进,使用 Sigmoid 函数建立信贷风险的量化模型,利用多目标非线性规划模型建立信贷策略优化模型;问题二在 C109、C002 的基础上改进使用含 Dropout XGBoost 算法建立信誉评级确定模型,并基于问题一的信贷策略优化模型确定 302 家企业的最佳信贷策略;问题三在 C109、C002、C227 的基础上,从宏观和微观上分析疫情对各类企业的影响,根据影响程度建立贷款调整方案,在问题二模型的基础上建立考虑信贷风险和突发因素影响的信贷策略优化模型。

针对问题一,建立了基于企业信誉和实力两个综合指标的信贷风险评估模型,建立了多目标非线性模型,通过人工鱼群算法求解得到最佳信贷策略。首先建立**信贷风险评估模型**,基于附件一数据特征计算出八项指标,通过动态加权得到企业实力和信誉的综合指标,利用 Sigmoid 函数计算得到信贷风险指标。然后建立**信贷策略模型**,本文以银行收益最大化、风险最小化为目标,考虑信贷风险约束、贷款额度约束、年利率约束、贷款比例约束,建立**多目标非线性规划模型**,利用人工鱼群算法求解得到银行的最佳信贷策略。计算得到最优**银行总利息收益为 194.43 万元,单位贷款额度风险值为 0.0302。**

针对问题二,建立基于含 Dropout 的 XGBoost 模型确定企业风险评级和违约记录,基于问题一现有模型求解得到最佳信贷策略。首先对附件二进行与附件一相同的处理,增加并计算发票作废率和退货率两个指标,并使用 SMOTE 合成少数类过采样技术对数据进行过采样,然后以附件一数据为训练集,利用含 Dropout 的 XGBoost 确定企业风险评级并评判违约记录。然后应用问题一所建立的信贷风险评估模型和信贷策略优化模型,得到针对各类型的最佳信贷策略。结果表明**银行总利息收益为 412.67 万元。**

针对问题三,以新冠疫情为例,从**宏观和微观层面**分析不同类型企业受疫情影响的经济变化,建立了政策补贴调整下信贷策略模型。在宏观层面,依据 2020 年各行业经济变化数据,以固定资产投资作为分析企业经济实力变化的指标,分析疫情对不同行业发展的正负面影响。在微观层面,本文从附件二随机抽取部分企业,分析 2020 年 1 月和 2 月与 2018 年、2019 年的毛利润的同比增长率,以反映疫情对企业经济上的影响。综合考虑宏微观层面,将企业按影响程度划分为四类,给予不同政策补贴。调整信贷策略优化模型,加入将**增大贷款额度和减小利率约束**。求解得到**银行贷款总额度为 1 亿元,总利息收益为 384.58 万元**,相较于问题二减少了 6.90%,但带来了巨大社会效益。

关键词: 中小微企业信贷决策 XGBoost 算法 人工鱼群算法 SMOTE

目录

1 真题解析	1
1.1 真题重述	1
1.1.1 问题背景	1
1.1.2 问题概述	1
1.2 优秀论文赏析	1
1.2.1 C109 论文赏析	1
1.2.2 C002 论文赏析	3
1.2.3 C227 论文赏析	4
2 模型假设	6
3 符号说明	6
4 改进问题一模型的建立与求解	6
4.1 问题一分析	6
4.2 信贷风险评估模型的建立与求解	7
4.2.1 数据分析处理	7
4.2.2 数据特征提取	8
4.2.3 风险评估指标	9
4.2.4 信贷风险评估	10
4.2.5 企业规模划分	10
4.2.6 结果分析	11
4.3 信贷策略优化模型的建立与求解	13
4.3.1 客户流失率与年利率关系	13
4.3.2 多目标非线性规划模型	14
4.3.3 人工鱼群算法求解	16
4.3.4 结果分析	17
5 改进的问题二模型的建立与求解	18
5.1 问题二分析	18
5.2 预备工作	18
5.2.1 信誉等级预测指标选取	18
5.2.2 数据过采样	19

5.3 基于 Dropout XGBoost 算法的信誉评级确定模型.....	20
5.3.1 模型建立	20
5.3.2 模型求解	20
5.4 信贷风险评估模型.....	21
5.5 信贷风策略优化模型.....	21
5.6 结果分析.....	22
6 改进的问题三模型的建立与求解	24
6.1 问题三分析.....	24
6.2 突发因素对企业的影响分析.....	24
6.2.1 宏观层面	24
6.2.2 微观层面	26
6.2.3 银行受疫情影响分析及贷款调整政策建立	27
6.3 调整后的信贷策略规划模型.....	27
6.4 模型求解与分析.....	28
7 模型评价	29
7.1 对优秀论文做出的改进.....	29
7.2 模型推广.....	29
附录 A 问题一 123 家企业风险指标量化表.....	31
附录 B 问题一 96 家贷款比例和贷款利率信息表.....	35
附录 C 问题二 302 家企业信誉评级预测表.....	38
附录 D 问题三 69 家企业 2020 年 1 月 2 月与前两年的同比增长率表.....	47
附录 E 问题一源代码	49
附录 F 问题二 XGBoost 算法源代码	51
附录 G 问题二人工鱼群算法源代码.....	52
附录 H 问题三同比利润率计算代码.....	53

1 真题解析

1.1 真题重述

1.1.1 问题背景

在实际生活中，由于中小微企业规模相对较小，无较多抵押资产，相比于大型企业银行的信贷策略和传统的信贷策略会有所不同。银行通常根据信贷政策、企业交易票据信息以及上下游企业的影响力，向实力雄厚、供需关系稳定的企业提供贷款，对信誉好且信用风险低的企业可以给予优惠利率。一般情况下银行会首先根据中小企业的实力和信誉对其信用风险进行评估，然后根据信用风险等因素确定是否选行放贷和贷款额度、利率、期限等信贷策略。中小微企业的信贷策略问题是一个风险决策问题，具有很强的现实意义，尤其是新冠疫情之后更为重要。

1.1.2 问题概述

某银行对确定要放贷企业的贷款额度为 10 – 100 万元；年利率为 4% 15%；贷款期限为 1 年。该银行请你们团队根据 2019 年统计的有无信贷记录相关企业数据和贷款利率与客户流失率关系，通过建立数学模型，解决下列问题：

(1) 请对附件一有信贷记录的公司的信贷风险进行量化分析，给出信贷总额固定情况下的信贷策略；

(2) 试对附件二无信贷记录的公司的信贷风险进行量化分析，给出信贷总额为 1 亿元时的信贷策略；

(3) 综合考虑附件二无信贷公司的信贷风险与突发因素对企业的影响，给出信贷总额为 1 亿元时的信贷策略。

1.2 优秀论文赏析

本文对 2020 年全国大学生数学建模竞赛 C 题中小微企业的信贷决策进行较全面的真题解析，主要针对高教杯 C109、C002、C227 三篇优秀论文进行解析，并对其优缺点进行分析。

1.2.1 C109 论文赏析

论文解读

该篇论文针对中小微企业的信贷决策问题，采用基于改进的梯度下降算法对决策树模型进行优化，建立信贷风险评估判断预测模型，之后基于收益最大、风险最小、客户流失率的最小目标建立非线性规划模型并进行求解，最后给出最优的银行贷款策略。

针对问题一，首先基于附件的交易票据数据挖掘出各企业的多项经营与财务指标，并进行筛选；之后对传统的决策树模型进行改进，对于原模型添加正则项函数以抑制决策树的复杂性，利用集成学习思想将多个决策树模型进行叠加，基于梯度下降算法进行迭代优化，得到最终的集成模型。基于该监督学习模型，建立以违约概率为目标函数的信贷风险评估模型，对附件一中各企业的违约概率进行定量求解，即得到各企业的信贷风险。求解结果表明，该模型的预测准确率达到 96.75%，相较于传统的决策树算法和 BP 神经网络模型，预测效果分别提升了 15.36% 和 63.15%。基于得到的各企业的信贷违约概率，考虑到银行贷款期限为一年，以银行一年后的贷款收益最大化、客户流失率最小、客户违约风险最小为目标，考虑银行贷款额度、银行贷款利率、银行客户流失率、银行是否放贷、企业违约概率等约束，建立银行信贷策略的非线性多目标规划模型，最终通过求解模型以确定银行的最佳信贷策略。

针对问题二，首先基于前文的改进的梯度下降决策树算法建立企业信用评级模型，并基于附件一的数据进行训练，利用得到的模型对附件二中的 302 家企业进行信用评级。之后应用问题一所建立的信贷风险评估模型和信贷策略模型，对 302 家企业进行信贷风险评估，并确定银行针对 302 家企业具体的最佳信贷策略，最终结果表明银行的最优的期望收益为 199.635 万元，对于各企业贷款利率从 4% – 15% 不等。

针对问题三，首先将 302 家企业划分到 14 个行业大类、4 个规模大类，之后以新冠肺炎疫情作为突发性因素的代表，从宏观和微观层面分析了不同行业、不同规模企业所受到的疫情影响的严重程度。从宏观层面，该篇论文选择各行业具有代表性的上市公司股票，利用 *Python* 爬虫爬取了各行业从新冠疫情爆发以来至今的股票价格信息，以股票价格的波动作为分析该行业的发展与恢复情况的基本指标；在微观层面，从 302 家企业中随机抽取 72 家企业分析其 2020 年 1 月、2 月与 2018 年、2019 年同期的同比增长率，以此判定企业的运营状况。结合宏观与微观分析，将各类企业按照受疫情影响程度分为四类，同时考虑到国家政策对于中小微企业的信贷补贴优惠，从而确定银行的针对不同行业、不同类别企业的贷款调整政策。之后在问题二的基础上建立同时考虑信贷风险和突发因素影响的信贷决策规划模型，并且进行求解。最终银行贷款额度为 1 亿元，一年后的期望收益为 184.655 万元，相较于调整前减少 7.5%，但是由此银行产生巨大的社会效益，可以帮助众多中小企业渡过难关。

论文结构

该篇论文结构如下：

优缺点分析

• 该篇论文优点：

在对问题一的求解中，该篇论文针对中小微企业的信贷问题，基于改进的梯度下降算法对决策树模型进行了优化，建立了企业信贷风险评估判断预测模型，之后基于收益最大、风险最小、客户流失率的最小等目标建立了非线性模型并进行求解，考虑的目标

较为全面，得到的信贷策略也较为合理

在问题二的求解中，该篇论文依据实际，综合考虑银行放贷总额和总收益。企业贷款利率、额度、违约概率和利率与流失率的关系等，建立了以银行总期望收益最大化的优化决策模型，求解，给出了合理的信贷策略。

在问题三的求解中，该篇论文从宏观和微观两个层面上研究了突发因素的正负两方面的影响，给出了信贷决策的调整模型和策略，并对国家相关政策的作用做了讨论，想法新颖，考虑的十分全面。

• 该篇论文缺点：

在论文中多次采用主观性较强的查阅文献估值的方式确定参数如估计客户流失价值，这样做很有可能导致结果的可靠性和准确性较差。

1.2.2 C002 论文赏析

论文解读

该篇论文研究中小微企业的信贷风险与银行的放贷策略，建立了量化风险的灰色评价模型，采用 XGBoost 预测算法计算缺失指标，最后基于单目标优化模型分类别制定贷款策略，并进行了评价。

针对问题一，建立了基于灰色关联度的多层信贷风险评价模型，通过人工鱼群优化算法求解银行最佳信贷决策方案。建立基于月均进项额、月均销项额等 6 个指标的下层企业实力灰色评价模型，再基于信贷评级、企业实力、违约值 3 个指标建立上层企业可靠度灰色评价模型，并通过熵权法求得指标权重。为了简化信贷策略，对企业实力进行 K-means 均值聚类分析，将其分为中、中小、小、微四个类别，并结合信誉评级将最终的企业可靠度分成 A 中企业、A 中小企业等 16 类，并求出各的可靠度得分。最后以一般企业风险率为基准，将各类中的可靠度得分量化为风险率，并基于风险率和客户流失率建立单目标优化模型，通过人工鱼群算法对每一类的放款额度和利率进行寻优求解，据此给出每一类的最佳信贷策略，最后求得银行利润期望为 198.9799 万元。

针对问题二，通过 XGBoost 算法预测出附件二中企业的信誉评级，并结合问题一的信贷风险评价模型给出其量化信贷风险以及银行的最佳信贷策略。选取销项利润率、销项发票数等 13 个自变量，信誉评级为因变量，经过 SMOTE 过采样后，通过 xgboost 算法预测出 302 家企业的信誉评级，然后通过问题一的企业实力评价模型求出附件二中各企业的企业实力。选取信誉评级和企业实力为自变量，违约值为因变量利用附件一数据进行 xgboost 的算法训练，进而预测出附件二中各企业的违约值，并纳入到信贷风险评价模型中，最后通过人工鱼群算法求解出各类的放款额度和利率，并作出信贷决策，最后求得银行利润期望为 305.2666 万元。

针对问题三，在问题一的基础上新增新冠疫情下各行业经济的短期增长率、中长期季度增长率以及扶持力度三项指标重新建立多层信贷风险评价模型。首先将附件二中各

行业企业分为零售业、科技业、传统业、服务业和医疗业五类。然后选取二次函数、指数函数、样条函数平滑曲线三种函数对各行业经济变化进行拟合预测，求得疫情下经济的短期增长率和未来中长期的增长率。为了反映银行对弱势企业的帮扶政策，本文基于毛利润、企业实力等 6 项指标对扶持力度进行了灰色关联评价，得到指标值。最终通过新的风险评价模型得到最后的年度信贷调整策略。

论文结构

该篇论文结构如下：

优缺点分析

• 该篇论文优点：

在对问题一的求解中，将银行的信贷策略与多种因素联系在一起，考虑了多项指标，并通过人工鱼群优化算法进行求解，使策略更具说服力。在问题二的求解中，该篇论采用了 XGBoost 集成算法对多分类问题进行了预测，并通过 SMOTE 过采样解决了数据集不均衡的问题，并与其它如随机森林、GBDT 等集成算法进行了比较，验证了 XGBoost 对此题预测结果的准确率较为优越。

