

中图分类号：F832

论文编号：10006ZF2208417

北京航空航天大学  
硕士学位论文

基于机器学习的 A 公司企业  
客户信用评价研究

作者姓名 刘玉芳

学科专业 工商管理

指导教师 张军欢 副教授

培养学院 经济管理学院

# **Research on Customer Credit Evaluation of Company A Based on Machine Learning**

A Dissertation Submitted for the Degree of Master

**Candidate : Liu Yufang**

**Supervisor : Prof. Zhang Junhuan**

School of Economics and Management

Beihang University, Beijing, China

中图分类号：F832

论文编号：10006ZF2208417

## 硕 士 学 位 论 文

# 基于机器学习的 A 公司企业客户信用评价研究

作者姓名	刘玉芳	申请学位级别	工商管理硕士
指导教师姓名	张军欢	职 称	副教授
学科专业	工商管理	研究方向	金融科技
学习时间自	2022 年 09 月 01 日	起至	2024 年 06 月 30 日止
论文提交日期	2024 年 04 月 30 日	论文答辩日期	2024 年 05 月 23 日
学位授予单位	北京航空航天大学	学位授予日期	年 月 日

## 关于学位论文的独创性声明

本人郑重声明：所呈交的论文是本人在指导教师指导下独立进行研究工作所取得的成果，论文中有关资料和数据是实事求是的。尽我所知，除文中已经加以标注和致谢外，本论文不包含其他人已经发表或撰写的研究成果，也不包含本人或他人为获得北京航空航天大学或其它教育机构的学位或学历证书而使用过的材料。与我一同工作的同志对研究所做的任何贡献均已在论文中作出了明确的说明。

若有不实之处，本人愿意承担相关法律责任。

学位论文作者签名：\_\_\_\_\_ 日期：\_\_\_\_\_ 年 \_\_\_\_\_ 月 \_\_\_\_\_ 日

## 学位论文使用授权

本人完全同意北京航空航天大学有权使用本学位论文（包括但不限于其印刷版和电子版），使用方式包括但不限于：保留学位论文，按规定向国家有关部门（机构）送交学位论文，以学术交流为目的赠送和交换学位论文，允许学位论文被查阅、借阅和复印，将学位论文的全部或部分内容编入有关数据库进行检索，采用影印、缩印或其他复制手段保存学位论文。

保密学位论文在解密后的使用授权同上。

学位论文作者签名：\_\_\_\_\_ 日期：\_\_\_\_\_ 年 \_\_\_\_\_ 月 \_\_\_\_\_ 日

指导教师签名：\_\_\_\_\_ 日期：\_\_\_\_\_ 年 \_\_\_\_\_ 月 \_\_\_\_\_ 日

## 摘 要

信用评级是评估企业信用风险的核心工具，为商业活动和金融决策提供了不可或缺的参考依据。为了更准确地评价客户信用等级，企业需要识别出影响客户信用的相关因素。这些因素不仅繁多且复杂，需要运用科学的方法和手段进行筛选和识别。A 公司作为一个拥有全球客户的企业，更需要深入挖掘影响客户信用的关键因素。通过综合应用专业的技术手段和科学方法，A 公司可以更准确、客观地评估客户的信用状况，从而制定更为合理的还款账期和政策，降低逾期还款风险，保护企业利益。

本文以 A 公司的企业客户为研究对象，筛选了在财务、业务、企业资质方面的 23 个相关指标，使用二元 Logistic 回归进行单因素、多因素的分析，最终筛选出经营负荷情况、现金比率、应收账款周转率、是否有逾期记录、是否上市以及销售净利率等 14 个影响客户未来是否逾期的关键信用指标。并使用这 14 个指标分别基于卷积神经网络和 BP 神经网络构建了预测客户未来是否会逾期以及客户信用等级的一套评价模型，并进行了模型训练及参数调优。经过验证，基于卷积神经网络的模型准确度更高，更能能够客观反映客户的信用状况，同时研究发现以上 14 个指标能很好的预测客户未来是否会逾期, 而且在企业信用等级的评估中也发挥着重要作用。

**关键词：**信用评价指标，二元 Logistic 回归，卷积神经网络，BP 神经网络

## Abstract

Credit rating is a core tool for assessing corporate credit risk, providing an indispensable reference for commercial activities and financial decision-making. In order to more accurately evaluate customer credit ratings, enterprises need to identify relevant factors that affect customer credit. These factors are not only numerous but also complex, requiring the use of scientific methods and means for screening and identification. As a company with global customers, Company A needs to delve deeper into the key factors that affect customer credit. By comprehensively applying professional technical means and scientific methods, Company A can more accurately and objectively evaluate the credit status of customers, formulate more reasonable repayment terms and policies, reduce the risk of overdue repayment, and protect the interests of the enterprise.

This paper takes the corporate clients of Company A as the research object, selects 23 relevant indicators in finance, business, and corporate qualifications, uses binary logistic regression for single factor and multi factor analysis, and finally selects 6 key credit indicators that affect whether clients will be overdue in the future, including operating load, cash ratio, accounts receivable turnover rate, whether there is overdue record, whether they are listed, and sales net profit margin. Use these 6 indicators to train a set of evaluation models based on BP neural network to predict whether customers will be overdue in the future and their credit rating. After verification, the model can objectively reflect the credit status of customers and provide strong decision support for Company A. Research has found that the above six indicators can effectively predict whether customers will become overdue in the future, and also play an important role in evaluating corporate credit ratings.

**Key words:** Credit evaluation indicators, Binary Logistic Regression, Convolutional neural network, BP Neural Network

# 目 录

第一章 绪论 .....	1
1.1 选题背景和意义 .....	1
1.1.1 选题的背景 .....	1
1.1.2 选题的意义 .....	1
1.2 国内外研究现状 .....	2
1.2.1 国内研究现状 .....	2
1.2.2 国外研究现状 .....	5
1.2.3 文献综述总结 .....	6
1.3 研究问题和研究内容 .....	7
1.3.1 研究问题 .....	7
1.3.2 研究内容 .....	8
1.3.3 技术路线 .....	8
1.4 创新点 .....	8
1.5 章节结构 .....	9
第二章 相关基本理论 .....	11
2.1 信用评价的相关理论 .....	11
2.1.1 信用评价的目的 .....	11
2.1.2 信用评价的分类 .....	11
2.2 二元 Logistic 逻辑回归 .....	12
2.2.1 sigmoid 函数 .....	12
2.2.2 回归模型估计 .....	13
2.3 BP 神经网络 .....	15
2.3.1 BP 神经网络信号传播 .....	15
2.3.2 训练过程 .....	16
2.3.3 激活函数 .....	17
2.3.4 梯度下降法 .....	19
2.3.5 模型检验 .....	20
2.4 本章小结 .....	21

第三章 信用评价模型构建 .....	22
3.1 基于二元 Logistic 回归构建指标体系 .....	22
3.1.1 数据选取 .....	22
3.1.2 指标筛选 .....	22
3.1.3 数据预处理 .....	24
3.1.4 指标数据的处理 .....	27
3.1.5 单因素 Logistic 分析 .....	28
3.1.6 多因素 Logistic 分析 .....	30
3.1.7 指标体系合理性检验 .....	31
3.2 基于 BP 神经网络构建评价模型 .....	32
3.2.1 模型参数说明 .....	32
3.2.2 模型建立 .....	33
3.3 本章小结 .....	37
第四章 实证研究 .....	38
4.1 数据筛选与描述 .....	38
4.2 数据预处理 .....	39
4.3 模型训练 .....	40
4.4 模型检验 .....	40
4.4.1 模型准确率分析 .....	41
4.4.2 模型性能分析 .....	42
4.4.3 模型 R 值分析 .....	42
4.4.4 验证集结果分析 .....	43
4.4.5 实证分析结论 .....	44
4.5 本章小结 .....	45
结    论 .....	46
参考文献 .....	48
致    谢 .....	52



## 图 清 单

图 1 技术路线图 .....	9
图 2 Sigmoid 函数图像 .....	13
图 3 BP 神经网络信号传播图 .....	16
图 4 损失函数图 .....	19
图 5 BP 神经网络构建过程图 .....	34
图 6 BP 神经网络训练流程图 .....	35
图 7 迭代轮次性能图 .....	41
图 8 各组数据 R 值图 .....	42
图 9 验证结果结果对比图 .....	43

## 表 清 单

表 1 信用评级财务指标 .....	23
表 2 信用评级业务指标 .....	23
表 3 信用评级企业资质指标 .....	24
表 4 指标相关系数 $r$ .....	25
表 5 信用评级指标 .....	26
表 6 指标多重共线性分析 .....	27
表 7 企业成立年限的标准化处理 .....	28
表 8 客户逾期单因素回归结果 .....	29
表 9 客户逾期多因素回归结果 .....	30
表 10 霍斯默-莱梅肖检验 .....	32
表 11 分类表 .....	32
表 12 BP 神经网络输入输出参数 .....	33
表 13 基于机器学习的 A 公司企业客户信用评价体系 .....	38
表 14 信用等级划分及变量说明 .....	39
表 15 训练数据准确率 .....	41
表 16 信用评价模型验证数据 .....	44

# 第一章 绪论

## 1.1 选题背景和意义

### 1.1.1 选题的背景

随着全球性自由市场经济的迅速发展，企业的销售方式发生巨大的变化，由过去单一的计划客户转变为多元的信用销售为主。当企业产成品存货较多时，一般都采用较为优惠的信用条件进行销售，把存货转化为应收账款，减少产成品存货，节约相关的开支。然而，信用销售起到促进销售的同时，也会给企业在偿债能力以及应收账款的管理和现金流方面带来很大的挑战，稍有管理不善就容易因客户不能还款而导致坏账发生。因此，加强对客户的客户信用管理显得越来越重要。

信用评级是用于衡量企业信用风险的工具，是作为企业商业活动或是在金融市场上的重要参考指标。通过评估企业的等级，可以了解该企业的信用质量和偿还能力。客户信用评级是企业为了有效控制客户信用风险，实现信贷资金的安全性、流动性和收益性。

A 公司目前上千家客户，覆盖全球众多国家及地区，客户信用质量关系到企业经营成败，所以需要密切关注客户的信用状况，并采取相应的客户信用等级评价措施来实时掌握客户的信用状况，制定相应的还款账期，来降低逾期还款风险、保护企业的利益。

### 1.1.2 选题的意义

客户信用评级下降，意味着客户的偿付能力可能会降低，企业可能面临逾期付款或坏账的风险，从而对企业的财务状况产生不利影响。通过合理的信用风险评级，不仅可以得到客户信用等级的排序，为公司进行业务来往提供决策参考；还能确定每个信用等级贷款客户的违约损失率大小，为公司对每个客户的贷款额度的设定提供依据。对与公司来说，客户信用评价的意义，具体可以表现为以下几个方面：

- 1、风险控制。客户信用评价可以帮助企业识别潜在的信用风险客户，包括那些可能逾期付款或违约的客户，从而帮助企业提前采取风险控制措施，避免与高风险客户进行交易，从而降低逾期付款和坏账损失的发生。同时能帮助企业制定相应的风险管理策略，包括建立合理的信用额度、采取灵活的付款方式等，避免过度放宽信用条件，从而

提高资金利用效率。

2、收益提升。客户信用评价可以帮助企业识别高质量的客户，从而优化客户选择，帮助企业建立更加稳定的合作关系；通过信用评价也可以帮助企业了解客户的信用状况和偏好，从而优化销售策略，制定更加精准的营销计划，提高销售收入和市场份额。

3、决策支持。客户信用评价可以帮助企业根据客户的信用状况和历史付款记录，制定合理的信用额度和付款条件。

4、管理优化：企业可以建立完善的客户信用评价体系，引入自动化的客户信用评价流程，提高评价效率，减少人为因素的干扰，确保评价结果的客观性和一致性。

通过以上方式，企业可以实现对客户信用评价的管理优化，提高评价的准确性和效率，降低信用风险，同时更好地支持企业的决策和运营，提高客户满意度和忠诚度，从而增强企业的竞争力。

## 1.2 国内外研究现状

### 1.2.1 国内研究现状

在信用评价指标体系选择研究方面，对企业或者客户信用评价的研究上，一般是基于财务或者非财务的指。霍思彤等 [1]根据合作企业特点，从财务指标和非财务指标两个维度设计 7 个一级指标和 25 个二级指标，采用专家打分和层次分析法确定每个指标的权重，建立起 6 级信用等级体系。宋永明等 [6]基数据的可获得性、低相关性以及全面性等考虑，借鉴已有研究选取了基于企业的偿债能力、经营能力、盈利能力、发展能力四个方面共 13 项指标构成中小企业信用评价指标体系。徐德顺和马凡慧等（2021）基于复杂系统理论和三重积分原理，从主观履约意愿、客观履约能力及外部履约环境三个维度设计指标体系，构建 WCE 三维信用评价模型，将企业信用评价拓展至三维空间内。李晓寒（2021）基于 Kisgen 的研究方法，利用公司内部财务信息建立 CR-CS 模型（信用评级—资本结构模型），对中国信用评级是否会影响非金融上市公司的资本结构展开研究。

罗勇和陈治亚（2015）基于模糊综合法构建了供应链金融客户信用评价指标体系。通过对供应链金融进行信用风险分析，确定了信用评价指标体系和评价标准。采用实例验证了评价指标体系的计算结果。该指标体系由客户企业信用、供应链信用和客户企

业所在地信用状况等 3 个一级指标构成, 其中客户信用状况包括 4 个二级指标, 供应链信用状况包括等 4 个二级指标, 地区信用状况包括 2 个二级指标, 共计 10 个二级指标; 采用专家咨询法结合层次分析法确定了指标权重和评价标准。