在问题三的求解中，通过多种函数拟合不同行业在新冠疫情下的经济变化情况，且每种函数拟合优度较高、误差较小，充分表现了新冠疫情对不同行业的不同影响，并在最后引入了扶持力度来表现银行的社会救济责任，充分反映了实事政策的影响。

• 该篇论文缺点：

该篇论文运用灰色关联度评价模型来进行风险评估，相较于其他模型适用性较差。

大部分数据都来自于进销项发票信息，可能部分指标之间相关性较强，会有信息重复表达，对评价结果可能造成影响。

1.2.3 C227 论文赏析

论文解读

该篇论文该篇论文针对中小微企业的信贷决策，采用 SoftVoting 集成学习算法评估企业信贷风险，建立多目标规划模型，得到不同评级企业的贷款额度和贷款利率的信贷策略

针对问题一，对附件一中的数据进行特征提取，综合考虑了企业实力、发展潜力、上下游供求关系、企业抗风险能力四类指标，共提取出 20 个特征来衡量企业的信贷风险。本文以 Logistic 回归、Adaboost、GBDT、SVM 和随机森林为基分类器，建立 SoftVoting 集成学习算法，计算企业的违约风险。SoftVoting 在整个数据集上的综合准确率为 97.6%，ROC 曲线覆盖面积 (AUC 值) 达到 99.42%，接近完美模型，计算出的概率可信度高。根据附件一给出的信誉评级和得出的违约风险，根据银行的行为建立一个利润最高，风险和潜在损失最小的多目标规划模型，求解得出本问的银行信贷策略为对 D 级企业和有违约经历的企业拒绝提供贷款：剩下的企业中，对 A 级企业提供利率为 5.85%，额度为

999905 元的贷款；对 B 级企业提供利率为 7.5%，额度为 535178 元的贷款；对 C 级企业提供利率为 11%，额度为 100030 元的贷款。

针对问题二，由于本问为无信贷记录企业，银行需要需要根据贷款企业的已知信息和有信贷记录的信息对贷款企业进行信誉评级和违约风险判断。对附件二的数据做与附件一相同处理，根据提取特征，使用问题一训练的 SoftVoting 模型对附件二企业进行违约风险预测。将违约风险加入已有特征，用有信贷记录的企业相关数据训练 XGBoost 模型，实现对信誉评级的多分类预测。XGBoost 在训练集上综合准确率为 87.8%。完成信誉评级和违约风险判断后，就可以进行多目标规划模型求解，得出银行的信贷策略是：对 D 级企业和有违约经历的企业拒绝提供贷款；剩下的企业中，对 A 级企业提供利率为 7.05%，额度为 999717 元的贷款；对 B 级企业提供利率为 7.52%，额度为 260320 元的贷款；对 C 级企业提供利率为 11.01%，额度为 100048 元的贷款。

针对问题三，以新冠疫情这个突发事件为例，综合考虑企业的信贷风险和疫情对不同企业的不同影响。从企业在疫情中面临的系统性风险和非系统性风险的角度，在附件二中提取行业风险、企业经营状况、利润同比增长率、废票占比变动率和交易企业数量变动率 5 个特征，反映新冠疫情对企业的综合影响。运用系统聚类将企业聚成 3 类，不同类别的企业具有不同的风险乘数 Q ，风险乘数描述了疫情对不同类别企业的违约风险的放大作用。将信誉评级与聚类结果综合考虑，银行将对企业建立新的评价分级体系，将风险乘数 Q ，代入得到调整之后的多目标规划模型，求解可得出银行问题三情景下的信贷策略。

论文结构

该篇论文结构如下：

优缺点分析

• 该篇论文优点：

在对问题一的求解中，利用 SoftVoting 集成学习算法甲酸企业违约风险，模型准确率达到 97.6%，AUC 值接近完美模型，得到的违约风险可信度高。

在问题二的求解中，基于问题一用 XGBoost 模型实现对信誉评级的多分类预测，在训练集上准确率较高，给出的信贷策略具有可靠性。

在问题三的求解中，该篇论文考虑疫情对企业影响时综合考虑了系统风险和非系统风险影响，并将企业聚类为 3 类，考虑不同类别企业面对新冠疫情后违约风险的变化，考虑贴合实际，模型具有普适性。

• 该篇论文缺点：

该篇论文假设相同信誉评级的企业给予相同的贷款额度和利率，未考虑企业规模、实力等方面，考虑不全面，会使信贷策略的实用性较差；

在论文中问题一提取了 20 个特征来衡量企业的信贷风险，部分特征之间具有相关性，若减少相关指标会使模型简化，可能会得到更佳的结果；

在研究客户流失率与英航年利率关系时，没有考虑小范围年利率变化时的客户流失率，使结果有失偏颇。

2 模型假设

- 假设附件数据真实可信，可直接利用已知数据进行分析；
- 本文假设各企业的信息对等，即不会因为信息不对等而导致对其他条件相同的两个企业的信贷策略不同；
- 假设除因利率问题导致客户流失和因信誉等级问题导致银行不对其放贷的原因；
- 银行不会因为其他原因而中止贷款业务。

3 符号说明

符号	说明
x_1	年均毛利润
x_2	年均毛利润增长率
x_3	上游年均业务量
x_4	上游业务量增长率
x_5	下游年均业务量
x_6	下游业务量增长率
x_7	信誉评级
x_8	违约记录
S	企业实力指标
C	企业信誉指标
R	违约概率

注：表中未列出及重复的符号均以首次出现处为准。

4 改进问题一模型的建立与求解

4.1 问题一分析

问题一要求对附件一有信贷记录的 123 家企业的信贷风险进行量化，并给出银行的最优信贷策略。本题在求解过程中主要分为两个部分。

第一部分本文建立信贷风险评估模型。首先，对附件一数据进行分析处理；然后，从企业实力和企业信誉两个方面综合考虑信贷风险，并依据附件一数据提取得到八项数据

特征，分别为年均毛利润、年均毛利润增长率、上游年均业务量、上游业务量增长率、下游年均业务量、下游业务量增长率、信誉评级、违约记录，通过动态加权将前六项指标整合为企业实力综合指标，后两项整合为企业信誉综合指标;之后，本文用企业违约概率来衡量信贷风险，利用 Sigmoid 的函数根据企业实力 + 企业信誉指标求得各类型企业的违约概率，并根据违约率大小将企业划分成四个风险等级。

第二部分本文建立信贷策略优化模型。首先，研究客户损失率与银行年利率的关系，采用指数函数进行拟合求解;然后，以收益最大化和风险最小化为目标函数建立多目标规划模型，利用人工鱼群算法求解，从是否放贷、贷款比例、贷款年利率三个方面针对不同企业规模给出相应信贷策略。^[1]

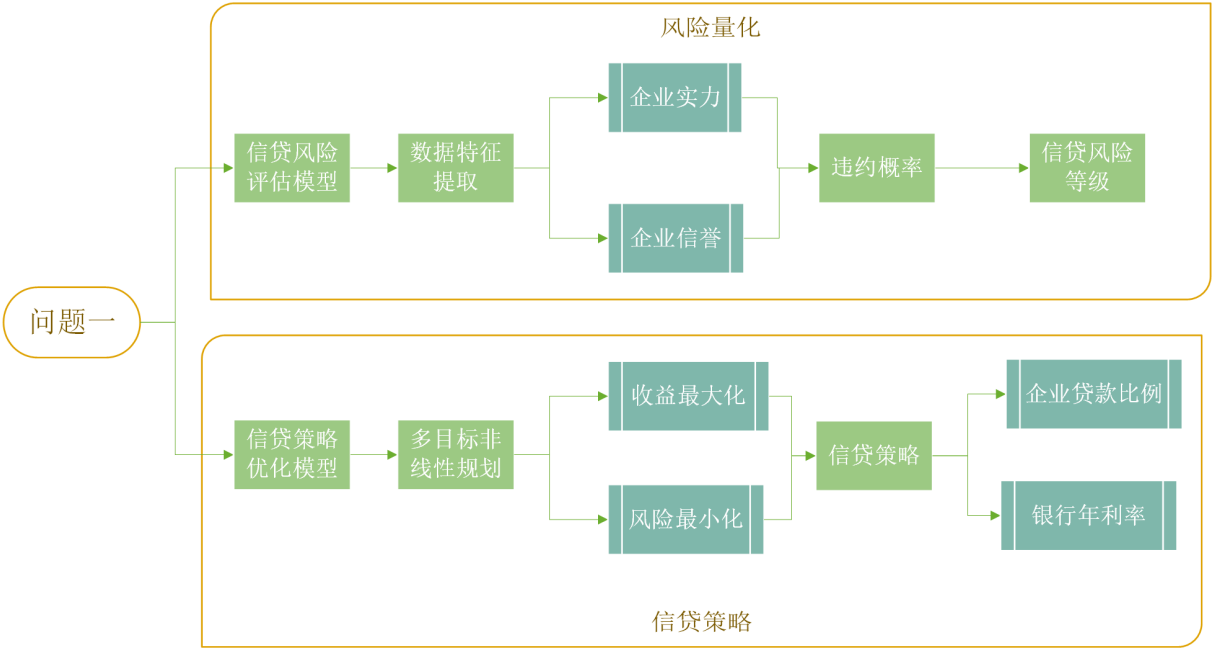


图 1 问题一思路图

4.2 信贷风险评估模型的建立与求解

信贷风险即信贷资金无法按照约定的还款方案按时归还的可能性，信贷风险的本质就是违约风险，因此本文用企业借贷的违约率来衡量中小微企业的信贷风险。在实际中，违约这一过程受主客观多种因素影响，面临诸多不确定性问题，来自企业自身、外部环境等因素都有可能导导致银行等金融机构无法收回贷款本息的风险，本文建立信贷风险评估模型试量化信贷风险。

4.2.1 数据分析处理

(1) 作废发票代表交易活动的取消，一般是由于发生错误、误填等情况产生，为不对后续风险评估产生干扰，本文剔除该部分数据；

(2) 分析数据发现附件一的企业数据之间的时间长度不同，部分年份数据缺失严重，后续计算指标数据需注意此问题。

4.2.2 数据特征提取

参考相关文献结合专家评分，中小微企业信贷风险受企业实力和企业信誉影响最大。综合附件一数据，本文选取年均毛利润、年均毛利润增长率、上游年均业务量、上游业务量增长率、下游年均业务量、下游业务量增长率六个指标来衡量企业实力，选取信誉评级和违约记录来衡量企业信誉。

由于题目所给为年度信贷，为更好的评价企业信贷风险，本文选取的指标都为年度指标。但在数据分析时发现不同企业数据的时间长度不同，为充分利用数据信息，增强结果的客观性和真实性，并使不同企业度量指标具有可比性，本文采用公式 (1) 对上述指标中的年平均值进行计算。

$$\bar{Y} = 12 \times \bar{M}. \quad (1)$$

其中 \bar{Y} 代表年平均指标值， \bar{M} 月平均指标值。

各项指标的具体意义和计算公式如下：

年均毛利润 $x_1 = (x_{1,1}, x_{1,2}, \dots, x_{1,n})$ ，可直接反映企业的实力，其计算公式为

$$x_{1,i} = \frac{E_i - I_i}{M_i} \times 12. \quad (2)$$

其中 n 表示附件企业数即 123， E_i 表示第 i 个企业的销项发票总金额， I_i 表示第 i 个企业的进项发票总金额， M_i 代表有第 i 个企业数据的月份数。

年均毛利润增长率 $x_2 = (x_{2,1}, x_{2,2}, \dots, x_{2,n})$ ，可反映企业前景发展状况，其计算公式为

$$x_{2,i} = \frac{12}{M_i} \sum_{k=1}^{M_i-1} \frac{p_{k+1,i} - p_{k,i}}{p_{k,i}} \quad (3)$$

其中 $p_{k+1,i}$ 表示第 i 个企业第 $k+1$ 个月的毛利润， $p_{k,i}$ 表示第 i 个企业第 k 个月的毛利润。

上游年均业务量 $x_3 = (x_{3,1}, x_{3,2}, \dots, x_{3,n})$ ，可从侧面反映企业实力，业务量越大，一定程度上偿还贷款的能力越强，银行可适当增大额度减小利率。其计算公式为

$$x_{3,i} = \frac{T_i}{M_i} \times 12 \quad (4)$$

其中 T_i 表示第 j 个企业上游交易总次数。

上游业务量增长率 $x_4 = (x_{4,1}, x_{4,2}, \dots, x_{4,n})$, 可反映企业的发展潜力, 计算公式为

$$x_{3i} = \frac{12}{M_i} \sum_{k=1}^{M_i-1} \frac{t_{k+1,i} - t_{k,i}}{t_{k,i}} \quad (5)$$

其中 $t_{k+1,i}$ 表示第 i 个企业第 $k+1$ 个月的交易次数, $t_{k,i}$ 表示第 i 个企业第 k 个月的交易次数。

下游年均业务量 $x_5 = (x_{5,1}, x_{5,2}, \dots, x_{5,n})$, **下游业务量增长率** $x_6 = (x_{6,1}, x_{6,2}, \dots, x_{6,n})$ 的计算公式与上游相同。

信誉评级 $x_7 = (x_{7,1}, x_{7,2}, \dots, x_{7,n})$, 将该指标量化来衡量公司信誉情况, 其公式为

$$x_{7,j} = \begin{cases} 5, & \text{信誉评级:A} \\ 3, & \text{信誉评级:B} \\ 1, & \text{信誉评级:C} \\ 0, & \text{信誉评级:D} \end{cases} \quad (6)$$

违约记录 $x_8 = (x_{8,1}, x_{8,2}, \dots, x_{8,n})$, 是反映企业信誉的重要指标。将其设置为 0—1 变量, 即

$$x_{8,i} = \begin{cases} 1, & \text{企业无违约记录;} \\ 0, & \text{企业有违约记录.} \end{cases} \quad (7)$$

4.2.3 风险评估指标

• 数据标准化

由于上述提取的指标量级差异较大, 本文对其进行 min-max 标准化处理, 即

$$x^* = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (8)$$

• 企业实力指标

年均毛利润、年均毛利润增长率、上游年均业务量、上游业务量增长率、下游年均业务量、下游业务量增长率八项数据大小一定程度上直接或间接地反映企业实力的强弱。因此, 本文将通过整合各数据得到衡量企业实力的综合指标, 将其作为企业信贷风险的评估指标之一。

考虑到数据间的“质差”和“量差”, 在确定综合指标时, 既要能体现不同数据项对企业实力影响的差异, 也要能体现所有指标共同影响企业实力, 本文采用动态加权法来整合企业实力指标, 某一指标的权重可以根据值变化而动态改变。

联系实际情况, 随着指标增长率的提高, 指标的增长速度越来越快, 对企业实力的

影响也会随之逐渐加强，其影响效果可近似为一条“S”曲线。因此，本文用 S 型分布函数来进行动态加权求取企业实力指标，即

$$\lambda(x) = \begin{cases} 2 - e^{-3x}, & x \geq 0; \\ e^{-3|x|}, & x < 0. \end{cases} \quad (9)$$

从而，企业实力指标可以用年均毛利润、上、下游业务量三项指标表示为

$$S_i = \lambda(x_{1,i})x_{2,i} + \lambda(x_{3,i})x_{4,i} + \lambda(x_{5,i})x_{6,i} \quad (i = 1, 2, \dots, n) \quad (10)$$

• 企业信誉指标

综合考虑是否有信贷违约记录和信誉评级两项指标来衡量企业信誉，对于有违约记录的企业信誉指标值记为 0，对于无违约记录的企业用评级来作为企业信誉指标值，即

$$C_i = x_{7,i} \times x_{8,i} \quad (i = 1, 2, \dots, n) \quad (11)$$

4.2.4 信贷风险评估

信贷风险由企业实力和企业信誉两项指标决定，其数据均已根据上述步骤求得，需将两项指标进行合并进行风险量化分析，则企业实力 + 企业信誉为

$$SC = (SC_1, SC_2, \dots, SC_n), \quad SC_j = S_j C_j. \quad (12)$$

类似于 Logistic 回归模型，本文将合并的指标带入 Sigmoid 函数即

$$SC'_j = \begin{cases} \frac{1}{1 + e^{-SC_j}}, & SC_j \neq 0 \\ 0, & SC_j = 0 \end{cases} \quad (13)$$

并利用如下公式计算各信誉评级的企业的违约概率，作为企业信贷风险指标 $R = (R_1, R_2, \dots, R_n)$ 。企业的实力 + 信誉指标 SC 越高，信贷风险指标越小，信贷风险越小。

$$R_j = 1 - SC'_j \quad (j = 1, 2, \dots, n) \quad (14)$$

4.2.5 企业规模划分

根据国家统计局发布的《关于印发〈统计上大中小微型企业划分办法(2017)〉的通知》，在企业规模上按照行业门类、大类、中类和组合类别，依据从业人员、营业收入、资产总额等指标或替代指标，将我国的企业划分为大型、中型、小型、微型等四种类型。但本题附件一的 123 家企业的相关信息有限，本文决定以上下游企业业务量总和为指标

对企业规模进行划分，并且依据题目所给企业已都为中小微企业，本文将对其进行进一步细分，划分为中型、中小型、小型、微型 4 个类别。

表 1 企业规模划分标准

上下游企业年均业务量总和	企业规模
[0,500)	微型
[500,1000)	小型
[1000,2500)	中小型
2500 及以上	中型

本文统计了不同规模企业的信誉评级情况如下，发现企业规模与信誉评级存在正相关关系。信誉评级为 D 的企业几乎全部为集中在微型企业，说明企业规模较小信贷风险可能更高。

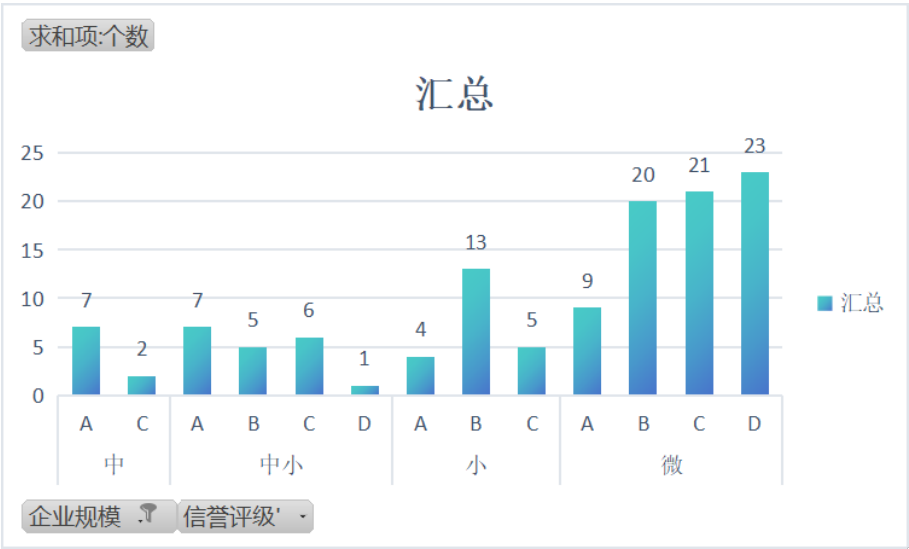


图 2 不同规模企业的信誉评级情况

4.2.6 结果分析

基于该风险量化模型，本文解出了 123 家企业的信贷风险指标，部分数据如下表，完整数据见附录 A。

表 2 问题一 123 家企业风险指标量化表（部分）

企业代号	实力指标	信誉指标	规模	风险指标
E1	0.761236661	5	中	0.043498647
E2	3.651396983	5	中	2.35571E-08
E3	0.628823142	1	中	0.695554869
E4	0.107520882	1	小	0.946291292
E5	0.235843301	3	中小	0.660289742
E6	1.833068448	5	中	0.000209183
E7	0.474017734	5	中	0.170964414
E8	1.204345287	5	中	0.00483922

基于建立的信贷风险评估模型求解得到各企业的信贷风险指标 R 数值，并按照大小划分信贷风险级别，结果如下

表 3 企业信贷风险等级结果

信贷风险 R	信贷风险等级	企业数量
$[0,0.4)$	1 级	47
$[0.4,0.8)$	2 级	25
$[0.8,1)$	3 级	24
1	4 级	27

本文统计了不同信誉评级和不同企业规模的平均违约概率如下。



图 3 不同信誉评级企业的平均违约概率

由上图可知该信贷风险模型所得违约概率与企业信誉评级较为相符，信誉评级为 A、B 的企业的违约概率均在 0.5 以下，信誉评级为 C、D 的企业违约概率较高，由于信誉评级为 D 的企业全部有违约记录，所以违约概率为 1. 而且可以看出信誉评级与企业的平均违约概率呈负相关关系，证明了评估模型的合理性和可靠性。



图 4 不同规模企业的平均违约概率

由上图可以看出小型和微型企业的违约概率较大，中型和中小型企业的违约率较小，这与实际相符。企业规模大小一定程度上反映了企业实力，企业规模较小可能资金周转能力较弱、销售渠道较少，这些原因都可能导致企业资金链出现问题，导致最终无法偿还贷款。但由上述统计可以看出微型企业占比很大，是银行的主要客户，银行应对这部分客户进行更加深入的风险研究制定合适的信贷策略。

4.3 信贷策略优化模型的建立与求解

4.3.1 客户流失率与年利率关系

为模型简洁，本文假设客户流失率仅与利率有关。通过分析附件三数据并作图发现客户流失率与利率之间呈现负相关关系，而且二阶导数小于 0，随着利率的增大客户流失率最终趋向于 1。因此，本文选取如下公式对附件三中 A,B,C 三类信誉评级的客户流失率与利率做拟合，并与 C109 的三次函数拟合作对比。

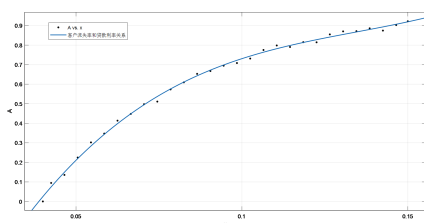
$$l_i = 1 - e^{-a_i(r+b_i)} (i = 1, 2, 3), \quad (15)$$

其中 $i = 1, 2, 3$ 分别表示企业信誉评级为 A,B,C; l 表示客户流失率， r 表示银行贷款年利率； a 和 b 表示待求系数。

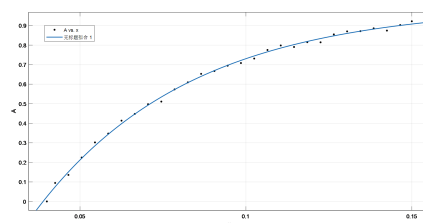
拟合结果如下所示。

表 4 客户流失率与银行贷款年利率的拟合关系

企业信誉评级	a	b	调整 R 方（指数函数）	调整 R 方（三次函数）
A	21.4553	-0.0389	0.9977	0.9975
B	18.8656	-0.0388	0.9981	0.9980
C	18.5321	-0.0393	0.9979	0.9979



(a) 三次函数



(b) 指数函数

图 5 信誉评级为 A 的客户流失率与银行年利率的拟合关系图像

以信誉评级 A 的企业为例，尽管三次函数的拟合度效果已经很好并且两者的调整 R 方也相差不大，但是从散点图上可看出，在银行贷款年利率大于 0.15 时，三次函数出现斜率激增的情况，若后续银行提高年利率上限，客户流失率将会大于 1，这显然不符实际。因此，指数函数相较于 C109 的三次函数拟合效果更佳，可用性更强。

4.3.2 多目标非线性规划模型

在信贷风险评估模型中已求得各企业的违约概率，当企业违约概率过大时，企业极有可能无法还款，导致银行遭受损失。因此，本文对信贷风险 R 为 1 的企业实行不予放贷策略，仅对信贷风险等级为 1、2、3 的三类企业进行信贷策略优化。

• 决策变量

针对不同信贷风险等级的企业，银行应设置不同的贷款额度。根据题目年度信贷总额固定，本文用 $\lambda_j (j = 1, 2, \dots, n_0)$ 表示不同企业的贷款比例。

• 目标函数

(1) 收益最大化——对于银行来说，向企业放贷的最直接目的就是获得利息回报。即要从未损失的客户中获取最大利息收益：

$$\max Y = \sum_{i=1}^{n_1} \sum_{j=1}^{n_0} A \lambda_j (1 - l_j) r_i \quad (16)$$

其中， n_0 表示信贷风险等级为 1、2、3 的三类企业的数量； n_1 表示企业三个信贷风险评级数目，由于评级为 D 的企业全部违约，信贷风险级别为 1 未被研究，余下企业评级

都为 A 、 B 、 C ，因此 n_1 取值 3； A 表示确定要放贷企业的贷款最大额度即 100 万元； l 表示客户流失率； r 表示银行贷款年利率。

(2) 风险最小化——若企业违约，银行不仅无法得到利息，还会损失本金。为银行损失最低，要求信贷风险最小化，本文取单位贷款额度的信贷风险最小：

$$\min W = \frac{\sum_{j=1}^{n_0} (1 - l_j) R_i}{\sum_{j=1}^{n_0} A x_j (1 - l_j)} \quad (17)$$

(3) 化多目标为单目标——本文采取线性加权的方式将两个目标化为一个目标函数。首先，将通过取倒数将目标函数 Y 最小化，使其与 W 相同；然后，给每个目标各赋予 50% 的权重。

$$\min Z = \frac{1}{2} \left(\frac{1}{Y} + W \right) \quad (18)$$

• 约束条件

(1) 信贷风险约束——为确保信贷风险不能过大，本文要求未流失的客户得信贷风险不能超过信誉风险为 1、2、3 级的三类企业信贷风险平均值，即

$$\overline{R_i} = \frac{1}{N_i} \sum_{j=1}^{N_i} R_i (i = 1, 2, 3) \quad (19)$$

$$\sum_{j=1}^{N_i} (1 - l_j) R_i \leq N_i \overline{R_i} \quad (20)$$

其中 N_i 三表示类企业的数量， $\overline{R_i}$ 表示三类企业信贷风险平均值。

(2) 企业贷款比例约束——不同信贷风险等级的企业贷款额度不同，风险等级越高，贷款额度越小，贷款比例越低。根据银行对确定要放贷企业的贷款额度为 10 至 100 万元，所有企业的贷款比例要 $[0.1, 1]$ 范围内。针对不同风险等级企业，本文作如下约束

$$\begin{cases} 0.1 \leq \lambda_j \leq 1.0, \text{信贷风险1级;} \\ 0.1 \leq \lambda_j \leq 0.9, \text{信贷风险2级;} \\ 0.1 \leq \lambda_j \leq 0.6, \text{信贷风险3级.} \end{cases} \quad (21)$$