在信用评价模型构建方面, 模型是多样化的, 不同的模型针对特定的评估需求和应用场景进行设计, 采用独特的评估方法和技术。有基于统计学的信用评价模型、基于机器学习的信用评价模型, 还有一些针对特定行业的信用评价模型。周颖和张志鹏(2023)提出了新的信用等级划分约束条件, 确保违约企业更多分布在较低等级, 改变了现有研究将违约企业更多划分在较高等级的不合理现象。通过巴塞尔协议关于预期损失与违约风险暴露、违约概率、违约损失率的关系, 近似得出上市公司的预期损失, 给出了解决上市公司缺乏违约损失数据的方法。姚定俊和顾越等(2023)从中小微企业的基本情况出发, 建立有针对性的信用风险评价指标体系; 从财务风险的视角选取财务指标, 并创新性地将中小微企业管理层特征纳入信用风险指标体系。选用 SVM 作为违约风险估计的基础模型, 以 RF 算法筛选指标体系来提升 SVM 模型的分类能力。在参数寻优上, 初步考虑使用 SMA 算法, 并对 SMA 进行改进, 最终形成 RF-LSMA-SVM 模型进行中小微企业信用风险评价, 进一步提升了信用风险判定的准确性和实用性。华泰证券固定收益部课题组(2023)将大数据、人工智能等技术引入信用评价与投资管理领域, 通过多源数据融合方案和数据自动化校验, 提高数据质量。在获取数据的基础上, 构建智能信用评价模型动态评估发债主体和交易对手的信用情况, 提升金融机构的风险管控能力。信用评价模型开发需经过建模准备、单因素分析、多因素分析、回归模型建立和验证五个步骤。

阳彩霞(2022)采用 BP 神经网络的数据挖掘算法建立一个供应商信用评价模型, 在测试数据集基础上采用混淆矩阵对模型进行检验, 该模型的准确度高达百分之 96.7。刘英杰和李光等(2022)针对当前工程监理市场存在的企业失信、恶意竞争、违规投标等信用缺失现象, 结合工程监理行业的发展, 构建了一套有效合理的工程监理企业信用评价指标体系, 并结合实例分析, 采用 AHP 和改进灰色关联模型的方法对其信用指标进行赋权, 进而确定工程监理企业的信用评价等级。曹轲(2022) IGWOBPDA 算法通过结合改进的灰狼算法和 BP 神经网络, 针对无线传感器网络数据融合中的关键问题提出了有效的解决方案, 并通过仿真实验验证了其性能优势。任向英(2022)针对中小企业经营中的信用风险, 提出灰色聚类评估方法, 重点分析该评估法的原理与应用方式,

并结合实证分析，评估中小企业的信用情况。

杨莲和石宝峰等（2022）利用 ClassBalancedLoss 中的平衡因子  $\omega$  弥补交叉熵函数无法调节两类样本损失权重的缺陷，克服由样本不均衡带来的评价模型对非违约样本识别过度、对违约样本识别不足。并通过考虑数据重叠，利用随机覆盖方法进行不放回采样，将图像识别领域中的 ClassBalancedLoss 函数引入信用评价领域。董秉坤和郑陈柔雨等（2022）研究基于 AHP-熵值组合赋权法的零售户信用评价模型。该模型采用主观赋权法（AHP）和客观赋权法（熵权法）相结合的组合赋权方法，以弥补单一赋权带来的不足，力求将主观随机性控制在一定范围内，确保主客观赋权中的中正，实现了零售户信用的定性和定量评估的有效结合。王灿（2022）运用 EasyEnsemble 方法解决信用数据集不平衡问题，再通过非对称误差成本的核 SVM、逻辑斯蒂回归、带有距离加权的 KNN 算法以及 C5.0 算法的决策树的 Bagging 集成得到个人组合评价模型。郭鑫（2022），提出层次分析法的科研信用评价模型。采用空间分布式信息重构方法进行科研信用评价的模糊特征重组，进行科研信用多元评价的信息挖掘和特征提取与进行科研信用评价的均衡配置和线性规划设计，实现科研信用评价的层次化分析和特征重建。采用二乘规划模型进行科研信用评价的大数据调度和模糊空间信息采样，结合资源优化调度方法，实现科研信用资源信息优化配置，通过博弈均衡控制方法，进行科研信用的优化评价和参数自适应评估。刘翠玲和胡聪等（2022）构建了一种基于 XGB 算法的客户多维信用评价模型，并基于特征重要度方法进行特征选择，采用极限梯度提升方法以及树模型对客户信用进行模型构建，计算不同节点上不同的增益值来获取最佳的预测效果，从而构建一个准确、稳定的客户信用评价模型。

程砚秋和徐占东（2019），借鉴 ELECTREIII 的评价原理，计算新增贷款客户的信用风险评价得分。以违约、非违约样本加权后的组内差异程度为基础，确定 ELECTREIII 的否决阈值；反映了不同评价指标数据差异大小对评价结果的影响，避免了现有阈值人为主观确定的不足。郑晓杰和王双（2019）采用德尔菲法、层次分析法、指数分析法、模糊综合分析相结合的方式，针对 M 银行对中小型企业客户的信用评价进行模型构建，并建立起一套科学、可行的中小企业信用评价体系。赵志冲和迟国泰（2017）通过对比只含有某一个指标的完整模型的对数似然值和只含常数项的零模型的对数似然值确定统计量，构造了某一特定指标与违约状态之间的逻辑回归方程。比较在有某指标时的两个对数似然值之间的差。最终建立了包括超速动比率、近三年企业授信情况、城市居民

人均可支配收入等 16 个指标的小企业信用风险评价指标。温小霓和韩鑫蕊（2017）针对科技型中小企业，首先建立了一个多层次的信用风险评价指标体系，并利用 t 检验和因子分析对指标体系进行简化和降维，然后在此基础上构建基于 MLP 神经网络的科技型中小企业信用风险评价模型。

### 1.2.2 国外研究现状

在信用评价指标体系选择研究方面，基于信用得分或违约概率的信用等级划分。Roy P.K.（2023）研究提出了一个易于实施的 ESG 信用评级模型，它考虑了 ESG 和财务表现来填补现有的研究空白，提高了整体经济能力，同时最小化环境、社会和治理风险。Duan JC 和 Li 基于上市公司违约概率 (PD) 将上市公司划分为 AAA、AA+ 等 21 个等级。Serano-Cinca C. 和 Cutierrez-Nieto B 基于企业的 FICO 分数将企业划分为 7 个信用等级。ChenY.S 和 Cheng C. H. 基于各等级违约概率 (PD) 的差异，将企业划分为 5 个信用等级。Moon T.Kim Y. 等根据企业的信用得分，划为 10 个信用等级。Mileris R(2012) 通过因子分析与 Probit 模型进行指标筛选，确定了包括工业生产指数等影响信用风险的关键指标。Dainelli,et al.(2013) 通过 logit 模型建立了包括盈利能力、偿付能力和信贷质量等指标的中小企业信用评级体系。

在信用评价模型构建方面，Baser 等 (2023) 提出一种基于聚类的模糊分类 (CBFC) 的信用风险评估方法，将模糊聚类与监督学习相结合的策略来构建信用评分模型。Roy Pranith Kumar(2022) 早期的可持续信用评分模型被扩展为一个基于 ESG 的信用评级模型，以对借款人进行分类。Hamido Fujita 等（2022）将企业信用状况分为低违约风险、中等违约风险和高违约风险三类。在采用 OVO 分解融合方法处理多类问题的前提下，提出了两种新的多类不平衡企业信用评价模型，即 OVO-AB-OVO-LightGBTGBM-M1 模型和 OVO-AB-GBM-M2 模型，将 AB 不平衡处理方法与 LightGBM 集成分类器相结合。Yujia Chen, Raffaella Calabrese(2023) 等人通过分析局部可解释模型无关解释和局部解释解释 (LIME)，并研究了它们在不同的类不平衡水平上的解释性能。使用由欧洲数据中心提供的住宅抵押贷款数据和另外两个开源的信用评分数据集来验证结果的稳健性。XGBoost 和随机森林被选为“黑盒”机器学习模型来生成预测。Furkan Baser(2022) 等提出一种基于聚类的模糊分类 (CBFC) 的信用风险评估方法，该方法根据 FKM 聚类分布给聚类，通过提高 ML 算法的预测能力。通过对具有模糊隶属度值的模型输出进

行加权,开发一种计算每个输入的单一违约概率(PD)的算法。构建了模糊聚类与监督学习相结合策略的信用评分模型。Boughaci 等人(2021)在估计随机森林集成模型之前也使用 k-means 聚类来提高财务困境预测性能。Golbayani 等(2020)对应用机器学习技术预测信用评级的文献结果进行了调查和比较分析,将之前研究中被认为有用的四种机器学习技术(袋装决策树、随机森林、支持向量机和多层感知机)应用于相同的数据集。我们使用 10 倍交叉验证技术评估结果,得出基于决策树的模型具有优越的性能。

Bao 人(2019)提出了一种将无监督学习与监督学习相结合的信用风险评估策略,提出一种基于聚类和分类方法相结合的算法来评估个体的信用风险。在此背景下,作者讨论了 7 种监督分类模型以及两种不同的聚类模型,k-均值和自组织映射。Terry Harris(2013)通过支持向量机的方法遴选出信用风险评价指标,构建了巴巴多斯信贷联盟的信用评价模型。

### 1.2.3 文献综述总结

在信用评价体系的筛选上,应用较为广泛的方法是偏相关分析、支持向量、因子分析与 Probit 模型等方法,这些方法受各个因素的影响,比如各个企业的特殊性,行业标准差异性,或者受限于相关数据的可获得性等,这些对因素对评价体系筛选的准确性有待验证,在筛选指标时有些方法像模糊层次分析法存在主观性,导致结论存在客观偏差。这些体系在实际应用中存在一些不足,企业信用评级模型需要考虑到不同行业间的经营特色,企业经营状况等因素,可以考虑加入企业的业务往来数据作为参考指标,提升评价结果的准确性。

在信用评价模型方面,选择预测信用评价的算法时,需要根据实际情况综合考虑各种因素。对于简单的二分类问题,逻辑回归可能是一个不错的选择;对于需要解释性的场景,决策树可能更合适;对于需要高预测准确率的场景,随机森林和支持向量机可能是更好的选择;而对于数据量大、关系复杂的场景,神经网络可能具有更好的性能。MLP 神经网络,持向量机,Logistic 回归信用评价模型,AHP 加熵权法,ESG,SVM 等方法在研究上应用较广泛,其中机器学习的相关方法能更为客观地度量企业真实的信用质量,基于机器学习的信用评价模型则更加侧重于对数据的挖掘和深度分析。通过运用复杂的算法和模型,机器学习模型能够自动从数据中提取有用的信息,并对信用风险进行更为精准和个性化的评估。这类模型在处理非线性关系、高维度数据以及动态变化



方面具有优势,适用于复杂多变的信用评估场景。且基于 BP 神经网络的数据挖掘算法,模型的准确度高。随机森林在信用评估中表现出较高的分类准确率和稳定性,并且具有较高的噪声容忍度,能够有效避免过拟合现象。

在信用等级划分方面,在不同的研究中,企业信用等级的划分标准和方法存在差异,不具备统一的标准和一致性。不同研究结果都是基于不同的行业或者业务模式下构建的划分方法,具有不可比性。如果是专业的测评机构需要选择具有统一标准的划分方法,但是对于企业来说,可以根据与客户的业务往来状况来选择适合自己业务的划分方式。

综上所述,在新的研究中,对于 A 公司来说,在指标的筛选上,可以在采用财务相关指标的基础上,加入 A 公司企业客户的业务往来相关的业务指标来研究这些指标对评价结果的影响。在筛选评价模型算法时,可以考虑使用机器学习算法,如 BP 神经网络和随机森林来进行信用评价体系的构建,减少主观性,通过两个模型的对比筛选存储更适合 A 公司的评价模型。

## 1.3 研究问题 and 研究内容

### 1.3.1 研究问题

A 公司目前的信用评价系统是基于客户的财务报表、业务相关数据,使用业务制定的评分规则来计算出相应的信用评分,在法务层面是人工审核企业信息并给予主观评分,系统得出评分后,后续会有主观的授信建议阶段,最终的评价模型涉及到的指标较多,而且存在主观因素。

随着业务的发展,该模型所得出的评价信息与实际客户的应收账款逾期情况逐渐趋于不一致情况,信用评级较高的客户应收账款逾期的情况发生概率并不低。同时经查看数据客户发生逾期的情况在各评级等级中发生概率差别较小,无法从评级中看出客户未来是否会发生逾期。

同时,信用等级高低与未来是否会逾期的关系并不是高度相关的关系。虽然信用等级可以提供一定的预测能力,但仍需要综合考虑其他因素,以更准确地评估实体的未来信用行为。信用等级并不是预测未来的唯一依据。尽管高信用等级的企业客户未来逾期的风险较低,但仍有可能受到市场变化、突发事件、政策调整等因素的影响而出现逾期

行为。同样，低信用等级的实体也有可能通过改善经营管理、加强信用管理等方式提高自身的信用状况，降低未来逾期的风险。

所以本次研究的问题是期望将影响企业客户实际违约的因素挖掘出来，计算出一套更为客观的信用评价体系，在能看到客户信用评价的同时，也能预测到该客户未来的逾期情况。同时使评级模型脱离主观评价，使客户的评级更为客观。

### 1.3.2 研究内容

本论文的主要研究内容如下：

（1）针对当前 A 公司在企业客户信用等级评价上遇到的问题，结合公司的业务背景及国内外相关文献的研究，进行影响企业客户信用指标的分析，预计采用财务相关指标、业务相关指标、及企业资质等信息分析影响客户未来逾期的情况。