其中 λ_j 表示企业 j 的贷款比例。

(3) 贷款额度约束——所有企业得贷款金额不能超过银行年度信贷总额，本文将问

题一的年度信贷总额设置为 5000 万，即

$$\sum_{j=1}^{n_0} A\lambda_j(1-l_j) \leq B \quad (22)$$

其中 B 表示银行年度信贷总额。

(4) 年利率约束——根据题目要求，银行年利率要在 $[4\%, 15\%]$ 范围内，即

$$r_0^1 \leq r_i \leq r_0^2 \quad (23)$$

其中， r_0^1 表示年利率下限即 4%， r_0^2 表示年利率上限即 15%。

• 模型综述

$$\min Z = \frac{1}{2} \left[\frac{1}{\sum_{i=1}^{n_1} \sum_{j=1}^{n_0} A\lambda_j(1-l_j)r_i} + \frac{\sum_{j=1}^{n_0} (1-l_j)R_i}{\sum_{j=1}^{n_0} A\lambda_j(1-l_j)} \right] \quad (24)$$

$$\left\{ \begin{array}{l} \sum_{j=1}^{N_i} (1-l_j)R_i \leq N_i \bar{R}_i; \\ \sum_{j=1}^{n_0} A\lambda_j(1-l_j) \leq B; \\ r_0^1 \leq r_i \leq r_0^2 \\ l_j = 1 - e^{-a_i(r+b_i)}; \\ 0.1 \leq \lambda_j \leq 1.0, \text{ 信贷风险1级;} \\ 0.1 \leq \lambda_j \leq 0.9, \text{ 信贷风险2级;} \\ 0.1 \leq \lambda_j \leq 0.6, \text{ 信贷风险3级;} \\ i = 1, 2, n_1; j = 1, 2, \dots, n_0. \end{array} \right. \quad (25)$$

4.3.3 人工鱼群算法求解

人工鱼群算法（AFSA）是一种基于模拟鱼群觅食、追尾和聚群等行为来搜索最优解的优化算法。在算法中，每个解被看作是一个鱼的位置，鱼会通过观察周围环境和与其他鱼的交互来不断调整自己的位置，所有鱼的位置构成了问题的解空间。其基本步骤包括初始化鱼群的位置和速度、计算适应度函数、鱼的移动、鱼的群聚和追尾等。通过不断迭代和优化，人工鱼群算法可以在解空间中搜索到较优的解。人工鱼群算法在求解优化问题中具有较好的全局搜索能力和快速收敛性，常用于组合优化、函数优化等问题。本文利用人工鱼群智能算法的循优得到银行受益最大、风险最小的最优解，从企业贷款额度和贷款利率两个方面给出放贷方案。

4.3.4 结果分析

在已有多目标非线性规划模型基础上，本文采用人工鱼群算法对最优方案进行了探寻，算法迭代曲线图如下图所示，可以看到算法在迭代 50 次后趋于平稳，即已搜寻至最优解。

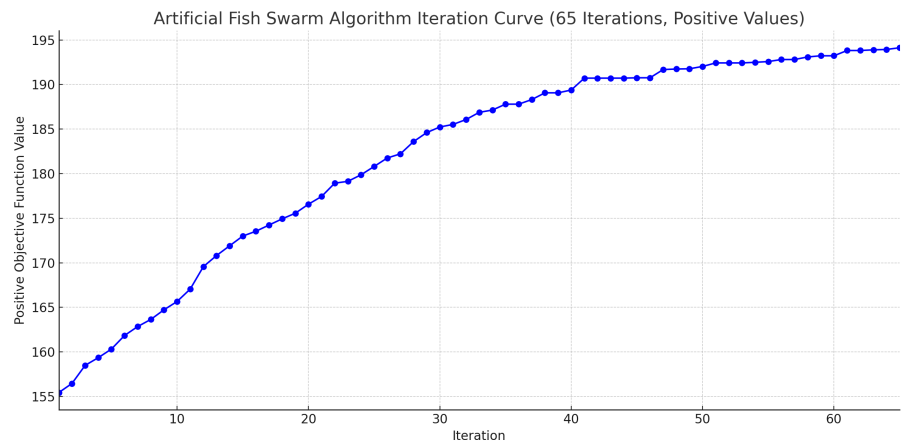


图 6 人工鱼群算法迭代图

经求解，得到银行受益最大、风险最小的最优信贷策略，在该方案中银行总利息收益为 194.43 万元，单位贷款额度的风险值 0.0302，信贷风险指标值小于 0.05，说明该方案较佳。对各类型企业制定的贷款比例以及贷款利率如下图所示：

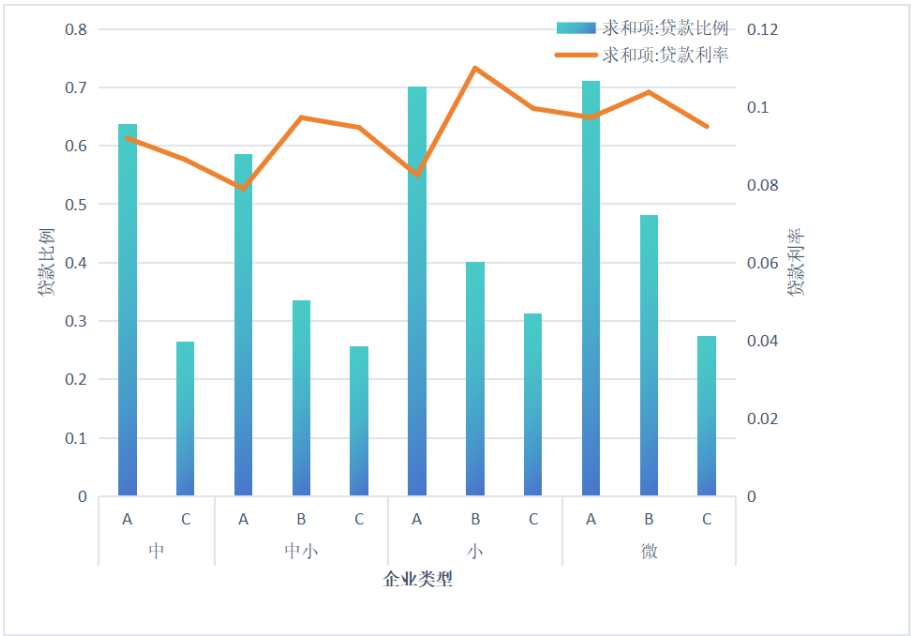


图 7 各类型企业的信贷策略

本文针对银行的信贷策略主要是从是否放贷、贷款额度、贷款利率三个方面进行考虑。为保证银行信贷风险小、损失少，对于企业风险等级为 4 和风险评级为 D 的企业不

予放贷。对于其他企业综合考虑风险小、收益高两个目标和客户损失问题，针对不同企业规模和信誉评级予以不同的贷款额度比例和贷款利率。这即是本文给出的最优信贷策略，最终计算得到附件一得到不违约的 96 家企业的贷款比例和贷款利率，可见附录 B。

5 改进的问题二模型的建立与求解

5.1 问题二分析

问题二要求对附件二无信贷记录的 302 家企业的信贷风险进行量化，并给出银行的最优信贷策略。通过问题一我们可知企业的违约概率与该企业信誉评级和违约记录存在高度关系，因此在进行风险评估之前，本文基于附件一数据利用引入 Dropout 的 XGBoost 模型，确定附件二风险评级并评判违约记录；之后，本文利用基于问题一的信贷风险评估模型和信贷策略优化模型针对不同类型企业设置相应的信贷策略。

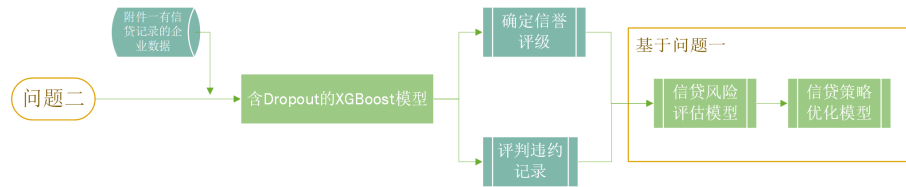


图 8 问题二思路图

5.2 预备工作

5.2.1 信誉等级预测指标选取

根据问题一计算的指标有年均毛利润、年均毛利润增长率、上游年均业务量、上游业务量增长率、下游年均业务量、下游业务量增长率，问题二中依旧利用这些指标作为信誉等级预测指标，并且还增加发票作废率、退货率两个指标。**发票作废率**代表着交易活动取消，若作废发票数越多，则表明交易活动取消数越多，表明该企业的运营越不稳定，相应表明该企业的信誉等级越低，因此作废发票率是企业信誉的负向指标，其计算公式为

$$\text{发票作废率} = \frac{\text{作废发票数量}}{\text{总销项发票数量}} \times 100\% \quad (26)$$

退货率用于衡量一定时期内退回商品的数量占销售总数量的百分比。它提供了有关客户满意度和企业信誉的重要信息，其计算公式为

$$\text{产品退货率} = \frac{1}{3} \cdot \frac{\sum_{\text{年份}=2017}^{2019} \text{负销项发票数量}}{\sum_{\text{年份}=2017}^{2019} \text{总销项发票数量}} \times 100\% \quad (27)$$

5.2.2 数据过采样

由于一些类别的企业数与其他类别企业相比过少，可能会造成数据类别不平衡问题，不平衡样本会导致最终训练模型侧重样本数目较多的类别，而“轻视”样本数目较少类别，这样模型在测试数据上的泛化能力就会受到影响。因此，为解决不平衡样本问题，需对数量少的数据进行过采样操作。

本文采用 SMOTE 合成少数类过采样技术对数据进行过采样，这是在随机采样的基础上改进的一种过采样算法^[3]，其实现过程为：

1. **选择少数类样本**：从少数类样本中随机选择一个样本 x 。
2. **找到其 k 个最近邻**：使用 K 近邻算法找到该样本的 k 个最近邻居，记为 $N_k(x)$ 。
3. **从 k 个最近邻中随机选择一个邻居**：从 $N_k(x)$ 中随机选择一个邻居 n 。
4. **在所选样本和所选邻居之间生成新样本**：选择两个样本之间的差异，并乘以一个在 0 到 1 之间的随机数，然后将其加到所选样本上。

$$\text{差异} = n - x,$$

$$\text{随机数} \sim U(0, 1),$$

$$\text{新样本} = x + \text{随机数} \cdot \text{差异}.$$

数据过采样前后的数据分布如图所示

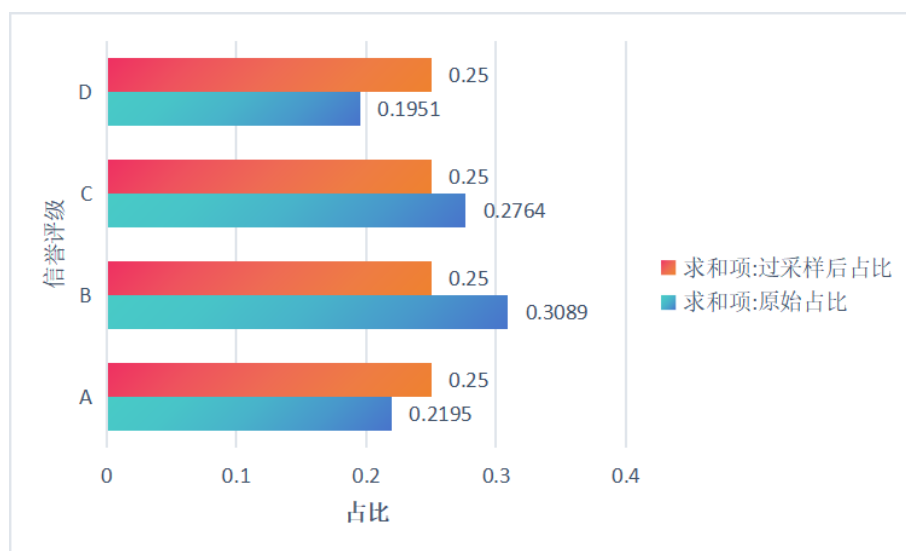


图9 过采样数据与原始数据占比对比图

由图可以看出，数据过采样后，各类企业数据量差别减少，数据类别达到平衡，从而避免模型在测试数据上的泛化能力受到影响。

5.3 基于 Dropout XGBoost 算法的信誉评级确定模型

5.3.1 模型建立

通过问题一求解过程和结果可知，企业的信贷风险与该企业的信誉等级之间存在高度正相关关系。但是附件二企业无信贷记录，根据现有数据难以直接评估风险。因此，本文为方便对附件二企业的信贷风险进行评估，需要先建立模型以确定企业的信誉评级和是否有违约记录。

本文需要找出附件一有信贷记录企业的信誉评级与企业实力指标的关系，从而利用附件二中企业实力指标以确定企业的信誉评级。实质是将附件二企业按照信誉评级 A、B、C、D 来进行分类，其本质可以看作是一个分类问题。经查阅资料发现 XGBoost 在分类问题上表现良好，本文采用 XGBoost(eXtreme Gradient Boosting) 来建立集成学习模型，并引入 Dropout 以避免出现过拟合的情况。^[5]

5.3.2 模型求解

XGBoost 是一种基于决策树的集成学习算法，其核心思想是，通过迭代地训练多个弱分类器（即决策树），并将它们组合成一个强分类器。在每一轮迭代中，XGBoost 可以根据上一轮的分类结果计算出每个样本的权重，并使用这些权重训练下一轮的分类器。随着迭代次数的增加，XGBoost 可以不断地改善分类的准确性，直到达到最优的分类结果。

利用问题一信誉评级 x_7 和违约记录 x_8 的量化结果，并用以下 O 和 P 分别表示附件二企业的信誉评级和违约记录。

$$O = (O_1, O_2, \dots, O_{n'}) \quad (28)$$

$$P = (P_1, P_2, \dots, P_{n'}) \quad (29)$$

其中 n' 表示附件二企业数目即 302。

下面以信誉评级为例介绍 XGBoost 模型求解过程，违约记录与其求解过程相同，不再重述。

(1) 建立信誉评级与企业实力的决策树模型，优化函数为

$$O_j = \varphi(x_1, x_2, x_3, x_4, x_5, x_6) = L(t) \quad (30)$$

$$L(t) = \sum_{i=1}^{n'} l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \delta(f_t) + \text{constant} \quad (i = 1, 2, \dots, 6) \quad (31)$$

其中 $L(t)$ 表示第 t 次迭代的优化函数； $x_1, x_2, x_3, x_4, x_5, x_6$ 分别表示附件二企业年均毛利润、年均毛利率增长率、上游年均业务量、上游业务量增长率、下游年均业务量、下