（2）基于构建出的信用评价指标体系，结合国内外相关的理论研究，使用收集到 A 公司企业客户的相关的历史数据作为训练数据，拟采用 BP 神经网络及二元 Logistic 逻辑回归方法来进行信用评价模型的建立，得出企业客户的信用等级和未来的逾期情况。

（3）开展应用研究。以 A 公司近期的业务数据作为验证数据，结合构建的信用等级评价模型进行应用实现，分析结果，得出结论。

### 1.3.3 技术路线

本文通过理论与实证研究相结合方法,研究了企业客户的信用风险评级指标体系的构建,评价模型的建立,研究的技术路线如下图所示。

## 1.4 创新点

本文的创新点有以下两点：

（1）在评价模型中增加企业客户法务相关的指标，法务因素对企业的信用评价有着重要的影响。目前 A 公司中的法务指标均为法务部门相关人员手动打分，存在较大主观性，此研究中将引入法务指标，研究这些法务指标对客户信用的影响性。

（2）此研究的信用评价模型得出的是客户的信用等级和未来是否会逾期这两个结果。在查看客户的等级的同时也能观测到该客户未来的信用情况，从而可以给该客户制定相应的应收账款的账期，减少逾期的发生。

（3）使用真实的客户逾期数据，逾期数据是信用评价模型中的重要指标之一，能够

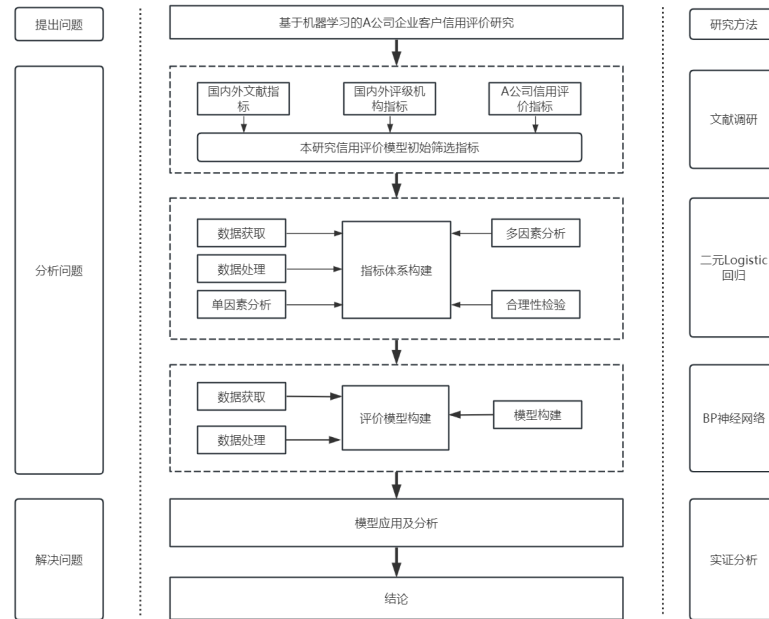


图 1 技术路线图

直接反映客户的信用风险和还款能力。使用真实的客户逾期数据可以帮助构建信用评价模型的训练和验证数据集。通过对逾期数据进行特征提取和模型训练，可以建立更贴近实际情况的信用评价模型，提高模型的预测准确性和泛化能力。同时通过使用逾期数据作为测试集，可以对模型的准确率指标进行评估，从而验证模型的预测能力和稳健性。

## 1.5 章节结构

本文分为四个章节，各章节研究框架与内容安排如下：

第一章为绪论部分。本部分依次对本研究的研究背景和意义、国内外文献研究、研究技术路线、研究内容、创新点进行阐述。

第二章为本文模型构建的理论部分，分别对本文中用到的二元 Logistic 回归及 BP 神经网络在模型构建过程中用到的核心理论及技术点进行阐述。

第三章是本文的核心章节，是信用评价模型构建的过程的描写。首先基于文献、权威机构及 A 公司现有的信用评价系统筛选出评价指标，再使用二元 Logistic 回归进行单因素、多因素的指标筛选，最终筛选出影响企业客户是否逾期的评价指标。再基于已筛选出的指标对是否逾期和客户信用等级使用 BP 神经网络构建信用评价模型。

第四章是基于构建的评价模型，使用 A 公司企业客户的数据进行实证分析，通过模型预测出客户的信用等级和未来是否逾期，再与实际情况进行对比分析，进行模型的检验。

最后为本文的结论部分，在该章节中将对研究成果进行总结，并深入剖析研究中存在的不足，同时提出相应的改进建议，以期为未来的研究和实践提供有价值的参考。

## 第二章 相关基本理论

### 2.1 信用评价的相关理论

#### 2.1.1 信用评价的目的

信用评价（Credit Evaluation）也称为信用评估、信用评级、资信评估、资信评级，是以一套相关指标体系为考量基础，标示出个人或企业偿付其债务能力和意愿的过程。信用评价是一个企业履约状况和偿债能力综合反映。这种评估基于一系列因素，包括财务健康状况、历史偿还记录、业务稳定性、管理质量、市场地位等。信用评价的目的是帮助贷款方、投资者和其他利益相关者了解借款方的信用风险，从而做出更明智的决策。

信用评价通常分为不同的等级，不同的信用评级机构采用不同的评估标准、方法和模型来进行信用评价，也会采用不同的等级划分标准和命名方式。信用评价可以由专门的信用评级机构（如穆迪、标准普尔、惠誉等）进行，也可以由银行、保险公司等金融机构内部进行。

信用等级并不是绝对的，它会随着市场环境、企业经营状况等因素的变化而发生变化。投资者或者企业在参考信用等级时，需要了解评级机构的评估标准和方法，并结合实际情况进行综合分析。

#### 2.1.2 信用评价的分类

信用评价的主要类型和分类可以根据不同的标准来划分。以下是几种常见的分类方式：

按评价对象分类：企业信用评价：主要评估企业的偿债能力、履约能力、守信程度等，以判断企业的信用状况。个人信用评价：主要评估个人的履约能力、信用记录、还款意愿等，以判断个人的信用状况。国家信用评价：主要评估一个国家的偿债能力、政治稳定性、经济前景等，以判断国家的信用状况。

按评价范围分类：综合信用评价：对企业或个人的整体信用状况进行全面评估，包括财务状况、经营能力、履约记录等多个方面。专项信用评价：针对某一特定方面或领域进行信用评估，如债券信用评价、担保机构信用评价等。

按评价方法分类：定性评价：主要依赖于评估人员的经验和专业知识，通过主观判断来评估信用状况。定量评价：主要依赖于数学模型和统计数据，通过客观分析来评估信用状况。

按评价机构分类：内部信用评价：由企业或机构内部设立的信用评价部门进行，主要用于内部管理和风险控制。外部信用评价：由独立的第三方信用评价机构进行，结果更具公正性和权威性，广泛应用于市场。

按评价等级分类：信用评价通常分为不同的等级，如 AAA 级、AA 级、A 级、BBB 级、BB 级、B 级、CCC 级、CC 级、C 级和 D 级等。不同机构对等级的定义可能略有差异，但一般来说，AAA 级表示信用状况最好，违约风险最低；而 D 级则表示信用状况最差，违约风险最高。

本文研究的信用评价的类别是企业客户的信用等级，将企业客户的信用等级进行划分，为 A 公司信用相关人员提供一个清晰的信用参考，帮助他们做出更明智的决策，降低客户的信用风险，促进商业合作。

## 2.2 二元 Logistic 逻辑回归

### 2.2.1 sigmoid 函数

二元 Logistic 回归假设输出变量服从伯努利分布，即取值为 0 或 1 的离散分布。模型的目标是通过给定的输入变量，预测输出变量为 0 或 1 的概率。模型的核心是 Logistic 函数，该函数可以将输入变量的线性组合映射到一个 0 到 1 之间的数值。这种映射使得 Logistic 回归能够处理二分类问题。logistic 函数也就是经常说的 sigmoid 函数，是一个 S 型曲线，公式如下：

$$y = \frac{1}{1 + e^{-x}} \quad (2.1)$$

sigmoid 函数图像如下：

因变量  $y$  为二分变量，其中阳性值取 1，阴性值取 0。假设  $p$  为取阳性的概率，那么  $1-p$  就是阴性的概率。阳性与阴性的比值就成为胜算（odds），定义胜算的公式为：

$$odds = \frac{p}{1 - p} \quad (2.2)$$

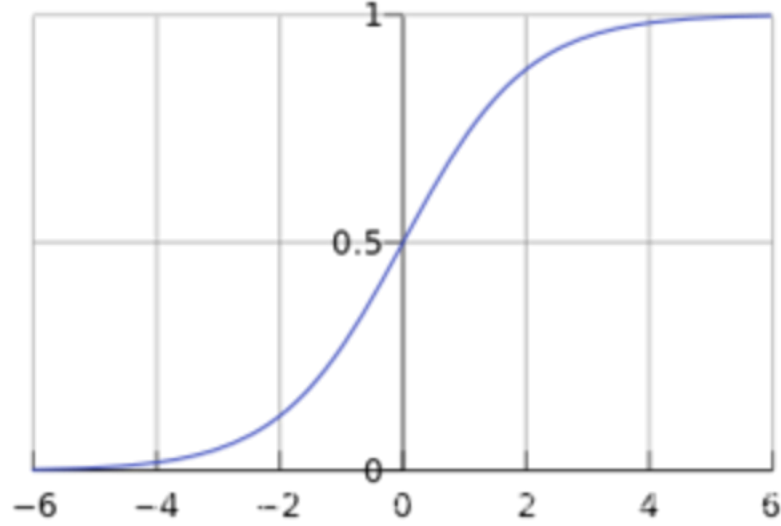


图 2 Sigmoid 函数图像

该公式表示阳性的概率是阴性概率的多少倍。对胜算取自然对数，就可以得到 Logistic 的逻辑回归模型。Logistic 回归模型是建立，在  $\ln(\frac{p}{1-p})$  与自变量的线性回归模型。公式如下：

$$\ln\left(\frac{p}{1-p}\right) = f(x) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k \quad (2.3)$$

其中  $\ln(\frac{p}{1-p})$  服从二元 logistic 分布，取值范围为  $(-\infty, +\infty)$ ,  $X_1$  到  $X_k$  是  $k$  个自变量，取值范围可以是任意范围。自变量可以是连续变量，也可以是分类变量。如果  $X$  是分类变量，则需要将其处理为虚拟变量。方程的右边是一个线性回归方程，左边为胜算 odds 自然对数，经过对数转换，即可得到 Logistic 回归模型的线性模型。其中系数  $\beta$  是回归系数值，即为二元 logistic 回归需要计算出来的值， $X$  增加 1 个单位，则胜算的自然对数会增加  $\beta$  的倍数。则取阳性的概率  $p$  可以表示为：

$$p = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k)}} \quad (2.4)$$

### 2.2.2 回归模型估计

#### (1) 模型的参数估计

Logistic 二元回归是使用极大似然估计（MLE）做参数估计。极大似然估计的基本思想是在给定一组观测数据的情况下，找到一组参数，使得这组参数能够最大化产生观测数据的概率。在 Logistic 回归的上下文中，这意味着找到一组参数（通常是权重和偏

置),使得根据这些参数计算出的观测对象属于某个类别的概率与实际观测到的标签最匹配。通过极大似然估计,可以得到 Logistic 回归模型的参数估计值,进而使用这些参数来预测新观测对象的类别概率。

## (2) 拟合优度的检验

评估 Logistic 回归模型拟合优度的统计检验方法为 Hosmer-Lemeshow 检验,Hosmer-Lemeshow 检验的基本思想是将模型预测的概率分成若干组,然后比较每组的实际观测值和预期值之间的差异。具体来说,它将样本按照模型预测的概率从低到高排序,然后分为若干个等概率的组。在每个组内,计算实际观测事件发生的比例和模型预测的事件发生的比例,并计算它们的差异。如果模型拟合得好,那么这些差异应该很小。

Hosmer-Lemeshow 检验的原始假设是测量值分布和期望值分布之间没有显著性差异。如果检验的显著性 P 结果大于显著水平(一般取 0.05),则接受原假设,模型拟合优度较好;当 P 值小于显著性水平时,拒绝原假设,模型拟合度较差。

(3) 模型的假设检验 Logistic 回归模型的假设检验,常用的检验方法有似然比检验(likelihood ratio test)、计分检验(score test)和 Wald 检验(Wald test),这些检验方法用于检验模型中各个自变量的系数是否显著,即判断自变量对因变量的影响是否存在统计学意义。

likelihood ratio 检验:用于比较一个完整模型和一个简化模型之间的拟合优度,从而判断模型是否需要包含特定的自变量。基本思想是比较两种不同假设条件下,对数似然函数值的差别大小。它可以用来检验整个模型是否具有统计学意义,即所有自变量的总体回归系数是否均为 0。在应用 LR 检验时,需要满足适用数据类型、样本容量要求、独立性假设、线性关系假设、无多重共线性等条件。此外,还需要对模型的拟合度和预测能力进行评估,并进行结果的解释和推断。只有在满足这些条件的情况下,才能获得准确可靠的 LR 检验结果,并对研究问题进行科学合理的判断和决策。LR 检验的基本假设之一是样本之间是相互独立的。即各个样本之间的观测值不会相互影响。此假设在实际研究中往往不容易满足,因此需要在样本选择和实验设计上进行合理的控制,以尽量减少样本之间的相关性。LR 检验的另一个重要假设是自变量和因变量之间存在线性关系。简单来说,就是自变量的变化对因变量的影响是线性的,而不是非线性的。如果自变量和因变量之间存在非线性关系,需要进行适当的数据变换或者选择其他适用的统计方法。