游业务量增长率; l 为损失函数, $constan$ 表示常数项; $\delta(f_t)$ 表示正则项。

(2) 为防止过拟合, 本文在 XGBoost 算法中引入 Dropout。在训练迭代过程中, 随机剪去部分枝叶训来避免过拟合的发生, 过程如下: Bernoulli 是概率 p 的伯努利分布, 取值为 0 或 1, 即该分布中的每一个元素取值为 0 的概率是 p , 当使用 dropout 时, 第 t 次迭代的节点样本集合乘以概率为 p 的 Bernoulli 分布, 即剪去部分枝叶。

$$r_i^t \sim \text{Bernoulli}(p) \quad (32)$$

$$\hat{y}_i^t = \hat{y}_i^{t-1} + f_t(x_i) \quad (x_i \in r_i^t I) \quad (33)$$

其中, I 表示节点样本集合。

(3) 移除 $L(t)$ 的常数项, 对其进行二阶泰勒展开得

$$\tilde{L}(t) \cong \sum_{j=1}^T \left[\left(\sum_{i \in I_j} \Phi_i \right) w_j + \frac{1}{2} \left(\sum_{i \in I_j} \Psi_i + \lambda \right) w_j^2 \right] + \gamma N \quad (34)$$

其中 T 表示迭代次数, Φ_i 表示 $\hat{y}(t-1)_i$ 的一阶导数, Ψ_i 表示 $\hat{y}(t-1)_i$ 的二阶导数, N 为叶子节点个数, w_j^2 为每个叶子节点的值的平方, I_j 为每个叶子节点上的样本集合, λ 和 γ 为防止过拟合设置的比重系数。

5.4 信贷风险评估模型

通过 XGBoos 模型求解得到的企业风险评级和违约记录, 按照问题一的信贷风险评估模型对附件二的 302 家企业进行风险量化, 方法与问题一相同, 不再赘述。

5.5 信贷风策略优化模型

问题二规定了银行年度信贷总额为 1 亿元, 将问题一中 B 取值改为 10000, 其余约束条件、决策变量、目标函数与问题一相同, 调整后规划模型为

$$\min Z = \frac{1}{2} \left[\frac{1}{\sum_{i=1}^{n_1} \sum_{j=1}^{n_0} A \lambda_j (1 - l_j) r_i} + \frac{\sum_{j=1}^{n_0} (1 - l_j) R_i}{\sum_{j=1}^{n_0} A \lambda_j (1 - l_j)} \right] \quad (35)$$

$$\left\{ \begin{array}{l} \sum_{j=1}^{N_i} (1 - l_j) R_i \leq N_i \bar{R}_i; \\ \bar{R}_i = \frac{1}{N_i} \sum_{j=1}^{N_i} R_i; \\ \sum_{j=1}^{n_0'} A \lambda_j (1 - l_j) \leq B; \\ r_0^1 \leq r_i \leq r_0^2 \\ l_j = 1 - e^{-a_i(r+b_i)}; \\ 0.1 \leq \lambda_j \leq 1.0, \text{ 信贷风险1级;} \\ 0.1 \leq \lambda_j \leq 0.9, \text{ 信贷风险2级;} \\ 0.1 \leq \lambda_j \leq 0.6, \text{ 信贷风险3级;} \\ i = 1, 2, n_1; j = 1, 2, \dots, n_0'. \end{array} \right. \quad (36)$$

其中 n_0' 附件二中信誉级别为 1、2、3 的企业数量。

5.6 结果分析

对于缺失信贷记录的企业，银行在制定信贷策略时，需要先对其确定信誉评级、评判违约记录，才能进行相应的信贷风险评估和信贷策略优化。基于上述模型求解得到 $E217$ 、 $E402$ 、 $E404$ 、 $E420$ 、 $E422$ 、 $E423$ 、 $E425$ 存在违约记录，其他企业无违约记录。不同规模企业的信誉评级情况如下图所示，详细结果见附件 C

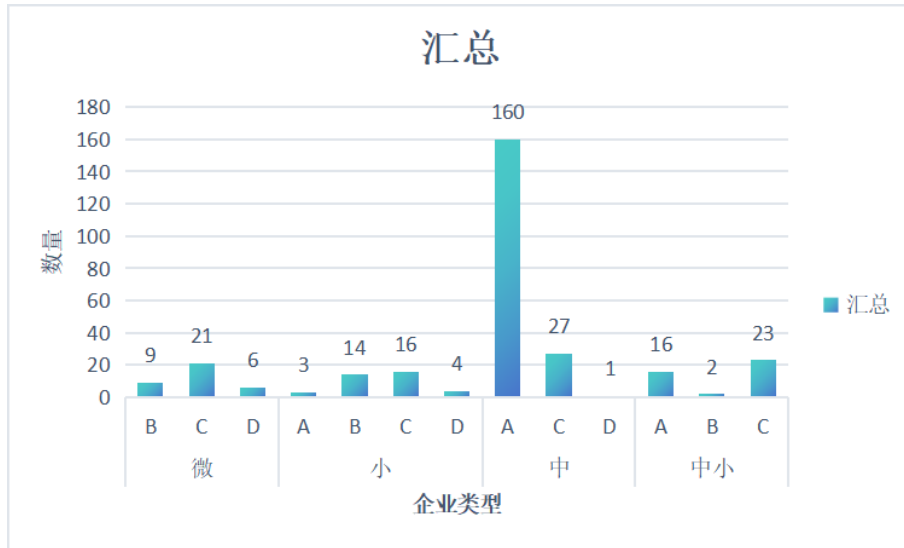


图 10 不同规模企业的信誉评级情况

通过求风险评估模型得到各信贷风险级别，结果如下通过求优化模型得到银行受益最大、风险最小的最优信贷策略，其中前 8 个企业对应的贷款额度、贷款利率如下表所示

表 5 问题二前 8 个企业信贷策略信息表

企业代号	贷款额度/万元	贷款利率
E124	29.1637126	0.081132056
E125	26.3990634	0.075034686
E126	29.2053792	0.0790967
E127	29.0483052	0.143656194
E128	15.6803373	0.111774588
E129	81.7173663	0.048780825
E130	87.893454	0.110831247
E131	39.5468749	0.069097392

在该方案中银行总利息收益为 412.67 万元，对各类型企业制定的贷款比例以及贷款利率如下图所示：

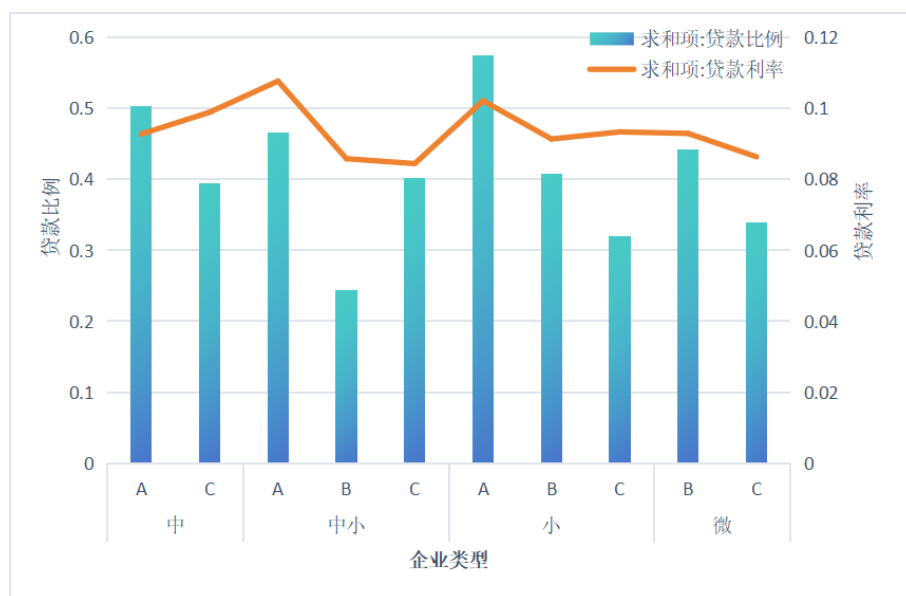


图 11 各类型企业的信贷策略

6 改进的问题三模型的建立与求解

6.1 问题三分析

问题三要求综合考虑信贷风险和可能的突发因素对各企业的影响，给出该银行在年度信贷总额为 1 亿元时附件二的信贷调整策略。本文以新冠疫情为例，从宏观和微观两个层面考虑突发因素对不同类型企业的影响。考虑国家政策补贴，对信贷策略优化模型进行调整，将增大企业贷款额度和减小借贷利率加入到约束中，并将求解结果与第二问对比，最后从企业、银行、社会三个层面分析调整后信贷策略效益。

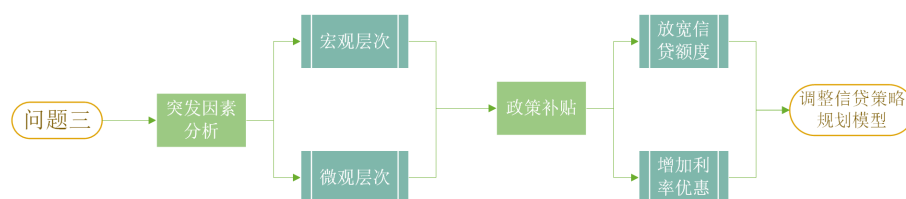


图 12 问题三思路图

6.2 突发因素对企业的影响分析

问题三需考虑突发事件对企业的影响。中小微企业由于企业规模较小，普遍存在资金链脆弱、风险承受能力差等问题。面对突发事件时，若没有外界援助，很容易出现资金周转不济、资金链断裂问题，最终导致破产。中小微企业仅依靠自身难以应对突发事件。

影响企业的突发因素可分为两类，分别是外界宏观层面和企业自身微观层面。宏观层面的突发因素一般是由经济结构失衡、生产与消费矛盾、政府宏观调控失误、全球性经济危机等不可控的条件引起的。微观层面的突发因素一般是产品质量缺陷、供应链断裂等问题，产生原因常是企业自身和上下游企业问题。本文将突发事件确定为 2020 年发生的新冠疫情，从宏观和微观两个层面上分别研究对信贷风险和信贷策略的影响及企业的抗突发能力。

6.2.1 宏观层面

为分析新冠疫情对各个行业经济的影响，本文按照《国民经济行业分类 GB T4754-2002》根据将企业按照如下标准划分第一、二、三产业类别，观察三次产业新冠疫情前后的经济上变化。

表 6 三次产业划分标准

产业类别划分	
第一产业	农、林、牧、渔业
第二产业	采矿业，制造业，电力、燃气及水的生产和供应业，建筑业
第三产业	除上述第一、第二产业以外的其他各业

根据国家统计局发布的《中华人民共和国 2020 年国民经济和社会发展统计公报》得到三大产业投资占固定资产投资的比重如下图所示，各行业的固定资产投资变化如下表所示。^[4]

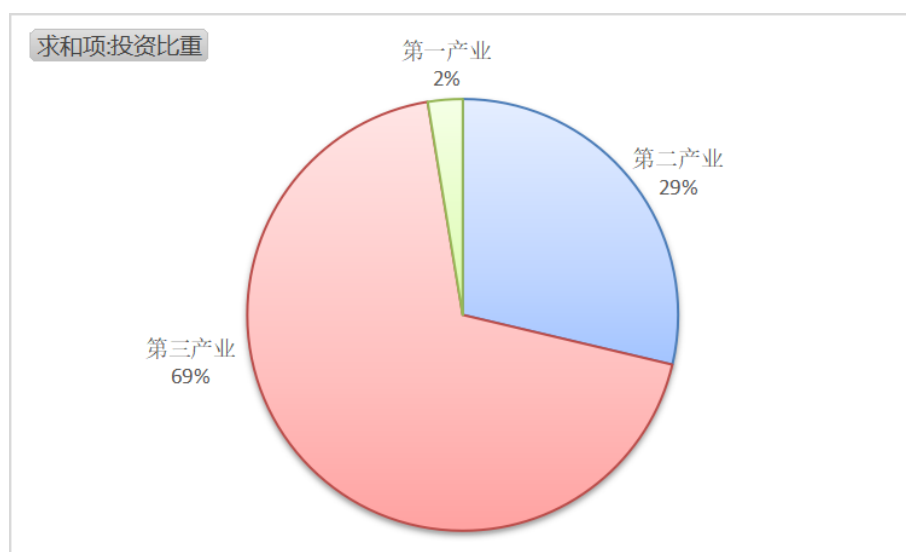


图 13 2020 年三大产业投资比重

固定资产投资是以货币形式表现的、企业在一定时期内建造和购置固定资产的工作量以及与此有关的费用变化情况。包括房产、建筑物、机器、机械、运输工具、以及企业用于基本建设、更新改造、大修理和其他固定资产投资等。企业的固定资产投资可以较好的反映该企业实力，其增长率的变化可以反映该企业的应对突发事件的能力。

本文还搜集了 2018 年至 2021 年的零售业、建筑业、工业、服务业与前年相比增长速度的变化折线图，如下图所示

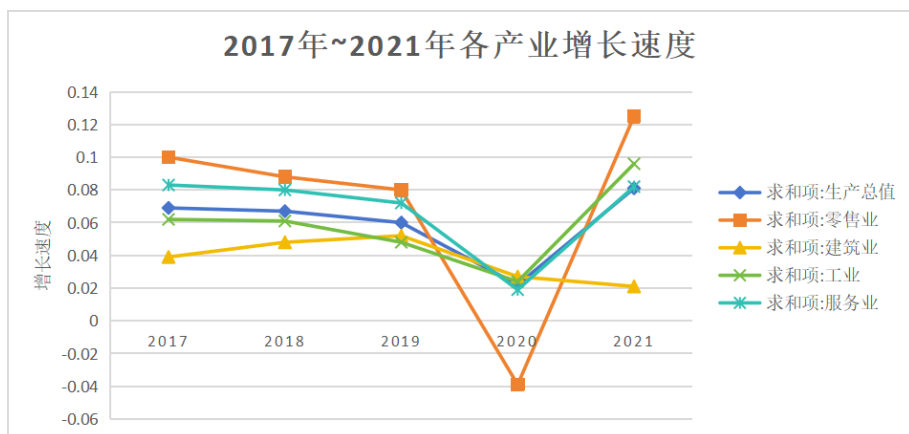


图 14 部分行业 2018 年至 2021 年的增长速度变化图

可见疫情的爆发对大部分行业都造成了不同程度的打击。特别是对零售业来说，2020 年的增长速度下降到了-0.039，许多零售实体店被迫关停，传统商超、便利店和街边店为保障民生需求维持开业，但流量依旧下行；银行贷款利息开支等使得现金流周转承压，经营利润大幅下跌，直接挑战企业生存。但也有部分行业趋势向好，比如医疗行业和科技行业。为研发应对病毒的检测、抑制、治疗的药品和试剂、研发合适的流感疫苗预防病毒，全国对医疗行业极度重视，同时全国人民的保健意识提高，进一步加强了健康防护，对医疗产品的购买也更加积极。同时，在疫情防控过程中科技行业也发挥了极大的作用，国家重视科技创新，许多科技行业的企业在疫情期间得到了更大的发展机遇。