计分检验也是一种常用的假设检验方法，其基本原理和似然比检验类似，都是通过比较不同模型的对数似然函数值来进行检验。计分检验可以用于检验模型的某些假设是否成立，例如检验自变量是否满足线性关系或检验模型是否满足比例性假设等。计分检验的基本步骤构建模型的得分函数，计算得分函数的期望值；构建检验统计量，最后进行假设检验：根据检验统计量的分布和给定的显著性水平，进行假设检验。如果检验统计量的值大于临界值，则拒绝原假设，认为模型的某些假设不成立；否则，接受原假设。

Wald 检验被广泛应用于医学、社会科学、金融等领域中，用于判断实验结果是否具有统计显著性，从而推断出实验中所探究的变量之间是否存在相关性。由于 Wald 检验的计算方法比较易懂，所以在统计学中也被广泛使用。本文采用 Wald 检验，Wald 检验是每个自变量是否与因变量显著性影响，如果有多重共线性时结果不准确。Wald 检验是先对原方程（无约束模型）进行估计，得到参数的估计值，再代入约束条件检查约束条件是否成立。Wald 检验的优点是只需估计无约束一个模型。用  $u$  检验或者  $x$  平方检验，推断各参数  $B$  是否为 0，其中使用 Wald 统计量来检验每个参数的显著性。Wald 统计量的计算公式为：

$$W = \frac{b - \beta}{SE(b)} \quad (2.5)$$

其中， $b$  为模型估计的系数， $\beta$  为假设的值， $SE(b)$  为  $b$  的标准误。计算出  $W$  值后，可以将其与标准正态分布进行比较，如果  $W$  值大于标准正态分布的临界值，则说明该参数显著。需要注意的是，Wald 检验通常适用于样本量较大的情况，如果样本量较小，则可能会出现  $W$  值过大而导致显著性判断错误的情况。

## 2.3 BP 神经网络

### 2.3.1 BP 神经网络信号传播

BP 神经网络（Back Propagation），作为一种多层前馈神经网络，核心特性在于信号的正向传递与误差的反向传播。该网络由输入层、隐含层和输出层构成，每一层均含有多个并行工作的神经元。在正向传播阶段，输入信号通过输入层与隐含层，最终到达输出层。训练过程中，一旦输出值与预期值之间存在误差，该误差将通过反向传播的方式

逐层传递,进而动态调整每个神经元间的阈值及其连接权值。经过多轮的训练迭代,模型的输出值会逐渐逼近预期值。以下是 BP 神经网络的简化结构图:

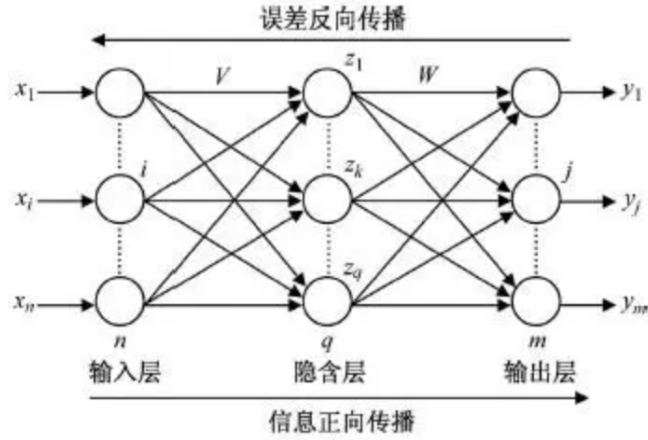


图 3 BP 神经网络信号传播图

上图中, $x_1$  到  $x_n$  是 BP 神经网络的输入值, $y_1$  到  $y_m$  是 BP 神经网络的预测值, $V$  和  $W$  为 BP 神经网络权值。从图可以看出,BP 神经网络可以看成是一个非线性函数,网络输入值和预测值分别为该函数的自变量和因变量。当输入节点数为、输出节点数为  $m$  时,BP 神经网络就表达了从  $n$  个自变量到  $m$  个因变量的函数映射关系。

### 2.3.2 训练过程

BP 神经网络学习过程包括以下步骤,

(1) 网络初始化: 首先,需要对神经网络的各连接权值和阈值进行初始化,通常赋予它们  $(0,1)$  区间内的随机数值。同时,设定误差函数和计算精度值。确定输入层、输出层、隐含层节点个数,假设分别为  $n$ 、 $m$ 、 $M$ ,其中  $M$  的公式一般取如下:

$$M = \sqrt{n + m} + a \quad (2.6)$$

其中  $a$  的取值范围为 1-10。BP 神经网络对于分类问题一般使用交叉熵损失函数作为误差函数。二分类时,交叉熵损失函数如下:

$$L = -[y \log(p) + (1 - y) \log(1 - p)] \quad (2.7)$$

其中,其中:  $y$  是实际标签 (0 或 1)。 $p$  是模型预测为正类的概率。 $\log$  是自然对数。多分类输出时,误差函数拓展如下:

$$L = - \sum_{i=1}^C y_i \log(p_i) \quad (2.8)$$

其中,  $C$  是类别的总数,  $y_i$  是样本实际属于第  $i$  个类别的指示函数 ( $y_i = 1$  如果样本属于第  $i$  类, 否则  $y_i = 0$ ),  $p_i$  是模型预测样本属于第  $i$  个类别的概率。

(2) 正向传播: 在正向传播过程中, 输入样本从输入层传入, 经过各隐含层神经节点的处理 (包括加权求和和非线性激活函数), 然后逐层传递, 最终从输出层输出信息。每个神经元接收来自前一层所有神经元的加权输入, 然后加上一个偏置 (Bias), 并通过激活函数 (Activation Function) 来产生输出信号。

(3) 计算输出层误差: 计算网络的实际输出与期望输出之间的误差。这个误差反映了当前网络对输入样本的处理能力, 也是后续权值调整的依据。

(4) 误差反向传播: 如果输出层的误差大于设定的精度值, 那么将误差逐层反向传播到之前的各层。在这个过程中, 根据误差信号调整各层神经元的权值和阈值, 使网络的输出向期望输出逼近。

(5) 权值更新: 根据误差反向传播的结果, 按照一定的学习率更新各层神经元的权值和阈值。这个过程是神经网络学习的核心, 通过不断调整权值和阈值, 使网络能够更好地适应输入样本。

(6) 迭代学习: 重复以上步骤, 直到整个训练样本集的误差减小到符合要求为止。在这个过程中, 神经网络的性能会逐渐提高, 对输入样本的处理能力也会逐渐增强。

以上就是 BP 神经网络学习过程的主要步骤。需要注意的是, 这个过程是一个迭代的过程, 需要反复进行权值调整和误差计算, 直到达到设定的精度要求或者达到预设的最大迭代次数。

### 2.3.3 激活函数

BP 神经网络的激活函数 (Activation Function) 是在神经网络的神经元上运行的函数, 负责将神经元的输入映射到输出端。激活函数对于神经网络模型去理解、学习复杂的非线性函数具有十分重要的作用。几种常用的激活函数如下:

#### (1) Sigmoid 函数

Sigmoid 函数已在二元 Logistic 回归中介绍过, 见公式 (2.1), BP 神经网络和 Logistic 回归中使用的 Sigmoid 函数在本质上是一样的, 都是 S 型的非线性函数, 可以将连续的

输入映射到 0 到 1 的输出范围内。然而，在 BP 神经网络和 Logistic 回归中，Sigmoid 函数的应用方式有所不同。

在 BP 神经网络中，Sigmoid 函数通常用作隐藏层和输出层的激活函数，用于将神经元的输出映射到 0 到 1 的范围内。这使得网络能够学习和模拟非线性函数，从而提高了网络的表达能力和泛化能力。在训练过程中，BP 算法通过反向传播误差来更新网络的权重和偏置，使得网络的输出逐渐逼近期望的输出。而在 Logistic 回归中，Sigmoid 函数则用作输出函数，将回归模型的输出映射到 0 到 1 的范围内，从而得到一个概率值。这个概率值表示了样本属于正类的概率。在训练过程中，Logistic 回归使用最大似然估计来求解模型的参数，使得模型能够最小化预测错误。

### (2) Tanh 函数

Tanh 函数是双曲正切（Hyperbolic Tangent）函数的缩写，是 Sigmoid 函数的一种变体，它将连续的输入映射到 (-1,1) 的输出范围内。与 Sigmoid 函数相比，Tanh 函数在零点附近的导数更大，从而可能在一定程度上减轻梯度消失的问题。

### (3) ReLU 函数（Rectified Linear Unit）

ReLU 函数是近年来非常受欢迎的激活函数之一。对于输入值小于 0 的情况，ReLU 函数的输出为 0；对于输入值大于 0 的情况，ReLU 函数的输出等于输入值。ReLU 函数具有简单的数学形式和计算效率高的特点，并且在一定程度上能够缓解梯度消失的问题。然而，它也可能导致神经元输出恒为 0 的问题。

### (4) Softmax 函数

Softmax 函数通常用于多分类问题的输出层。本文的输出就涉及到多分类的问题，所以会用到该函数。它将一组输入映射到一个概率分布上，使得所有输出的概率之和为 1。Softmax 函数在 BP 神经网络中用于将神经元的输出转换为概率值。Softmax 函数的定义如下：

$$Softmax(z_i) = \frac{e^{z_i}}{\sum_{k=1}^K e^{z_k}} \quad (2.9)$$

其中  $z_i$  为第  $i$  个节点的输出值， $K$  为输出节点的个数，即分类的类别个数。通过 Softmax 函数就可以将多分类的输出值转换为范围在  $[0, 1]$  和为 1 的概率分布。

### 2.3.4 梯度下降法

梯度下降法是一种优化算法，用于寻找最小化损失函数的参数。BP 神经网络中通过训练误差来逐步调整各层间的输入权重和偏置，这个调整的过程依据的算法一般有两种，一个是梯度下降法（Gradient Descent），另一个是最小二乘法。

训练误差（或称为损失函数）是依赖于神经网络中的权重和偏置的二次函数。为了最小化这个误差，需要计算损失函数关于权重和偏置的偏导数，这些偏导数组成了梯度向量。梯度向量的方向指示了训练误差增加最快的方向，而梯度向量的相反方向则是训练误差减少最快的方向。因此，沿着梯度向量的反方向进行权重和偏置的更新，可以更有效地找到损失函数的最小值，从而实现神经网络的优化。

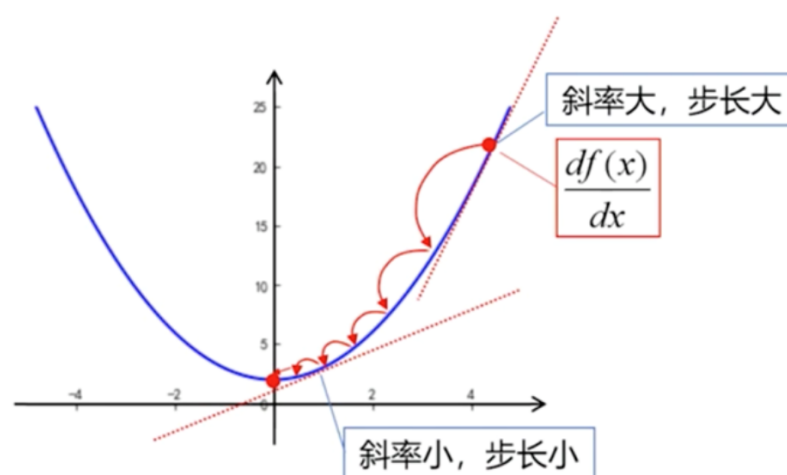


图 4 损失函数图

假设上图中描绘的曲线代表了损失函数的形状，如果设置的步长太小，迭代次数就会多，收敛慢，如果设置的步长太大，会引起震荡，可能导致无法收敛。该函数存在一个最小值点，是误差率最小，也就是斜率最小。梯度是指损失函数在某一点处的方向导数，沿着这个方向，损失函数的变化率最大。在 BP 神经网络中，梯度下降法通过计算损失函数关于权重和偏置的偏导数（即梯度），然后沿着梯度的反方向更新权重和偏置，以减小损失函数的值。这个过程不断迭代进行，直到损失函数的值收敛到一个最小值。这种方法就是通过不断地迭代更新参数，最终实现网络的优化。假设损失函数为  $E$ ，权重为  $w$ ，偏置为  $b$ ，学习率为  $\eta$ 。在梯度下降法中，按照以下公式更新权重和偏置：

对于权重：

$$w_{new} = w_{old} - \eta \frac{\partial E}{\partial w} \quad (2.10)$$

对于偏置:

$$b_{new} = b_{old} - \eta \frac{\partial E}{\partial b} \quad (2.11)$$

其中,  $(\frac{\partial E}{\partial w})$  和  $(\frac{\partial E}{\partial b})$  分别表示损失函数  $E$  对权重  $w$  和偏置  $b$  的偏导数,  $\eta$  是学习率, 它决定了参数更新的步长。

这两个公式表示, 在每次迭代中, 都按照损失函数在当前点的梯度方向 (即损失函数增加最快的方向) 的相反方向来更新权重和偏置, 从而逐渐逼近损失函数的最小值。

梯度下降法有多种变体, 如批量梯度下降法、随机梯度下降法和小批量梯度下降法等。这些方法的主要区别在于每次更新权重和偏置时所使用的样本数量不同。批量梯度下降法使用所有的样本数据进行计算, 而随机梯度下降法则每次只使用一个样本数据进行计算。小批量梯度下降法则是介于两者之间, 每次使用一部分样本数据进行计算。不同的方法在不同的应用场景下可能有不同的表现, 需要根据实际情况进行选择。

### 2.3.5 模型检验

如果要查看 BP 神经网络模型建立的效果如何, 可以通过以下几个方面进行检验:

**误差下降曲线图:** BP 神经网络预测模型结果需要查看误差下降曲线图。然而, 一般不必过分关注这条曲线, 除非是为了研究改进算法以提高收敛速度。相比之下, 更应关注网络的实际训练效果和实际应用能力, 例如预测能力等。

**训练集和测试集的表现:** 可以通过观察模型在训练集和测试集上的表现来评估模型的效果。如果模型在训练集上表现良好, 但在测试集上表现不佳, 可能出现了过拟合的问题。相反, 如果模型在训练集和测试集上的表现都不好, 可能是模型复杂度不足, 即欠拟合。

**过拟合和欠拟合的解决方法:** 针对过拟合问题, 可以尝试增加训练数据量、使用正则化约束、调整参数和超参数、降低模型复杂度或使用 Dropout 等方法。对于欠拟合问题, 可以尝试增加模型的复杂度、增加样本有效特征数、调整参数和超参数或使用更复杂的网络结构等方法。

实际应用效果：最终，模型的效果还需要通过在实际应用中的表现来评估。例如，可以观察模型在预测任务中的准确率、召回率、F1 分数等指标，以及模型在处理实际问题时的稳定性和鲁棒性。综上所述，BP 神经网络模型建立的效果需要综合考虑多个方面的因素来评估。在实际应用中，可能需要根据具体问题和数据集的特点来选择合适的评估方法和优化策略。

## 2.4 本章小结

本章探讨了客户信用评价的目的及分类标准，同时详细阐述了在模型构建过程中所使用的二元 Logistic 回归和 BP 神经网络的核心理论及技术点。通过对这些理论和技术点的细致剖析，旨在更深刻地理解它们在模型构建中所发挥的作用，从而为读者提供一个全面而深入的理论框架。

在二元 Logistic 回归部分，重点介绍了在筛选指标过程中所使用的函数、参数估计方法以及模型检测的相关技巧。这些函数和方法的运用，能够有效地识别并筛选出与信用评价密切相关的关键指标，进而构建出更加精准和可靠的预测模型。

在 BP 神经网络部分，详细阐述了信号传播的原理及过程，并介绍了在训练过程中所使用的步骤及函数。此外，还介绍了梯度下降法的损失函数相关理论，以及 BP 神经网络模型检验的方法。这些理论和方法的运用，有助于更好地理解和应用 BP 神经网络。

## 第三章 信用评价模型构建

### 3.1 基于二元 Logistic 回归构建指标体系

#### 3.1.1 数据选取

本研究的数据来源是从 A 公司的信用管理系统及 ERP 系统中获取到的数据，这些数据是关于企业客户在信用管理系统中做过信用评价的客户数据，部分数据是通过企查查第三方信息查询机构中获取的，比如企业是否上市、不良记录等信息。涉及了从 2019 年至 2023 期间 A 公司所有进行过信用评价的客户，考虑到国外客户的信息无法获取的因素，该二元 Logistics 研究只取了 A 公司国内客户的数据，排除无财务相关指标的数据及错误数据的客户，获取到的初始数据有 1861 条。其中下一节中介绍的指标均为模型中的自变量，本模型的因变量为是否逾期 FutureOverdueFlag。是否逾期的定义为：客户在信用管理系统进行过信用评价的申请后，如果在 6 个月内存在应收账款逾期记录，则就被标记为逾期。

是否逾期说明：本文基于 A 公司的信用评价业务基础进行研究，该业务中，客户在申请信用评价后，会得出该客户的账期，如果客户在账期内未付清账款，发生逾期，逾期产生七天后仍逾期，则该客户的逾期记录被归档。客户是否有历史逾期的定义：若一个客户在申请信用评价时，若过去 6 个月以内存在逾期归档记录，则该客户就被记录为有逾期记录。本文延用该定义，在此基础上，研究客户未来是否会产生逾期。未来是否会逾期定义为：若一个客户在申请信用评价时，未来 6 个月以内会产生逾期且会被归档，则该客户未来是否会逾期将被记为是，否则为否。

#### 3.1.2 指标筛选

##### 1. 财务指标

在参考了国内外评级机构及信用评价相关文献中，本研究选择的财务指标最核心的四大指标，有盈利能力、变现能力、偿债能力、资产管理能力，这些是企业信用评价非常重要的财务指标，它们能够为信用评价时提供企业是否有全面的财务健康状况，企业在面对短期偿债压力时的应对能力，债务偿还能力以及风险应对情况的评估。指标详见表 1：

##### 2. 业务指标



表 1 信用评级财务指标

一级指标	二级指标	指标符号
盈利能力	销售净利率	NPM
	销售毛利率	GPM
	净资产收益率	ROE
变现能力	流动比率	CR
	速动比率	QR
	现金比率	CASHR
资产管理能力	存货周转率	ITR
	应收账款周转率	ARTR
偿债能力	资产负债率	DAR

业务指标是筛选出与 A 公司有业务往来的客户的客观数据，分为三个指标：信用记录、增信措施、经营指标。信用记录指标反映的是企业在过去的与 A 公司的经营活动中所形成的信用行为和信用历史，可以反映企业过去的信用表现，展现客户的信用行为和信用管理能力，信用记录良好的客户通常能够获得更好的信用评级。增信措施是客户为提高信用而采取的措施，例如提供担保、投保信用险等。增信措施可以帮助企业降低债权人的信用风险，增强债务偿还能力。经营指标是企业的各项经营数据和，这些指标可以反映企业的经营状况。良好的经营指标通常可以为企业赢得更好的信用。本文采用的业务指标主要有以下指标：

表 2 信用评级业务指标

一级指标	二级指标	指标符号
信用记录	是否有逾期记录 不良记录	OverdueFlag BADR
增信措施	是否有银行保函 是否投保信用险 是否关联担保公司	BGFlag CTFlag GCFlag
经营指标	客户经营状态	CES
	经营负荷情况	OLC
	内销比例	DSR
	外销比例	ESR
	是否为新客户	NewFlag

### 3. 企业资质指标

企业资质指标可以反映企业的经营实力和规模水平，也可以反映企业的历史背景和稳定性，也能展示出企业的发展潜力和未来增长空间。拥有相关的资质认定可以表明企业在特定领域具有一定的专业能力和技术实力，同时也有助于提升企业的品牌形象和知名度，这些对于企业的信用评价具有积极的影响。企业成立年限是衡量企业稳定性

和可靠性的重要指标之一，成立年限较长的企业在面对市场波动和风险挑战时，往往具有更强的抵抗能力和适应能力，其信用状况也更为可靠。获得 ISO 认证意味着企业已经通过了一系列国际标准的审核，证明其在质量、环境、安全等方面达到了一定的要求。在信用评价中，拥有 ISO 资格认证的企业往往更容易获得较高的信用评级。上市公司需要遵守更为严格的法规和监管要求，包括定期公开财务报告、接受审计和遵守证券市场的相关规定。这种透明度和规范性有助于提升公司的信用评级。净资产代表企业所有者对企业资产的净权益，它反映了企业的自有资金实力。一个拥有较大净资产规模的企业，通常意味着其具备较强的财务实力和抵御风险的能力。在信用评价中，这种财务实力是企业获得较高信用评级的重要保障。本文采用的企业资质指标如下表：

表 3 信用评级企业资质指标		
一级指标	二级指标	指标符号
基本信息	企业成立年限	YOE
	ISO 资格认证	ISOFlag
	是否上市	ListedFlag
	净资产规模	NAS

3.1.3 数据预处理

在做二元 Logistic 回归分析之前，通常需要先进行一系列的数据分析工作。这些分析不仅有助于确保数据的准确性和可靠性，还能为后续的 Logistic 回归分析提供重要的背景信息和指导。本文将对企业数据进行深入的相关性分析和多重共线性分析，旨在揭示企业不同变量之间的关联程度以及变量间可能存在的潜在问题。

通过相关性分析，可以了检测探究不同变量之间是否存在某种关联，以及这种关联的程度和方向。这对于理解研究领域的内在机制和影响因素至关重要，为后续模型解释提供基础。同时相关性分析是进行多因素 Logistic 回归分析的基础。在进行多因素分析之前，需要首先了解各自变量与因变量之间的相关性，以便选择最具影响力的自变量进行进一步的分析。同时，多重共线性分析将帮助处理变量间的共线性问题，避免自变量之间存在高度相关的情况，当存在多重共线性时，回归模型的估计结果可能会失真，导致模型对数据的拟合效果不佳，降低模型的稳定性和可靠性。通过综合运用这两种分析方法，将更全面地了解数据的内在关系，为构建准确、可靠的模型打下坚实的基础。

Logistic 回归模型本身是一种概率模型，它分析的是某一事件发生与否的概率与自

变量之间的关系。因此，即使变量之间存在高度相关性，Logistic 回归模型仍然可以处理并给出相应的结果。但是，为了获得更稳定、更准确的模型，需要在进行 Logistic 单因素分析之前先进行相关性分析。

#### (1) Pearson 相关系数分析

本文采用 Pearson 相关系数分析，计算公式为

$$r = \frac{cov(X, Y)}{\sigma X * \sigma Y} \quad (3.1)$$

其中  $cov(X, Y)$  表示 X 和 Y 的协方差， $\sigma X$  和  $\sigma Y$  表示 X 和 Y 的标准差。相关系数的取值范围在-1 到 1 之间，当  $r>0$  时表示正相关， $r<0$  时表示负相关， $r=0$  时表示无相关关系。

使用 Excel 数据分析工具对 3.1.2 章节得出的 23 个指标数据进行分析得出下表：

表 4 指标相关系数 r

相关系数 (r)	流动比率 (CR)	速动比率 (QR)	内销比例 (DSR)
速动比率 (QR)	0.98	1	0.088
现金比率 (CASHR)	0.86	0.88	0.1
外销比例 (ESR)	-0.088	-0.088	-1

经过深入的相关性分析，发现速动比率（QR）与流动比率（CR）的相关系数高达 0.98，现金比率（CASHR）与流动比率（CR）的相关系数为 0.86，而现金比率（CASHR）与速动比率（QR）的相关系数也达到了 0.88，均显示出强烈的正相关性。同时，外销比例（ESR）与内销比例（DSR）之间的相关系数为-1，呈现出极强的负相关性。

鉴于速动比率（QR）与流动比率（CR）以及现金比率（CASHR）之间的强相关性，优先选择了保留现金比率（CASHR），因为它直接反映了企业可立即用于偿还短期债务的能力，是信用评价中的关键指标。较高的现金比率意味着企业短期偿债能力更强，从而有助于提升信用评级。

在内销比例（DSR）和外销比例（ESR）的选择上，保留了外销比例（ESR）作为代表性指标。最终，剔除了速动比率（QR）、流动比率（CR）和内销比例（DSR）。

此外，在相关性分析中，注意到指标“是否有银行保函”（BGFlag）在排除无财务相关指标的数据后，该指标的数据均变为无银行保函，因此不再具有研究价值，删除该指标。

经过这一系列的分析与筛选，最终保留了 19 个二级指标，这些指标将为后续的研究

究提供更加精准和有效的数据支持，如下表：

表 5 信用评级指标

指标大类	一级指标	二级指标	指标符号
财务指标	盈利能力	销售净利率	NPM
		销售毛利率	GPM
		净资产收益率	ROE
		变现能力	CASHR
	资产管理能力	现金比率	CASHR
		存货周转率	ITR
		应收账款周转率	ARTR
		偿债能力	DAR
	信用记录 不良记录	资产负债率	DAR
		是否有逾期记录	OverdueFlag
		BADR	
业务指标	增信措施	是否投保信用险	CTFlag
		是否关联担保公司	GCFlag
		经营指标	CES
		客户经营状态	OLC
		经营负荷情况	ESR
		外销比例	NewFlag
		是否为新客户	
企业资质	基本信息	企业成立年限	YOE
		ISO 资格认证	ISOFlag
		是否上市	ListedFlag
		净资产规模	NAS

## (2) 多重共线性分析

为了避免自变量之间存在高度相关性的情况，需要对变量进行多重共线性检测。多重共线性可能会导致回归系数估计不准确，使得模型的解释性和预测能力降低。本文采用方差膨胀因子 VIF 进行检查，计算公式如下，

$$VIF = 1/(1 - R^2) \quad (3.2)$$

$R^2$  为该自变量与其他自变量的线性相关系数的平方和。

本论文对保留的 19 个指标进行了共线性检测。计算各指标之间的方差膨胀因子 (VIF)，检测结果如下：

当  $0 < VIF < 10$ ，不存在多重共线性；当  $10 \leq VIF < 100$ ，存在较强的多重共线性；当  $VIF \geq 100$ ，存在严重多重共线性。在上述表中，VIF 值最大的为 2.347 < 3，发现这些指标之间不存在严重的共线性问题。这意味着这 19 个指标能够独立地为本研究提供有价值的信息，避免了因共线性导致的分析结果失真。

表 6 指标多重共线性分析

指标	VIF
是否为新客户 NewFlag	1.447
企业成立年限 YOE	1.592
ISO 资格认证 ISO	1.112
是否有逾期记录 OverdueFlag	1.152
其他不良记录 BADR	1.348
是否上市 ListedFlag	1.446
经营负荷情况 OLC	1.099
是否已投保信用险 CTFlag	1.216
是否关联担保公司 CTFlag	1.212
销售净利率 NPM	1.352
销售毛利率 GPM	1.574
净资产收益率 ROE	1.308
资产负债率 DAR	2.347
净资产规模（人民币万元）NAS	1.482
现金比率 CASHR	1.905
存货周转率 ITR	1.171
应收账款周转率 ARTR	1.048
外销比例 ESR	1.242