6.2.2 微观层面

本文随机抽取附件二的 50 家企业进行分析，计算 2018、2019、2020 三年 1 月份和 2 月份的毛利润及其同比增长率，并对企业类别进行划分。

表 7 问题三部分企业 2020 年 1 月 2 月与前两年的同比增长率表

企业代号	2018 年	2019 年	2018 年	2019 年	行业
	1 月	1 月	2 月	2 月	
E141	-18.00%	5.69%	-38.81%	-60.07%	零售业
E171	435.05%	7.92%	-99.87%	-99.06%	建造行业
E176	-48.28%	23.72%	57.63%	236.83%	科技行业
E225	21.65%	22.79%	-91.54%	-193.93%	服务业
E311	26.48%	53.78%	-102.24%	-114.56%	服务业
E420	-99.91%	-99.90%	-34.60%	67.07%	医疗行业

从统计的结果可以看出，2020 年 1 月的利润增长和减少相对保持平衡，可看出此

时各行业还未受到疫情的影响。但 2020 年 2 月相比往年表现出差别，绝大多数企业出现利润负增长的情况，可看出大部分企业受疫情影响遭受损失，甚至濒临破产，急需外界的扶持渡过难关；当然也存在某些科技行业、医疗行业的相关企业出现一定的利润提升。

6.2.3 银行受疫情影响分析及贷款调整政策建立

经数据统计，旅游、住宿、餐饮、零售、交通还输、娱乐文化、线下教育等行业受疫情影响比较直接，这些行业多以中小企业为主，企业营业能力、现金流以及还款能力的减弱可能会出现一定数量的不良贷款甚至违约坏账。由于银行通常采用预期信用损失模型进行风险管理，疫情使得银行也需要对上半年的贷款拨备进行相应调整。同时，疫情使得很多企业和个人收入出现短期大幅度下降，对银行存贷款业务产生一定冲击，银行中间业务的拓展受阻。综合来看，疫情将对银行业短期营收会造成一定的负面影响。但长期来看，经济在疫情过后的修复与反弹也将改善受影响企业的资金状况，对银行业的长期影响有限。^[2]

基于前文划分的受影响严重程度的等级划分情况，充分考虑国家政策倾斜、银行盈利、银行长期发展等因素，银行的具体的调整策略如下所示

疫情影响程度	行业	贷款额度调整政策	信贷利率调整政策
严重影响	零售业、服务业	贷款额度上限提升 30%	贷款利率降低 30%
较大影响	工业	无变化	贷款利率降低 10%
部分影响	建造业	无变化	无变化
利好	医疗行业、科技行业	贷款额度上限提升 20%	无变化

央行等五部门联合发布了《关于进一步强化金融支持防控新型冠状病毒感染肺炎疫情的通知》。各地银行机构应加大对受疫情影响、生产经营遇到困难的中小企业融资支持力度，对受疫情影响严重的企业到期还款困难的，可通过展期或续贷、适当下调贷款利率等方式予以支持，帮助他们渡过难关。当然，银行的信贷重点也要向医疗、科技等行业倾斜，通过发放专项再贷款，向疫情防控重点保障企业提供更优惠的信贷支持。大力加强国家疫情防护、科技创新等方面的力度，让这类企业能抓住机遇更好地发展。

6.3 调整后的信贷策略规划模型

基于问题二，综合考虑信贷风险和突发因素对信贷策略的影响，将突发因素下的对中小微企业借贷政策倾斜政策量化，将增大企业贷款额度和减小借贷利率加入到信贷策

略优化模型中，放款约束条件，调整结果如下

$$\min Z = \frac{1}{2} \left[\frac{1}{\sum_{i=1}^{n_1} \sum_{j=1}^{n_0} A\lambda_j (1-l_j) r_i} + \frac{\sum_{j=1}^{n_0} (1-l_j) R_i}{\sum_{j=1}^{n_0} A\lambda_j (1-l_j)} \right] \quad (37)$$

$$\left\{ \begin{array}{l} \sum_{j=1}^{N_i} (1-l_j) R_i \leq N_i \overline{R}_i; \\ \overline{R}_i = \frac{1}{N_i} \sum_{j=1}^{N_i} R_i; \\ \sum_{j=1}^{n_0} A\lambda_j (1-l_j) \leq B; \\ r_0^1 - g_i \leq r_i \leq r_0^2 - g_i \\ l_j = 1 - e^{-a_i(r+b_i)}; \\ 0.1 \leq \lambda_j \leq 1.0 + h_j, \text{ 信贷风险 1 级;} \\ 0.1 \leq \lambda_j \leq 0.9 + h_j, \text{ 信贷风险 2 级;} \\ 0.1 \leq \lambda_j \leq 0.6 + h_j, \text{ 信贷风险 3 级;} \\ i = 1, 2, n_1; j = 1, 2, \dots, n_0. \end{array} \right. \quad (38)$$

其中 h_j 表示银行对不同企业贷款额度的提升系数， g_j 表示银行对不同企业年度利率优惠系数。

6.4 模型求解与分析

基于上述调整后的信贷策略规划模型，利用人工鱼群算法求解得到最优信贷策略。所求得的前 8 个企业的贷款额度和贷款利率如下表所示

表 8 问题三前 8 个企业信贷策略信息表

企业代号	贷款额度/万元	贷款利率
E124	20.1643	0.0611
E125	26.3990	0.0603
E126	29.6223	0.0569
E127	24.5312	0.0654
E128	24.3403	0.0345
E129	84.7231	0.0632
E130	82.3422	0.1039
E131	59.3546	0.0582

最终银行的总的贷款额度为 1 亿元, 总利息收益为 384.58 万元, 政策补贴后银行贷款策略中的银行的总的贷款额度未发生改变, 总利息收益相较于减少了 28.47 万元, 比

政策补贴前减少了 6.90%，虽然银行有少许损失，但是对于中小微企业得到帮扶效果是十分显著的，可以帮助他们顺利度过突发事件。从整个社会效益考虑，该信贷策略的调整十分必要。

7 模型评价

7.1 对优秀论文做出的改进

本文对优秀论文做出了如下改进：1. 问题一针对优秀论文 C002 的信贷风险指标量化模型进行改进，利用 Sigmoid 函数定义风险指标，并利用数据处理后计算得到的实力指标、信誉指标进行计算，模型更加简洁，避免了指标过多造成风险指标量化的不准确性。

2. 问题二针对优秀论文 C109、C002 进行模型优化，在分析贷款利率和客户流失率的关系是改用指数拟合，提高了拟合关系式的准确性；并利用改进的含有 Dropout 的 XGBoost 算法改进信誉等级预测模型，使得训练结果更加合理。

3. 问题三总结分析了 C109、C002、C227 三篇优秀论文的解题方法：对于突发因素（疫情）对中小微企业信贷的影响，可以先从宏观分析各类行业受其影响发生的变化总趋势，再从微观考虑具体时间段内各类行业的企业受到突发因素影响经济方面发生的变化，并综合考虑国家政策调整、银行的长久发展等角度建立合适的贷款调整策略模型。

7.2 模型推广

本文在模型优化过程中运用的方法与思想可以推广到其他领域，例如：1. 本文在问题一中对风险指标量化模型的改进可以运用到保险行业中，对保险产品的风险进行量化，以更好地定价和理解潜在损失。

2. 本文在问题一、问题二使用的模型优化中包括了指数拟合和 XGBoost 算法的改进，这些可以运用到医疗健康中，通过对生物统计数据拟合分析，可以更准确地诊断和预测疾病的发展趋势。

3. 本文在问题三中对突发因素的宏观与微观分析思想，可以运用到灾害管理领域，可以对自然或人为灾害的宏观和微观影响进行分析，并制定相应的应急响应计划。

参考文献

- [1] 黄鑫淼, 朱瑞, 秦一凡. 银行对中小微企业的信贷策略研究 [J]. 中国商论, 2022.
- [2] 肖虹, 裘益政, 刘巧瑜. 新冠肺炎疫情下逆周期信贷政策效用研究: 基于企业资产负债表渠道与银行风险承担渠道视角 [J]. 商业经济与管理, 2022, 42(6): 81-97.
- [3] 严远亭, 朱原玮, 吴增宝等. 构造性覆盖算法的 SMOTE 过采样方法 *[J]. 计算机科学与探索, 2020, 第 14 卷 (6): 975-984
- [4] 中华人民共和国统计局. 中华人民共和国 2020 年国民经济和社会发展统计公报 [M]. 北京: 中国统计出版社, 2020.
- [5] Huo H, Cui J, Hein S, et al. Predicting dropout for nontraditional undergraduate students: A machine learning approach[J]. Journal of College Student Retention: Research, Theory Practice, 2023, 24(4): 1054-1077.

附录 A 问题一 123 家企业风险指标量化表

企业代号	实力指标	信誉指标	规模	风险指标
E1	0.761236661	5	中	0.043498647
E2	3.651396983	5	中	2.35571E-08
E3	0.628823142	1	中	0.695554869
E4	0.107520882	1	小	0.946291292
E5	0.235843301	3	中小	0.660289742
E6	1.833068448	5	中	0.000209183
E7	0.474017734	5	中	0.170964414
E8	1.204345287	5	中	0.00483922
E9	0.567922867	5	中	0.11044131
E10	1.633380519	3	中小	0.014781006
E11	0.246923802	1	中小	0.877161601
E12	0.138131527	3	小	0.795717705
E13	1.195658693	5	中	0.00505349
E14	1.305316795	1	中	0.426543138
E15	1.328294966	5	中小	0.0026068
E16	1.318746977	5	微	0.002734093
E17	0.45859202	5	中小	0.183415366
E18	1.478623083	5	中小	0.001230194
E19	0.354141396	5	中小	0.290908863
E20	1.259134465	3	小	0.044740311
E21	1.34302186	3	小	0.034959923
E22	1.272442259	5	小	0.003445149
E23	1.599808244	3	中小	0.01633446
E24	1.377011641	5	中小	0.002043822
E25	1.024617614	1	小	0.528257843
E26	0.074001849	5	微	0.817077576
E27	0.261577952	5	中小	0.425680355
E28	1.275345429	3	微	0.042661824
E29	1.255813917	0	微	1
E30	0.143377533	3	小	0.788189337

企业代号	实力指标	信誉指标	规模	风险指标
E31	0.44633094	5	小	0.193887365
E32	1.328965117	3	中小	0.036438283
E33	0.043668731	3	小	0.934590427
E34	0.158376366	3	小	0.766805951
E35	1.209606294	3	微	0.051721954
E36	1.301959454	0	中小	1
E37	0.03827747	3	小	0.942646805
E38	0.925233489	3	小	0.117302979
E39	1.179219647	1	微	0.47038507
E40	1.326403317	1	中小	0.419509842
E41	1.198028528	1	小	0.463652232
E42	1.20296445	5	微	0.004872664
E43	0.044742307	3	微	0.932987123
E44	1.419093706	1	中小	0.389607402
E45	1.19010724	0	微	1
E46	1.31638858	1	小	0.422839534
E47	1.608977128	1	中小	0.333461349
E48	1.168731811	5	小	0.005779686
E49	0.394633404	1	中小	0.805204805
E50	1.181077329	1	微	0.46971709
E51	1.197699656	3	微	0.053552487
E52	1.175318839	0	微	1
E53	0.64920921	1	微	0.686335524
E54	0.499619477	5	中小	0.151983336
E55	0.169457978	1	微	0.915473187
E56	0.079144888	1	小	0.9604482
E57	1.145066345	3	微	0.062426669
E58	1.184191623	3	小	0.055705539
E59	0.905107271	5	小	0.021425792
E60	0.090486554	3	微	0.865097571
E61	0.202341041	3	小	0.705477125
E62	0.128673902	3	微	0.809350719
E63	0.08739801	3	小	0.869648887
E64	0.262688459	5	微	0.423822787

企业代号	实力指标	信誉指标	规模	风险指标
E65	1.085259356	3	微	0.07423963
E66	0.109382936	3	微	0.83738223
E67	1.030838187	3	微	0.08683416
E68	0.084334171	1	微	0.957857888
E69	1.212296165	1	微	0.458590046
E70	0.227445102	3	小	0.671480572
E71	0.098573651	3	小	0.853207727
E72	0.012273758	1	微	0.993863198
E73	0.14347993	1	微	0.928382855
E74	1.173095181	3	微	0.057536903
E75	0.208415924	1	中小	0.896167615
E76	0.771490724	3	微	0.179862907
E77	0.155000288	1	微	0.922654646
E78	0.113651855	1	微	0.943235161
E79	1.105271477	3	微	0.070065256
E80	0.080527291	1	微	0.959758098
E81	0.42283983	5	微	0.215448759
E82	0.566502042	0	微	1
E83	0.966636984	3	中小	0.104315931
E84	0.507339603	5	微	0.146650446
E85	1.090738127	3	微	0.073073596
E86	0.01645555	1	微	0.991772411
E87	1.095261529	0	微	1
E88	0.545420687	5	微	0.122785724
E89	0.081773765	5	微	0.798366774
E90	0.447458041	1	微	0.779930619
E91	0.56768086	5	微	0.110567636
E92	0.017855308	1	微	0.991072583
E93	0.520432717	3	微	0.346920899
E94	1.066505304	1	微	0.512136493
E95	0.203986677	3	微	0.703224436
E96	1.105342694	1	微	0.497480347
E97	0.998465366	3	微	0.095268592
E98	0.515675778	3	微	0.351032057