### 3.1.4 指标数据的处理

数据归一化是将数据归一化至  $[0,1]$  区间，消除不同变量之间性质、量纲、数量级等特征属性的差异，将其转化为一个无量纲的相对数值，也就是标准化数值，使各指标的数值都处于同一个数量级别上，从而便于不同单位或数量级的指标能够进行综合分析和比较。数据标准化在模型运行之前具有重要意义，可以提高模型的稳定性、准确性和收敛速度。数据标准化有以下重要功能：

（1）统一量纲：数据标准化可以将原始数据的量纲转化为无量纲的纯数值，便于不同单位或量级的指标能够进行比较和加权。这样做可以消除由于量纲不同导致的误差，使得模型能够更好地获取数据中的有效特征。

（2）消除异常数据的影响：标准化能够减小异常数据对于模型训练的影响，从而加快模型的收敛速度。标准化处理可以消除奇异数据对模型训练的负面影响，提高模型的稳定性和准确性。

本文 19 种指标中有不同的数据类型，有分类变量和连续变量。分以下步骤进行数据的标准化：

（1）数据清洗，为减少数据冗余提高数据质量和准确率，本文对原始数据进行了删除重复项、删除异常值，删除了缺少财务相关指标数据的处理后，剩余 239 条数据。

（2）数据转换和规范化，其中企业成立年限需要根据企业的注册日期计算出相应的

年限；对于分类变量客户经营状态、ISO 资格认证、是否有逾期记录、其他不良记录、是否上市、经营负荷情况、是否已投保信用险、是否关联担保公司的指标数据规范成为 0 和 1；是否为新客户，新客户值为 1，老客户值为 0；因变量是否逾期也规范为 0 和 1，其中 0 表示否定，1 表示肯定。

本研究将企业成立年限作为分类变量，根据成立年限的长短，将企业成立年限划分为不同的级别，每一级别给定一个具体的分数范围，以便进行量化考核。以下表格是企业成立年限处理后的结果。

表 7 企业成立年限的标准化处理		
指标名称	评分标准	分值
企业成立年限 YOE	$y \geq 10$ 年	1.0
	$5 \text{ 年} \leq y < 10 \text{ 年}$	0.7
	$3 \text{ 年} \leq y < 5 \text{ 年}$	0.4
	$1 \text{ 年} \leq y < 3 \text{ 年}$	0.2
	$y < 1 \text{ 年}$	0

(3) 对连续变量数据的归一化。

对其他连续变量本文采用最小-最大归一化 Min-Max Normalization 标准化，是一种常用的数据预处理技术，用于将数据缩放到一个特定的范围，通常是 0 到 1 之间。这种方法对于消除数据中的量纲和量纲单位的影响非常有用，使得不同特征或指标之间可以进行比较和加权。最小-最大归一化的数学公式如下：

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (3.3)$$

其中， $x$  是原始数据集中的某个特征值， $\min(x)$  是该特征在所有数据中的最小值， $\max(x)$  是该特征在所有数据中的最大值， $x'$  是归一化后的值，它将落在 0 到 1 的范围内。

### 3.1.5 单因素 Logistic 分析

单独考察每一个自变量对因变量的净效应，有助于更深入地理解每一个因素在整体模型中的作用和贡献。为了检测单一因素对是否逾期这一因变量的影响，判断这些自变量和因变量之间的关系是否显著，了解自变量如何影响因变量的概率，保留显著的自变量。单因素的数学模型构建如下：

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_i X_i \quad (3.4)$$

本公式为公式 2.3 变量个数为 1 的情况， $\beta_0$  为常量， $X_i$  是自变量， $\beta_i$  是自变量的系数。

对标准化后的 19 个自变量逐一与是否逾期进行二元 Logistic 逻辑回归，利用 SPSS 工具进行回归得到的显著性 P 值表如下：

表 8 客户逾期单因素回归结果

	B	标准误差	瓦尔德	自由度	显著性	Exp(B)	EXP(B) 的 95% 置信区间	
							下限	上限
是否关联担保公司	1.419	0.726	3.817	1	0.051	4.133	0.995	17.162
经营负荷情况	-1.454	0.602	5.838	1	0.016	0.234	0.072	0.76
是否上市	-1.377	0.497	7.668	1	0.006	0.252	0.095	0.669
是否有逾期记录	0.959	0.436	4.835	1	0.028	2.609	1.11	6.135
应收账款周转率	-3.18	2.316	1.885	1	0.17	0.042	0	3.894
客户经营状态	-1.454	0.602	5.838	1	0.016	0.234	0.072	0.76
资产负债率	1.734	0.718	5.833	1	0.016	5.663	1.387	23.132
销售毛利率	-1.568	1.199	1.709	1	0.191	0.208	0.02	2.187
现金比率	-0.78	0.337	5.348	1	0.021	0.458	0.237	0.888
销售净利率	-4.547	2.485	3.348	1	0.067	0.011	0	1.382
ISO 资格认证	-0.292	0.407	0.515	1	0.473	0.747	0.336	1.658
净资产收益率	-0.192	0.711	0.073	1	0.788	0.826	0.205	3.328
是否已投保信用险	-0.387	0.569	0.464	1	0.496	0.679	0.223	2.07
企业成立年限	-0.083	0.646	0.016	1	0.898	0.921	0.259	3.268
净资产规模（人民币万元）	0	0	0.066	1	0.797	1	1	1
ISO 资格认证	-0.292	0.407	0.515	1	0.473	0.747	0.336	1.658
存货周转率	-0.018	0.05	0.133	1	0.715	0.982	0.891	1.083
其他不良记录	0.267	0.792	0.113	1	0.736	1.306	0.277	6.162
外销比例	-0.356	0.507	0.493	1	0.483	0.701	0.259	1.893

在统计学中，通常使用 P 值（P-value）来衡量某一变量或因素对研究结果的影响是否显著。Logistic 二元逻辑回归的显著性水平通常是通过统计检验来确定的，例如使用 p 值来判断某个自变量是否对因变量有显著影响。显著性水平是一个预设的阈值，用于决定是否拒绝原假设（即认为自变量对因变量没有影响）。常见的显著性水平有 0.05、0.01 等，这意味着如果 p 值小于这些阈值，就拒绝原假设，认为自变量对因变量有显著影响。

由于本次研究所涉及的数据量相对较小，而且单因素分析也是初步筛选指标，为了充分利用有限的的数据资源并尽可能保留与因变量“是否逾期”相关的指标，决定将显著性水平扩展到 0.3。若变量的 p 值  $\leq 0.3$  则就被纳入下一步骤的多因素分析中。之所以设置  $p=0.3$  作为筛选标准，是考虑到数据量较少，避免遗漏某些重要的自变量<sup>[41]</sup>。观察上述表格的数据，可以发现是否关联担保公司、经营负荷情况、是否上市、是否有逾期记录、应收账款周转率、客户经营状态、资产负债率、销售毛利率、现金比率、以及销售净利率这十个自变量在统计检验中的显著性水平均低于 0.3。这意味着这些变量在

模型中具有较高的解释力，能够有效地反映出因变量的变化。因此，将这十个指标纳入 Logistic 多因素回归模型中，不仅是统计上的合理选择，更是为了更全面、更准确地揭示各因素之间的复杂关系及其对因变量的综合影响。

3.1.6 多因素 Logistic 分析

通过单因素分析，筛选出了 10 个与因变量是否逾期 FutureOverdueFlag 显著性较高的自变量。本文将充分利用这 10 个经过初步筛选和显著性水平放宽后保留的重要指标，进行多因素的二元 Logistic 回归分析。多因素二元 Logistic 逻辑回归的数学模型为公式 2.3。利用 SPSS 工具，对数据进行了多因素的二元 Logistic 回归分析，采用了向后 LR（向后似然比）的方法，设定步进概率为 0.3，即当某个自变量的显著性水平大于 0.3 时，该自变量将被从模型中移除。模型最终迭代了 4 次后收敛，得到了包含 6 个自变量的模型，详情如下：

表 9 客户逾期多因素回归结果

	B	标准误差	瓦尔德	自由度	显著性	Exp(B)	EXP(B) 的 95% 置信区间	
							下限	上限
经营负荷情况	-1.384	0.625	4.909	1	0.027	0.25	0.074	0.852
现金比率	-0.557	0.299	3.47	1	0.062	0.573	0.319	1.029
应收账款周转率	-3.792	2.903	1.706	1	0.191	0.023	0	6.67
是否有逾期记录	0.61	0.459	1.766	1	0.184	1.841	0.748	4.528
是否上市	-1.404	0.521	7.276	1	0.007	0.246	0.089	0.681
销售净利率	-2.94	2.639	1.241	1	0.265	0.053	0	9.329

多因素分析继续使用显著性 <0.3 作为检测标准，筛选了自变量，以确保所选指标对因变量“是否逾期”具有一定的解释力。在分析后得出的上表中，经营负荷情况、现金比率、应收账款周转率、是否有逾期记录、是否上市以及销售净利率这六个指标的显著性均小于 0.3。

这些指标的显著性水平低于设定的阈值，表明它们在模型中起到了重要的作用，能够显著影响因变量“是否逾期”的变化。经营负荷情况反映了企业运营的压力和效率，其显著性说明了经营状况对逾期风险的影响；现金比率则体现了企业的流动性状况，对逾期风险具有直接的指示作用；应收账款周转率则揭示了企业应收账款的回收速度，与逾期风险密切相关。

此外，是否有逾期记录作为衡量企业信用状况的重要指标，其显著性小于 0.3，进一步验证了信用历史对逾期风险的影响；是否上市则反映了企业的资本结构和市场监管情况，对逾期风险也有一定的影响；销售净利率则体现了企业的盈利能力和经营效率，对逾期风险同样具有解释力。



综上所述，这六个指标的显著性均小于 0.3，基于这些重要指标，成功构建了一个回归模型。该模型不仅准确地反映了这些自变量与因变量之间的关系，还为后续的风险评估和决策制定提供了有力的数据支持。构建的回归模型如下：

$$\ln\left(\frac{p}{1-p}\right) = 0.913 - 1.138OLC - 0.557CASHR - 3.792ARTR + 0.61OverdueFlag - 1.404ListedFlag - 2.94NPM \quad (3.5)$$

### 3.1.7 指标体系合理性检验

#### 1. 完备性检验

筛选出的 6 个指标经营负荷情况、现金比率、应收账款周转率、是否有逾期记录、是否上市以及销售净利率，在构建逾期风险模型时，涵盖了逾期风险的主要方面：

**经营负荷情况：**这一指标反映了企业的运营压力和效率。过高的经营负荷可能意味着企业面临较大的运营风险，从而增加逾期风险。因此，经营负荷情况是评估企业逾期风险的重要因素之一。

**现金比率：**现金比率是企业现金及现金等价物与流动负债的比值，反映了企业的短期偿债能力。现金比率较低的企业可能面临现金流紧张的问题，从而增加逾期风险。因此，现金比率是衡量企业逾期风险的重要指标。

**应收账款周转率：**应收账款周转率反映了企业应收账款的回收速度。周转率较低可能意味着企业应收账款回收困难，导致现金流紧张，进而增加逾期风险。因此，应收账款周转率也是评估企业逾期风险的重要方面。

**是否有逾期记录：**这一指标直接反映了企业的信用状况。有逾期记录的企业在未来发生逾期的可能性相对较高。因此，是否有逾期记录是判断企业逾期风险的重要依据。

**是否上市：**上市与否反映了企业的资本结构和市场监管情况。上市公司通常受到更严格的监管和信息披露要求，这有助于降低逾期风险。因此，是否上市也是评估企业逾期风险的一个方面。

**销售净利率：**销售净利率反映了企业的盈利能力和经营效率。净利率较低的企业可能面临较大的经营压力，从而增加逾期风险。因此，销售净利率也是衡量企业逾期风险的重要指标之一。

总之，选出的此 6 个指标能够很好地解释逾期风险。这些指标从不同维度反映了企

业的运营状况、财务状况、信用状况以及市场状况，从而全面评估了企业的逾期风险。

2. 拟合程度

基于 3.1.6 构建的二元 Logistic 回归模型，该模型的 Hosmer-Lemeshow 检验显著性如下表：

表 10 霍斯默-莱梅肖检验		
卡方	自由度	显著性
10.287	8	0.245

Hosmer-Lemeshow 检验的显著性为 0.245 表明该检验的 p 值（概率值）为 0.245, 在统计学中，通常将 p 值与一个预定的显著性水平（如 0.05）进行比较，以判断模型拟合优度的好坏。如果 p 值大于显著性水平，则通常认为模型的拟合程度是可以接受的，即模型预测的结果与实际观测结果之间的差异并不显著。本模型的 p 值为 0.245，模型的拟合程度较好。3. 模型的准确率根据表可以看到：

表 11 分类表				
是否逾期	预测		预测	
		0	1	正确百分比
是否逾期	0	178	12	93.6
	1	6	43	87.7
总体百分比				92.4

模型的准确率为 92.4%，有较强的预测能力。在预测客户未来信用情况时，该模型的结果与客户的信用等级相结合，将提供更具参考价值的洞察。因此，此模型在预测 A 公司企业客户未来是否会发生逾期行为方面，具有较高的实用性和准确性。

3.2 基于 BP 神经网络构建评价模型

3.2.1 模型参数说明

在指标构建章节的基础上，已经筛选了六个关键指标，这些指标对客户未来是否会发生逾期行为具有显著影响。本章节将利用这六个指标，构建一套全面而精准的信用评价模型。该模型不仅能够生成客户的信用等级，还能预测客户未来是否存在逾期风险。之所以同时提供这两个结果，是因为即使客户的信用等级很高，也不能百分之百地保证他们未来不会逾期。信用评价等级主要是基于客户过去的信用记录和行为进行预测和评估的，然而，过去的记录并不能完全代表未来的行为。因此，在给出信用等级的时候，

还提供了未来逾期风险的判断，旨在为 A 公司在设定客户信用额度和制定账期时提供更为全面、科学的参考依据。这样，A 公司可以更加精准地管理信用风险，优化资源配置，从而实现更为稳健和可持续的发展。输入输出字段详见表 12：

表 12 BP 神经网络输入输出参数

类型	参数	符号	参数类型
输入	现金比率	CASHR	连续变量
	应收账款周转率	ARTR	连续变量
	经营负荷情况	OLC	二分类 0-否 1-是
	是否有逾期记录	OverdueFlag	二分类 0-否 1-是
	是否上市	ListedFlag	二分类 0-否 1-是
	销售净利率	NPM	连续变量
输出	未来是否逾期	FutureOverdueFlag	二分类 0-否 1-是
	信用等级	CreditRating	多分类值为 1-5

在此 BP 神经网络模型中，输入层选取了六个关键指标，这些指标旨在全面捕捉客户信用的多个维度。而输出层，则聚焦于两个核心变量：“是否逾期”和“信用等级”。这两个变量对于评估客户的信用风险具有至关重要的作用。

“信用等级”这一关键输出，因为目前无法在各大评级机构中获取到 A 公司客户的信用等级。A 公司已经建立了一套完善的客户等级评价体系，该体系在每次信用申请时都会为客户进行详尽的测评。这套评价系统是由大公国际信用评级集团有限公司倾力打造的，确保了评价标准的专业性和权威性。从而得出具有参考价值的信用等级。

同时，在评级过程中，A 公司的多个关键部门，包括财务部门、信用管理部门以及法务部门等，都会进行逐层审核。这种跨部门的合作与审核机制，进一步提升了信用评级的准确性和可信度。因此，将这些经过严格审核的信用等级作为 BP 神经网络输出层的一部分，将为此模型提供强有力的支持，使其能够更加精准地预测客户的信用风险。

### 3.2.2 模型建立

BP 神经网络模型 BP 网络模型包括其输入输出模型、作用函数模型、误差计算模型和自学习模型。1. 模型网络结构设置

基于 BP 神经网络的信用评价构建流程图如下：

BP 神经网络是一种强大的多层前馈网络，其结构主要由输入层、隐藏层和输出层构成。每个神经元都与前一层的所有神经元通过权重连接，形成了密集的全连接模式。这种结构使得 BP 神经网络能够捕捉和表达复杂的非线性关系。

在构建 BP 神经网络时，网络的结构设计是至关重要的。它需要根据系统的输入输

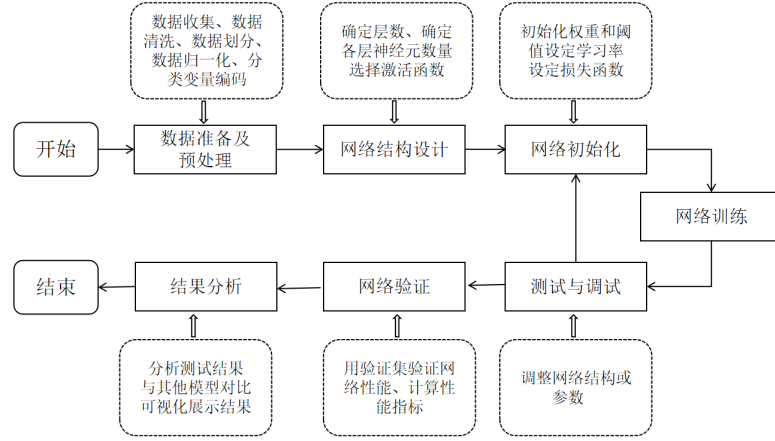


图 5 BP 神经网络构建过程图

出数据特点来确定。具体而言，输入层的节点数应当与输入特征的维度完全匹配，以确保每一个特征都能被网络有效地捕获。对于输出层，其节点数则通常根据问题的性质来设定：在二分类问题中，输出层通常有一个节点；而在多分类问题中，则会有多个节点。

隐藏层，其节点数和层数的选择则相对灵活，往往需要根据实际情况进行调整。隐藏层的存在为网络提供了“深度”，使其能够学习并表达更复杂的模式和关系。通过适当地增加隐藏层的节点数或层数，网络可以更好地拟合数据，但也可能导致过拟合的风险增加。因此，在选择隐藏层的结构时，需要权衡好模型的复杂度和泛化能力<sup>[2]</sup>。

在本模型中，输入层有 6 个指标，因此输入层应设置 6 个节点。而输出层由于有两个指标需要预测，故应设置 2 个节点。对于隐藏层，可以参考一些经验公式或理论来设定其节点数。根据公式 2.6，设定隐藏层的节点数为 10。因此，本模型的 BP 神经网络结构最终确定为 6-10-2。

假设中  $x_1, x_2, \dots, x_n$  是输入值，输出值为  $y_1, y_2, \dots, y_m$ ， $\omega_{ij}$  和  $a_j$  为输入层与隐含层之间的权值和阈值， $\omega_{jk}$  和  $b$  是隐含层与输出层间的权值和阈值。以下是各层输入输出的计算公式。

隐藏层的输出值计算公式如 3.6。

$$H_j = f\left(\sum_{i=1}^n x_i \omega_{ij} - a_j\right) \quad j = 1, 2, \dots, r \quad (3.6)$$

其  $H_j$  为隐含层输出值， $f$  是传递函数， $r$  是隐含层节点数，在该模型中的值为 10。

输出层输出值计算，根据隐含层的输出值  $H$ ，连接权值  $\omega_{jk}$  和阈值  $b_j$ ，计算出输出公式如下：

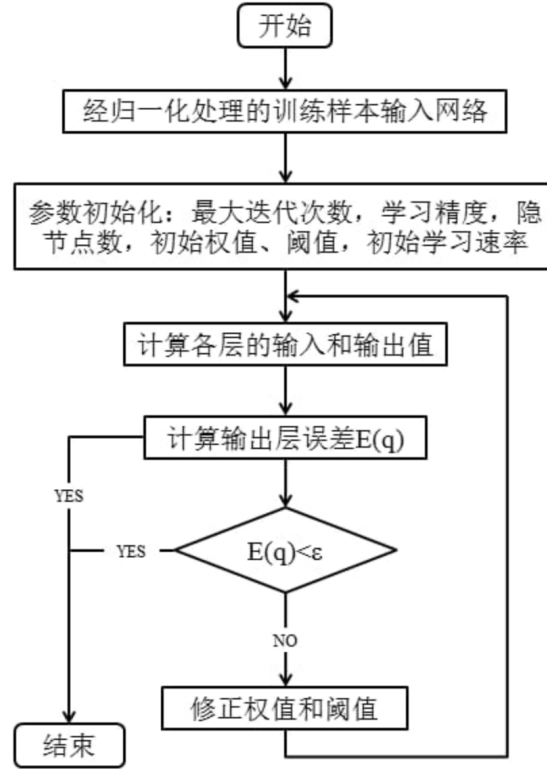


图 6 BP 神经网络训练流程图

$$O_k = f\left(\sum_{i=1}^r H_j \omega_{jk} - b_k\right) \quad k = 1, 2, \dots, m \quad (3.7)$$

其中  $O_k$  是输出层输出值， $m$  为输出层的节点个数，即 2。

## 2. 函数选取

图 6 是网络在训练过程中所需要计算的流程图，BP 神经网络在构建以及训练过程中，涉及多个关键的函数数，这些函数的选择会影响到模型建设效果。

BP 神经网络在构建过程中，涉及多种函数，主要涉及激活函数、学习函数、性能函数、初始化函数、传递函数。这些函数在神经网络的构建、训练和性能评估过程中起着关键作用。

激活函数是神经网络中非常关键的一个组件，它决定了神经元如何对其输入进行非线性变换。因为本研究的输出层两个指标，一个指标为是否逾期，该指标作为输出层是为二分类变量，选取 Sigmoid 函数，公式见 (2.1)，该函数常用于二分类问题的输出层，将数据映射到 0 和 1 之间；另一个指标是客户的信用等级，该指标作为输出层是多分类变量，可选择 Softmax 函数，公式见 (2.9)，该函数常用于多分类问题，输出层将数据映射到一个概率分布上。

学习函数定义了如何根据网络的误差来调整权重，本文采用梯度下降法来作为学习函数，梯度下降法是根据误差的梯度更新权重，是一种通用的优化算法，可以应用于多种类型的神经网络和机器学习模型，并且允许用户设置学习率，这是一个重要的超参数，可以控制权重更新的步长，通过调整学习率，可以平衡算法的收敛速度和稳定性。较小的学习率可能导致收敛速度较慢，而较大的学习率可能导致算法在最优解附近振荡。

针对多分类问题，通常会使用交叉熵损失函数作为误差函数，来评估模型的预测结果与真实标签之间的差异。公式见 (2.8)，作为模型的性能函数，本文研究的输出有二分类和多分类。不同的性能函数对应着不同的误差度量方式，如均方误差、交叉熵误差等。交叉熵误差是专门为分类问题设计的性能函数。对于多分类问题，交叉熵误差能够衡量预测概率分布与真实概率分布之间的差异，从而指导网络进行正确的分类。与均方误差等性能函数相比，交叉熵误差在训练初期不会导致梯度消失问题。这是因为交叉熵误差的梯度与激活函数的导数无关，从而避免了因激活函数饱和而导致的梯度消失。而且交叉熵误差的梯度计算简单且稳定，网络在训练过程中能够更快地收敛到最优解。

若没有达到预设的迭代次数或误差没有达到预设的最小值，需更新网络的权值和阈值，输入层与隐藏层权值  $\omega_{ij}$  更新公式见 (3.8)，隐藏层与输出层权值  $\omega_{jk}$  更新公式见 (3.9)，其中  $\eta$  为学习率。

$$\omega_{ij} = \omega_{ij} + \eta H_j(1 - H_j)x(i) \sum_{k=1}^m \omega_{jk}e_k \quad i = 1, 2, \dots, n; j = 1, 2, \dots, l \quad (3.8)$$

$$\omega_{jk} = \omega_{jk} + \eta H_j(1 - H_j)e_k \quad j = 1, 2, \dots, l; k = 1, 2, \dots, m \quad (3.9)$$

输入层与隐藏层的阈值  $a_j$  更新公式见 (3.10)，隐藏层与输出层阈值  $b_j$  更新公式见 (3.11)，其中  $\eta$  为学习率， $e_k$  为误差。

$$a_j = a_j + \eta H_j(1 - H_j) \sum_{k=1}^m \omega_{jk}e_k \quad j = 1, 2, \dots, l \quad (3.10)$$

$$b_k = b_k + e_k \quad k = 1, 2, \dots, m \quad (3.11)$$

### 3.3 本章小结

本章是模型构建章节，主要是说明网络的构建模型，具体的训练过程将放置第四章中进行。本章主要内容是基于广泛的文献研究、权威机构的报告以及 A 公司现有的信用评价系统，进行了评价指标的初步筛选。利用二元 Logistic 回归算法及 SPSS 工具，通过单因素及多因素相结合的方式，筛选出在财务、业务、企业资质方面的 6 个影响客户未来是否会逾期的关键指标，在单因素分析中，逐一考察了每个指标与客户是否逾期之间的关联性，通过统计检验确定每个指标对客户逾期行为的预测能力。在多因素分析中，综合考虑多个指标之间的相互作用，通过构建多元回归模型，识别出对客户逾期行为具有显著影响的指标组合。

基于之前分析筛选出的关键评价指标，进一步利用 BP 神经网络构建了信用评价模型，期望预测出客户未来的逾期情况及客户的信用等级。在模型构建过程中，设计了网络结构，确定输入层、隐藏层和输出层的神经元数量及连接方式。为了使网络能够快速、准确地学习并优化，使用梯度下降法等优化算法对网络进行训练，这些算法能够通过不断迭代，调整网络中的权重和阈值，使模型的预测结果与实际数据之间的误差逐渐减小。

第四章 实证研究

4.1 数据筛选与描述

本文所研究的信用评价模型，其核心在于采用了第 3.1 节中详细阐述的指标评价体系。这一体系经过精心构建和深入研究，旨在更全面地反映客户的信用状况。值得注意的是，这一新体系中包含的部分指标，在 A 公司原有的信用评价体系中并未涉及。在 A 公司原有的信用评价体系中，存在部分冗余指标。这些指标在评价客户的信用状况时，其贡献度和影响力相对较低，对整体信用评价结果的贡献并不显著。因此，这些冗余指标在信用评价过程中的存在意义不大，不仅增加了评价体系的复杂性，还可能干扰评价结果的准确性。原有评价体系虽然在一定程度上能够评估客户的信用水平，但受限于其指标选择的范围，可能无法全面捕捉客户的信用特征。而新加入的指标，不仅涵盖了更多维度的信息，还能够深入挖掘客户信用状况的内在规律和趋势。这些新增指标的存在，使得信用评价模型更加完善和科学，为 A 公司提供了更准确的客户信用评估依据。以下为各指标在 A 公司原有信用评价体系中的存在情况：

表 13 基于机器学习的 A 公司企业客户信用评价体系

指标	是否为 A 公司元评价体系指标
是否上市	否
经营负荷情况	否
现金比率	是
应收账款周转率	是
是否有逾期记录	是
销售净利率	是

本研究选取 A 公司 2019 年以来 A 公司所有进行过信用评价的数据，包含 A 公司所有国内与国外的所有客户数据。经过二元 Logistic 回归的筛选确定的六个指标，A 公司的国外客户同样能够提供这六个指标的数据，本文将这部分客户的数据也纳入 BP 神经网络的训练中。这一举措不仅丰富了数据集，使其更具多样性和代表性，而且有望进一步提升 BP 神经网络模型的准确性和泛化能力。通过融入国外客户的数据，能够更全面地捕捉不同市场环境下的信用特征，从而为 A 公司提供更为精准和可靠的信用风险评估服务。