企业代号	实力指标	信誉指标	规模	风险指标
E99	1.161450706	0	微	1
E100	0.237027359	0	微	1
E101	0.096448246	0	微	1
E102	0.242373483	0	微	1
E103	0.018071215	0	微	1
E104	1.091154166	1	微	0.502802007
E105	0.200573341	1	微	0.900048191
E106	0.665554779	3	微	0.239107181
E107	1.103136152	0	微	1
E108	0.024314432	0	微	1
E109	1.103568257	0	微	1
E110	1.009930282	1	微	0.53398699
E111	0.523234778	0	微	1
E112	0.024160092	0	微	1
E113	1.103891495	0	微	1
E114	0.01509833	0	微	1
E115	0.000464882	0	微	1
E116	0.403093873	0	微	1
E117	1.097736698	0	微	1
E118	0.095623292	0	微	1
E119	1.123022116	0	微	1
E120	0.773636338	0	微	1
E121	0.0686558	0	微	1
E122	0.001809972	0	微	1
E123	1.104253594	0	微	1

附录 B 问题一 96 家贷款比例和贷款利率信息表

企业代号	风险等级	贷款比例	贷款利率
E1	1	0.97613969	0.068092245
E2	1	0.200416885	0.074118174
E3	2	0.360646858	0.109479401
E4	3	0.221674335	0.061352025
E5	2	0.219050892	0.086048088
E6	1	0.804676373	0.120029159
E7	1	0.635228105	0.117745476
E8	1	0.51588519	0.064192323
E9	1	0.530523647	0.123126624
E10	1	0.267710707	0.042675019
E11	3	0.191803231	0.111251481
E12	2	0.502255502	0.142111349
E13	1	0.799732668	0.076344124
E14	2	0.169136344	0.063348276
E15	1	0.496916232	0.048958691
E16	1	0.628757248	0.130508611
E17	1	0.265191829	0.046098506
E18	1	0.326547156	0.080578692
E19	1	0.711271102	0.06230438
E20	1	0.564623444	0.122606975
E21	1	0.651304877	0.101325754
E22	1	0.895374799	0.120766552
E23	1	0.530363405	0.098186607
E24	1	0.861990009	0.117979548
E25	2	0.258647886	0.126546077
E26	3	0.650842653	0.070169398
E27	2	0.813609459	0.05901563
E28	1	0.720860802	0.114855445
E30	2	0.266387273	0.145681344
E31	1	0.385169337	0.086791486
E32	1	0.11590877	0.127980586

企业代号	风险等级	贷款比例	贷款利率
E33	3	0.248950241	0.149193533
E34	2	0.244457137	0.110924554
E35	1	0.233987596	0.142422866
E37	3	0.153830917	0.111023341
E38	1	0.726848206	0.069925079
E39	2	0.372808539	0.111170231
E40	2	0.279084726	0.055006437
E41	2	0.349469225	0.106383964
E42	1	0.48767627	0.070973851
E43	3	0.544103507	0.148906947
E44	1	0.281388664	0.140367
E46	2	0.337459617	0.061066829
E47	1	0.218657457	0.095725083
E48	1	0.592932916	0.07103335
E49	3	0.224692462	0.121595208
E50	2	0.205375609	0.115457677
E51	1	0.387864548	0.079320068
E53	2	0.35941802	0.115900861
E54	1	0.624612858	0.137731424
E55	3	0.194455519	0.061606126
E56	3	0.398889195	0.14269549
E57	1	0.188033788	0.061603451
E58	1	0.213193343	0.146134311
E59	1	0.934473091	0.051250351
E60	3	0.638503621	0.042907772
E61	2	0.370322596	0.041567482
E62	3	0.597081632	0.13633318
E63	3	0.414276964	0.123795543
E64	2	0.834505886	0.064159807
E65	1	0.477734266	0.066308531
E66	3	0.420726917	0.140587657
E67	1	0.613035481	0.126695732
E68	3	0.369866902	0.085322279

企业代号	风险等级	贷款比例	贷款利率
E69	2	0.288392104	0.068902836
E70	2	0.658835983	0.124697179
E71	3	0.199782394	0.04024213
E72	3	0.119788102	0.106184027
E73	3	0.398946803	0.048882677
E74	1	0.423838838	0.148765111
E75	3	0.344754659	0.044160337
E76	1	0.649937242	0.116249517
E77	3	0.342663204	0.114065893
E78	3	0.359816546	0.12421331
E79	1	0.5699523	0.042808019
E80	3	0.262751341	0.13729335
E81	1	0.749296073	0.147539214
E83	1	0.545426336	0.131274685
E84	1	0.943384573	0.09254472
E85	1	0.534805259	0.074759441
E86	3	0.154848959	0.148763206
E88	1	0.990523392	0.049983835
E89	2	0.985787751	0.123489436
E90	2	0.165707882	0.091044735
E91	1	0.133888479	0.125904592
E92	3	0.384949078	0.101787039
E93	1	0.574456505	0.117678602
E94	2	0.303475617	0.069635344
E95	2	0.321111247	0.095820051
E96	2	0.289297523	0.074434913
E97	1	0.403557409	0.131975726
E98	1	0.266096607	0.097078456
E104	2	0.126752234	0.042765967
E105	3	0.173510908	0.126313679
E106	1	0.590866842	0.086362117
E110	2	0.339083143	0.059757013

附录 C 问题二 302 家企业信誉评级预测表

企业代号	信誉等级	违约情况
E124	A	否
E125	A	否
E126	A	否
E127	A	否
E128	A	否
E129	C	否
E130	A	否
E131	A	否
E132	A	否
E133	A	否
E134	A	否
E135	A	否
E136	A	否
E137	A	否
E138	A	否
E139	C	否
E140	A	否
E141	A	否
E142	A	否
E143	A	否
E144	A	否
E145	A	否
E146	A	否
E147	A	否
E148	A	否
E149	A	否
E150	A	否
E151	A	否
E152	A	否
E153	C	否
E154	A	否

企业代号	信誉等级	违约情况
E155	A	否
E156	C	否
E157	A	否
E158	A	否
E159	C	否
E160	A	否
E161	A	否
E162	C	否
E163	C	否
E164	A	否
E165	A	否
E166	A	否
E167	A	否
E168	C	否
E169	A	否
E170	A	否
E171	A	否
E172	A	否
E173	A	否
E174	A	否
E175	A	否
E176	A	否
E177	A	否
E178	A	否
E179	A	否
E180	A	否
E181	A	否
E182	A	否
E183	A	否
E184	A	否
E185	C	否
E186	A	否
E187	A	否
E188	A	否

企业代号	信誉等级	违约情况
E189	A	否
E190	A	否
E191	A	否
E192	B	否
E193	A	否
E194	A	否
E195	C	否
E196	C	否
E197	A	否
E198	A	否
E199	A	否
E200	A	否
E201	A	否
E202	A	否
E203	A	否
E204	A	否
E205	C	否
E206	A	否
E207	A	否
E208	A	否
E209	A	否
E210	A	否
E211	A	否
E212	C	否
E213	A	否
E214	A	否
E215	A	否
E216	A	否
E217	D	是
E218	A	否
E219	A	否
E220	A	否
E221	A	否
E222	A	否

企业代号	信誉等级	违约情况
E223	A	否
E224	A	否
E225	C	否
E226	A	否
E227	A	否
E228	A	否
E229	A	否
E230	A	否
E231	A	否
E232	A	否
E233	A	否
E234	A	否
E235	A	否
E236	A	否
E237	A	否
E238	A	否
E239	A	否
E240	A	否
E241	A	否
E242	A	否
E243	A	否
E244	C	否
E245	A	否
E246	A	否
E247	A	否
E248	A	否
E249	A	否
E250	C	否
E251	A	否
E252	A	否
E253	A	否
E254	C	否
E255	A	否
E256	C	否

企业代号	信誉等级	违约情况
E257	C	否
E258	A	否
E259	A	否
E260	A	否
E261	A	否
E262	A	否
E263	C	否
E264	A	否
E265	A	否
E266	C	否
E267	C	否
E268	A	否
E269	B	否
E270	A	否
E271	A	否
E272	A	否
E273	A	否
E274	A	否
E275	C	否
E276	A	否
E277	A	否
E278	C	否
E279	A	否
E280	A	否
E281	A	否
E282	A	否
E283	A	否
E284	A	否
E285	A	否
E286	A	否
E287	C	否
E288	A	否
E289	A	否
E290	A	否

企业代号	信誉等级	违约情况
E291	A	否
E292	A	否
E293	B	否
E294	A	否
E295	A	否
E296	A	否
E297	C	否
E298	A	否
E299	A	否
E300	A	否
E301	A	否
E302	A	否
E303	C	否
E304	A	否
E305	A	否
E306	A	否
E307	C	否
E308	A	否
E309	A	否
E310	A	否
E311	A	否
E312	A	否
E313	C	否
E314	A	否
E315	C	否
E316	A	否
E317	A	否
E318	C	否
E319	C	否
E320	A	否
E321	C	否
E322	A	否
E323	A	否
E324	A	否

企业代号	信誉等级	违约情况
E325	A	否
E326	C	否
E327	C	否
E328	A	否
E329	B	否
E330	A	否
E331	B	否
E332	A	否
E333	A	否
E334	B	否
E335	C	否
E336	C	否
E337	C	否
E338	C	否
E339	C	否
E340	C	否
E341	B	否
E342	C	否
E343	C	否
E344	B	否
E345	C	否
E346	A	否
E347	B	否
E348	C	否
E349	B	否
E350	A	否
E351	B	否
E352	A	否
E353	C	否
E354	C	否
E355	A	否
E356	C	否
E357	B	否
E358	A	否

企业代号	信誉等级	违约情况
E359	C	否
E360	C	否
E361	B	否
E362	B	否
E363	B	否
E364	C	否
E365	B	否
E366	C	否
E367	C	否
E368	C	否
E369	C	否
E370	C	否
E371	C	否
E372	C	否
E373	A	否
E374	C	否
E375	A	否
E376	C	否
E377	C	否
E378	A	否
E379	C	否
E380	C	否
E381	A	否
E382	C	否
E383	C	否
E384	B	否
E385	C	否
E386	C	否
E387	C	否
E388	C	否
E389	C	否
E390	C	否
E391	C	否

企业代号	信誉等级	违约情况
E392	C	否
E393	C	否
E394	A	否
E395	B	否
E396	B	否
E397	C	否
E398	B	否
E399	C	否
E400	C	否
E401	C	否
E402	D	是
E403	B	否
E404	D	是
E405	C	否
E406	C	否
E407	B	否
E408	C	否
E409	B	否
E410	C	否
E411	B	否
E412	C	否
E413	C	否
E414	C	否
E415	D	否
E416	D	否
E417	D	否
E418	C	否
E419	D	否
E420	D	是
E421	C	否
E422	B	是
E423	D	是
E424	D	否
E425	D	是

附录 D 问题三 69 家企业 2020 年 1 月 2 月与前两年的同比增长率表

企业代号	2018 年	2019 年	2018 年	2019 年	行业
	1 月		2 月		
E141	-18.00%	5.69%	-38.81%	-60.07%	零售业
E143	-120.30%	-207.03%	-501.24%	-10.51%	科技行业
E162	-366.37%	-233.21%	-135.57%	-78.72%	科技行业
E163	-354.67%	-21.80%	-101.52%	-99.48%	服务业
E171	435.05%	7.92%	-99.87%	-99.06%	建造行业
E175	-95.53%	-95.33%	-89.04%	-94.47%	零售业
E176	-48.28%	-64.16%	57.63%	236.83%	科技行业
E180	-95.06%	11.45%	1101.53%	-1647.50%	建造行业
E194	-330.11%	-42.16%	-74.97%	-118.99%	服务业
E196	59.37%	138.85%	-95.78%	-87.60%	零售业
E212	118.92%	39.14%	-83.26%	-17.92%	服务业
E216	9.40%	-61.87%	-96.36%	-94.87%	建造行业
E225	21.65%	22.79%	-91.54%	-193.93%	服务业
E232	-470.00%	-1123.91%	-96.27%	-95.64%	服务业
E243	842.74%	23.72%	-483.54%	-55.37%	科技行业
E247	-1993.08%	-143.73%	-101.10%	-110.49%	建造行业
E250	45.69%	-22.00%	-51.74%	-70.62%	科技行业
E258	-75.21%	-61.07%	-64.68%	131.87%	服务业
E265	-24.23%	-70.87%	-203.46%	-1856.51%	科技行业
E269	5.13%	-97.78%	-99.39%	-80.51%	服务业
E275	-23.41%	-48.35%	-543.43%	-4283.92%	科技行业
E276	-46.54%	-67.40%	-97.97%	-102.04%	零售业
E278	-62.23%	37960.33%	-100.71%	-99.86%	零售业
E279	-147.25%	-108.52%	0.58%	138.06%	服务业
E282	-139.05%	-162.43%	-197.86%	-282.56%	服务业
E284	-13.58%	-66.55%	-102.63%	-100.94%	科技行业
E289	-65.89%	-76.25%	-100.44%	-100.41%	服务业
E290	-67.12%	-97.41%	-100.71%	-103.26%	科技行业
E291	134.42%	214.96%	-99.91%	-99.58%	建造行业
E298	142.87%	-4193.40%	3884.00%	-268.50%	服务业

企业代号	2018 年	2019 年	2018 年	2019 年	行业
	1 月		2 月		
E299	-1968.25%	107.67%	7261.86%	2134.98%	零售业
E301	-473.83%	-250.07%	-102.53%	-100.89%	服务业
E302	-201.93%	-64.62%	-93.96%	-100.41%	科技行业
E303	-4607.10%	52.52%	-12744.29%	-77.81%	零售业
E305	-87.30%	2519.00%	-98.26%	-97.95%	服务业
E307	-14899.19%	-74.14%	-52.99%	-100.09%	科技行业
E310	77.73%	-2518.52%	2.57%	-35.17%	服务业
E311	26.48%	53.78%	-102.24%	-114.56%	服务业
E313	-39.70%	-21.98%	207.23%	-524.90%	服务业
E314	231.72%	-92.98%	-100.68%	-105.31%	服务业
E315	-45.70%	-64.04%	-101.78%	-100.83%	零售业
E318	-50.49%	-51.82%	-50.08%	-46.74%	零售业
E320	-19.13%	-101.34%	-89.50%	-100.48%	建造行业
E325	169.70%	-51.97%	-102.68%	-101.82%	零售业
E326	89.15%	-1959.65%	-148.84%	-136.71%	零售业
E330	28.33%	-18.99%	-101.94%	-104.09%	服务业
E333	-88.44%	533.04%	-101.30%	-93.82%	服务业
E334	-134.31%	133.94%	-100.44%	-38.74%	科技行业
E335	832.29%	462.50%	-136.29%	824.54%	零售业
E342	-5943.84%	-19.56%	-1337.98%	16.98%	科技行业
E344	-180.68%	-105.86%	-10352.33%	-124.78%	科技行业
E348	-269.41%	-172.39%	-79.91%	-65.76%	服务业
E350	-36.40%	-980.92%	-67.70%	-109.73%	零售业
E352	268.55%	58.96%	-90.80%	-85.94%	服务业
E353	-108.97%	19.67%	-98.44%	-97.77%	服务业
E364	-87.84%	-75.99%	-88.34%	-107.62%	科技行业
E370	-26.36%	19.13%	866.50%	904.24%	服务业
E372	44.66%	9.56%	-101.22%	-124.89%	服务业
E377	-121.81%	-104.51%	-185.71%	-77.27%	服务业
E378	-4260.03%	1669.70%	-752.46%	-8.64%	零售业
E379	92.60%	-52.58%	-100.83%	-69.73%	医疗行业
E391	-78.34%	-80.14%	208.03%	692.61%	零售业

企业代号	2018 年	2019 年	2018 年	2019 年	行业
	1 月		2 月		
E392	588.90%	825.88%	-191.87%	-122.39%	科技行业
E403	-92.46%	-27.21%	10.83%	40.14%	科技行业
E407	-36752.39%	-18804.26%	-98.32%	-100.75%	服务业
E410	-61.39%	-59.92%	-106.47%	-89.10%	科技行业
E417	-40.00%	-50.93%	-81.36%	-74.86%	建造行业
E419	1727.53%	28.17%	1603.33%	169.79%	科技行业
E420	-99.91%	-99.90%	-34.60%	67.07%	医疗行业

附录 E 问题一源代码

```

import numpy as np
import pandas as pd

# Parameters
N_fish = 100 # Number of fish
N_iter = 1000 # Number of iterations
N_dim = 192 # Dimensionality of the problem (96 loan ratios + 96 loan interest rates)
visual = 0.2 # Visual distance
step_individual = 0.01 # Individual step
step_follow = 0.01 # Follow step
step_swarm = 0.01 # Swarm step
try_number = 5 # Number of tries in observation behavior

# Bounds for the variables
bounds = [(0.1, 1)] * 96 + [(0.04, 0.15)] * 96

# Objective function
def objective_function(x_r, credit_ratings):
    loan_ratios = x_r[:96]
    loan_interest_rates = x_r[96:]
    # Define the objective function based on the given formulas
    # ...

# Constraints
def combined_constraints(x_r):
    loan_ratios = x_r[:96]
    loan_interest_rates = x_r[96:]
    constraints = []

```

```

# Define the constraints based on the given formulas
# ...
return constraints

# Check if all constraints are satisfied
def constraints_satisfied(x_r):
    return all(value >= 0 for value in combined_constraints(x_r))

# Observation behavior
def observe(x_r, credit_ratings):
    # ...

# Follow behavior
def follow(x_r, swarm, credit_ratings):
    # ...

# Swarm behavior
def swarm_behavior(x_r, swarm, credit_ratings):
    # ...

# Main loop
swarm = np.random.uniform(low=0.1, high=1, size=(N_fish, 96)) # Initialize the swarm with
    random values within bounds
swarm = np.hstack((swarm, np.random.uniform(low=0.04, high=0.15, size=(N_fish, 96))))
best_x_r_global = None
best_f_global = float('inf')
best_f_values = []
for iteration in range(N_iter):
    # Apply behaviors
    new_swarm = []
    for x_r in swarm:
        new_x_r = observe(x_r, credit_ratings)
        new_x_r = follow(new_x_r, swarm, credit_ratings)
        new_x_r = swarm_behavior(new_x_r, swarm, credit_ratings)
        new_swarm.append(new_x_r)
    # Update the best solution
    # ...
    swarm = np.array(new_swarm)
    best_f_values.append(best_f_global)

# Results
print("\nFinal Best Objective Function Value:", best_f_global)
print("Final Best Solution:", best_x_r_global)

# Saving to Excel
result_df = pd.DataFrame({
    "企业代号": df["企业代号"],

```

```

    "贷款比例": best_loan_ratios,
    "贷款利率": best_loan_interest_rates
})
result_file_path = "/path/to/your/directory/best_loan_solution.xlsx"
result_df.to_excel(result_file_path, index=False)

```

附录 F 问题二 XGBoost 算法源代码

```

##数据加载和预处理
# 读取Excel文件的第一个工作表 (Sheet1)
sheet1_path = '/path/to/sheet1.xlsx'
sheet1_data = pd.read_excel(sheet1_path)

# 添加新特征'产品退货率.1'
sheet1_data['产品退货率.1'] = sheet2_data['产品退货率.1']

# 分离特征和目标变量
X = sheet1_data.drop(columns=['企业代号', '信用评级', '是否有违约'])
y_rating = sheet1_data['信用评级']
y_default = sheet1_data['是否有违约']

##信用评级模型训练
# 信用评级模型参数
params_rating = {
    'objective': 'multi:softprob',
    'max_depth': 5,
    'num_class': 4,
    'eval_metric': 'mlogloss',
    'seed': 42
}

# 训练信用评级模型
dtrain_rating = xgb.DMatrix(X_train_resampled_rating, label=y_train_resampled_rating)
dval_rating = xgb.DMatrix(X_val_resampled_rating, label=y_val_resampled_rating)
model_rating = xgb.train(params_rating, dtrain_rating, num_boost_round=200,
    evals=[(dtrain_rating, 'train'), (dval_rating, 'val')], early_stopping_rounds=10)

##违约情况模型训练
# 违约情况模型参数
params_default = {
    'objective': 'binary:logistic',
    'max_depth': 5,
    'eval_metric': 'logloss',
    'seed': 42
}

```

```

# 训练违约情况模型
dtrain_default = xgb.DMatrix(X_train_default, label=y_train_default)
dval_default = xgb.DMatrix(X_val_default, label=y_val_default)
model_default = xgb.train(params_default, dtrain_default, num_boost_round=200,
    evals=[(dtrain_default, 'train'), (dval_default, 'val')], early_stopping_rounds=10)

##特征调整和预测

# 加载Sheet2数据
sheet2_data = pd.read_excel(sheet2_path)

# 调整特征
features_to_divide = ['上游业务量年均增长率', '下游业务量年均增长率', '年均毛利润增长率']
X_sheet2[features_to_divide] = X_sheet2[features_to_divide] / 100

# 使用信誉评级和违约情况模型进行预测
predicted_rating = model_resampled_rating.predict(xgb.DMatrix(X_sheet2)).argmax(axis=1)
predicted_default = (model_default.predict(xgb.DMatrix(X_sheet2)) > 0.5).astype(int)

# 将预测结果保存到Excel文件
sheet2_data['预测信誉评级'] = predicted_rating
sheet2_data['预测是否有违约'] = predicted_default
sheet2_data.to_excel("/path/to/output.xlsx", index=False)

```

附录 G 问题二人工鱼群算法源代码

```

#Non-linear Optimization Code
#1. Import Necessary Libraries
import numpy as np
from scipy.optimize import minimize
import pandas as pd
#2. Read Data

data = pd.read_excel("人鱼2.xlsx")
Ri = data["风险等级"].values
# ... other variables ...

#3. Define Objective Function and Constraints

# 计算银行利润
def calculate_bank_profit(loan_proportions, interest_rates, Ri, A=10000):
    return np.sum(A * loan_proportions * (1 - (1 - np.exp(-22.0393 * (interest_rates - 0.04))))
        * interest_rates)

```



```

# 计算单位贷款风险
def calculate_unit_loan_risk(loan_proportions, interest_rates, Ri, A=10000):
    risk_factors = (1 - (1 - np.exp(-22.0393 * (interest_rates - 0.04)))) * Ri
    return np.sum(risk_factors) / np.sum(A * loan_proportions * (1 - (1 - np.exp(-22.0393 *
        (interest_rates - 0.04)))))

# 目标函数
def objective(x):
    loan_proportions = x[:290]
    interest_rates = x[290:]
    bank_profit = calculate_bank_profit(loan_proportions, interest_rates, Ri)
    unit_loan_risk = calculate_unit_loan_risk(loan_proportions, interest_rates, Ri)
    return -(bank_profit - unit_loan_risk)

# 约束函数
def constraint1(x):
    return 10000 - np.sum(x[:290])

# ... other constraints ...

#4. Define Optimization Problem

x0 = np.concatenate([np.full(290, 0.5), np.full(290, 0.075)]) # 初始解
bounds = [(0.1, 1)] * 290 + [(0.04, 0.15)] * 290 # 变量界限
con1 = {'type': 'ineq', 'fun': constraint1}
# ... other constraints ...
cons = [con1] # 约束列表

# 使用minimize函数求解
solution = minimize(objective, x0, method='SLSQP', bounds=bounds, constraints=cons)

#5. Extract Results

best_loan_proportions = solution.x[:290]
best_interest_rates = solution.x[290:]
final_bank_profit = calculate_bank_profit(best_loan_proportions, best_interest_rates, Ri)
final_unit_loan_risk = calculate_unit_loan_risk(best_loan_proportions, best_interest_rates, Ri)

```

附录 H 问题三同比利润率计算代码

```

# 导入所需的库
import pandas as pd

# 读取Excel文件中的三个表格

```

```

file_path = "/mnt/data/附件2: 302家无信贷记录企业的相关数据.xlsx"
sheet1 = pd.read_excel(file_path, sheet_name=0)
sheet2 = pd.read_excel(file_path, sheet_name=1)
sheet3 = pd.read_excel(file_path, sheet_name=2)

# 从302家企业中随机抽取60家
selected_companies = sheet1.sample(n=60, random_state=42)

# 定义一个函数来计算所选企业在特定月份的利润
def calculate_profit(selected_companies, sales_invoices, purchase_invoices, year, months):
    valid_sales_invoices = sales_invoices[sales_invoices['发票状态'] == '有效发票']
    valid_purchase_invoices = purchase_invoices[purchase_invoices['发票状态'] == '有效发票']
    valid_sales_invoices =
        valid_sales_invoices[valid_sales_invoices['企业代号'].isin(selected_companies['企业代号'])]
    valid_purchase_invoices =
        valid_purchase_invoices[valid_purchase_invoices['企业代号'].isin(selected_companies['企业代号'])]
    profits = []
    for month in months:
        monthly_sales = valid_sales_invoices[(valid_sales_invoices['开票日期'].dt.year == year)
            & (valid_sales_invoices['开票日期'].dt.month == month)]
        monthly_purchases = valid_purchase_invoices[(valid_purchase_invoices['开票日期'].dt.year
            == year) & (valid_purchase_invoices['开票日期'].dt.month == month)]
        monthly_profit = monthly_sales['价税合计'].sum() - monthly_purchases['价税合计'].sum()
        profits.append(monthly_profit)
    return profits

# 计算2020年1月和2月的利润
profits_2020 = calculate_profit(selected_companies, sheet2, sheet3, year=2020, months=[1, 2])

# 计算2018年1月和2月的利润
profits_2018 = calculate_profit(selected_companies, sheet2, sheet3, year=2018, months=[1, 2])

# 计算2019年1月和2月的利润
profits_2019 = calculate_profit(selected_companies, sheet2, sheet3, year=2019, months=[1, 2])

# 计算同比增长率
growth_rate_2019 = [(profits_2019[i] - profits_2018[i]) / profits_2018[i] * 100 for i in
    range(2)]
growth_rate_2020 = [(profits_2020[i] - profits_2019[i]) / profits_2019[i] * 100 for i in
    range(2)]

# 创建一个DataFrame来保存所有302家企业的计算结果
all_growth_rate_df = pd.DataFrame(columns=['企业代号', '企业名称', '2018年1月利润',
    '2019年1月利润', '2020年1月利润', '1月同比增长率2019', '1月同比增长率2020',
    '2018年2月利润', '2019年2月利润', '2020年2月利润',
    '2月同比增长率2019', '2月同比增长率2020'])

```

```
# 分别计算每个企业的同比增长率
for index, company in sheet1.iterrows():
    # ...
    # 计算利润和同比增长率
    # ...

    all_growth_rate_df.loc[index] = [company_code, company_name, profits_2018[0],
                                     profits_2019[0], profits_2020[0], growth_rate_2019[0], growth_rate_2020[0],
                                     profits_2018[1], profits_2019[1], profits_2020[1],
                                     growth_rate_2019[1], growth_rate_2020[1]]

# 保存为Excel文件
all_output_path = 'all_growth_rate_analysis.xlsx'
all_growth_rate_df.to_excel(all_output_path, index=False)
```