收集到的初始数据集为 1843 条，经过去除重复及缺失数据后，共计 1042 条有效数据，通过对这些数据的深入分析和学习，期望构建的 BP 神经网络模型将能够更准确地



捕捉客户信用的内在规律，为 A 公司提供更为精准、可靠的信用风险评估和决策支持。本文的输入特征变量只有 6 个，而且这 6 个变量对客户的信用情况具有很高的解释性和预测能力，而且这些数据均为真实业务数据，使用 BP 神经网络能够捕捉到这些变量中的潜在规律和模式，实现对客户信用情况的准确预测，也能为模型提供有效的学习信息。将这些数据划分为训练集、测试集和验证集。其中，训练数据有 900 条，用于训练网络并调整其权重和偏置，寻找到最优参数；测试数据有 100 条，用于在训练过程中实时评估模型的性能；验证数据有 42 条，用于在第四章对模型进行最终的评估。这种数据划分方式有助于全面、客观地了解模型的性能，并为其后续的优化提供有力的支持。

## 4.2 数据预处理

在进行 BP 神经网络模型构建之前，数据预处理工作显得至关重要。为了确保神经网络训练的稳定性 and 高效性，需要对数据进行一系列预处理操作，包括标准化、归一化等。这些步骤能够消除不同特征之间的量纲差异，使得网络能够更容易地捕捉到数据中的内在规律。

在本文中，“是否逾期”和“信用等级”作为网络的输出层，都是分类变量。对于“是否逾期”这一变量，将此变量表示为 0 和 1，其中 0 代表未逾期，1 代表逾期。这种简洁的表示方法有助于神经网络更好地学习和预测逾期情况。

而对于“信用等级”这一输出层变量，本文沿用了 A 公司现有的评级制度。A 公司的信用等级分为五个层级，从高到低分别为 AAA、AA、A、B、C。为了方便神经网络处理，将这些等级转换为相应的数值表示，即 1、2、3、4、5。这种数值化的表示方法不仅保留了信用等级的原始信息，还使得神经网络能够更容易地处理和理解这一输出层变量。通过这样的预处理操作，可以为后续的 BP 神经网络训练奠定坚实的基础，提升模型的性能和准确率。等级设置规则和变量设置如下表：

表 14 信用等级划分及变量说明

等级	变量设置	等级含义说明
AAA	5	偿还债务的能力极强，基本不受不利经济环境的影响，违约风险极低。
AA	4	偿还债务的能力很强，受不利经济环境的影响不大，违约风险很低。
A	3	偿还债务能力较强，较易受不利经济环境的影响，违约风险较低。
B	2	偿还债务能力一般，受不利经济环境影响较大，违约风险一般。
C	1	偿还债务能力较弱，受不利经济环境影响很大，有较高违约风险。

现金比率、应收账款周转率、销售净利率为连续变量，需做归一化处理。其余三个

输入变量均为二分类变量，分类变量的输入数据是不需要进行归一化处理的。

对于连续变量，现金比率、应收账款周转率和销售净利率，为了确保这些变量在神经网络中能够被平等对待，并防止某些特征对模型训练产生过大的影响，需要对这些连续变量进行归一化处理。归一化可以将这些变量的数值范围缩放到一个共同的标准区间  $[0,1]$ ，从而消除量纲差异，使得网络能够更好地捕捉数据中的内在模式。

本文中的其余三个输入变量，情况则有所不同。这些分类变量通常表示某种属性或状态的存在与否，其取值往往是离散的，如 0 和 1。由于这些变量本身就不具有连续的数值范围，因此不需要进行归一化处理。在神经网络中，这些二分类变量可以直接以独热编码（one-hot encoding）或标签编码（label encoding）的形式输入，以便模型能够正确地处理和理解这些分类特征。通过这样的处理方式，可以确保不同类型的数据在 BP 神经网络模型中都得到适当的处理，从而提升模型的性能和准确性。

### 4.3 模型训练

学习率用于控制参数更新速度的超参数。它决定了权重调整的幅度，过大可能导致参数更新过快而错过最优解，过小可能导致收敛速度过慢。通常学习率的取值范围在 0.001 到 0.1 之间，本文设置的学习率为 0.01。

迭代次数的设置是网络在停止训练之前可以进行的最大迭代次数。迭代次数过多可能导致过拟合，过少可能导致训练不足。通常，初始时可以设置较大的迭代次数，然后根据训练收敛情况逐渐减小。本文经过多轮训练测试，设置的迭代次数为 1000。

误差阈值（通常称为“停止训练条件”或“误差目标”）是一个用于决定何时停止训练过程的参数。它代表网络训练过程中期望达到的误差水平。当网络的训练误差低于这个阈值时，训练过程将停止。本研究在实验过程中对误差阈值经过多轮的调试，初始时，设置一个相对较高的误差阈值 0.01。这样的设置可以为网络提供足够的训练时间，使其能够从数据中学到更多的信息。后面根据模型结果展示的趋势进行逐步减小误差阈值，最终模型设置的误差阈值为 0.001。

### 4.4 模型检验

BP 神经网络的训练过程是通过反向传播算法来不断调整网络中的权重和偏置，使得网络的输出尽可能接近真实值，从而实现模型的优化和学习。本研究将使用归一化处

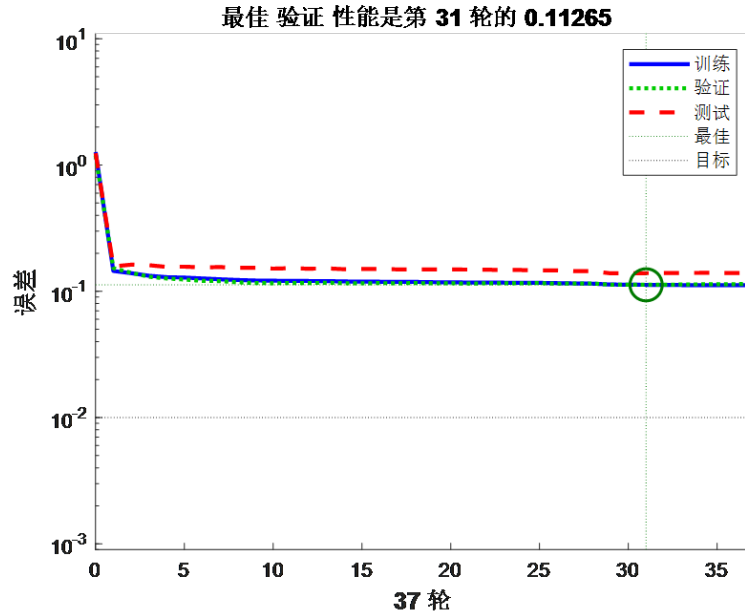


图 7 迭代轮次性能图

理后的 462 个样本数据来进行模型的训练，使用 MATLAB R2023 工具箱来进行训练。

#### 4.4.1 模型准确率分析

表 15 训练数据准确率

数据类型	输出	实际数量	预测正确数量
训练集	是否逾期	900	816
	信用等级	900	794
	正确率		89.4%
测试集	是否逾期	100	90
	信用等级	100	88
	正确率		85%
验证集	是否逾期	42	34
	信用等级	42	33
	正确率		79.7%

使用模型训练最终得出的各组数据的准确率如上表。其中型在不同数据集上的性能表现相对一致，但也有一些细微的差别。训练集准确率较高，说明模型在训练数据上学习得较好，能够识别并正确分类大部分样本。测试集准确率略低于训练集，但仍然是一个相对较高的值。这通常意味着模型具有较好的泛化能力，能够在未见过的数据上表现良好。验证集准确率略低于测试集和训练集，但仍然是一个可接受的性能水平。验证集通常用于模型选择和超参数调整，其准确率有助于确定模型在未知数据上的预期性能。综合来看，模型的性能表现相对较好。

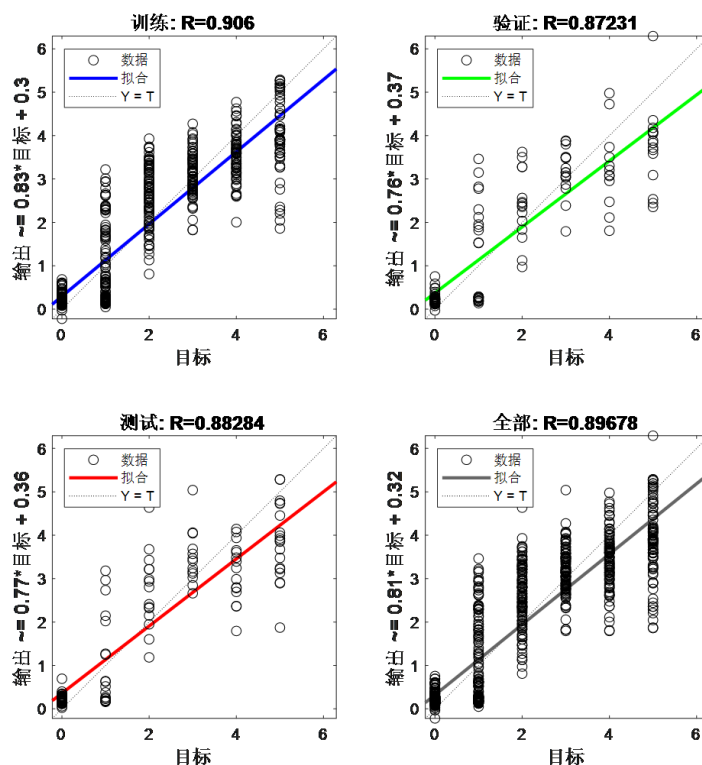


图 8 各组数据 R 值图

#### 4.4.2 模型性能分析

模型在训练过程中会逐渐调整其参数，以最小化损失函数。该在第 31 轮次后达到最佳性能，如图 7，说明模型在此之前一直在进行有效的学习，并且没有出现明显的过拟合或欠拟合现象。这也暗示了训练策略的有效性，包括学习率的选择、优化算法的使用以及正则化技术的应用等，都对模型的最终性能起到了积极作用。图中展示的损失误差是一个相对较低的数值，模型的输出为分类输出，此数值是可以接受的范围。模型在预测时的误差较小。这通常意味着模型能够准确地捕捉数据的内在规律和模式，从而做出准确的预测。损失误差的绝对值并不是唯一衡量模型性能的指标。还需要结合其他评估指标，如准确率，等来综合判断模型的性能。

#### 4.4.3 模型 R 值分析

图 8 为模型中的回归图形，在分类问题中，回归  $r$  值本身并不直接具有解释性。但是回归分析可以帮助在训练模型时更好地理解数据的分布和特征之间的关系，在分类问题中，回归的  $r$  值，即相关系数，通常用于衡量两个变量之间的线性关系强度。然而，它本身并不直接对分类任务的性能或准确性提供明确的解释。分类问题关注的是将数

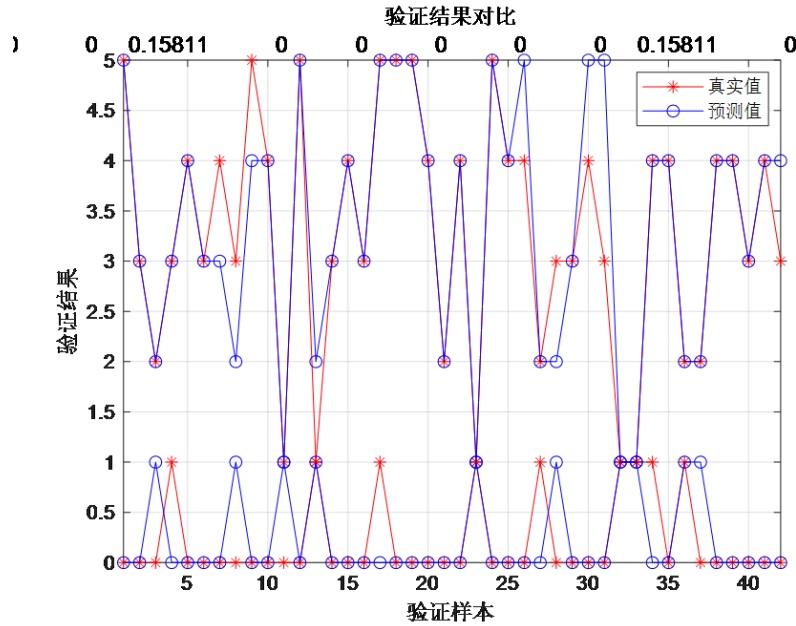


图 9 验证结果结果对比图

据点划分为不同的类别，而  $r$  值更多地反映了连续变量之间的线性趋势。尽管如此，回归分析在分类问题的训练过程中仍然扮演着重要的角色。通过回归分析，可以更好地理解数据的分布和特征之间的关系，从而有助于构建更有效的分类模型。

该模型的训练  $R$  值为 0.906：这表明在训练数据集上，模型的预测值与实际值之间具有很高的线性相关性。这可能意味着模型在训练数据上表现良好，能够很好地拟合训练数据的特征与目标变量之间的关系。测试  $R$  值为 0.88：测试集是用来评估模型在未见过的数据上的泛化能力的。测试  $R$  值略低于训练  $R$  值，但仍然是一个相对较高的值，说明模型在测试数据上的预测性能也较好。验证  $R$  值为 0.87：验证集用于调整模型的超参数，以避免过拟合。验证  $R$  值与测试  $R$  值相近，说明模型在验证数据上的性能也比较稳定。模型总  $R$  值为 0.89：这个值可能是训练、测试和验证  $R$  值的某种加权平均，它提供了一个整体性能的概览。0.89 的  $R$  值表明模型在整体上具有较好的预测能力。

#### 4.4.4 验证集结果分析

验证数据共有 42 条记录，详见表 16 中，这些验证数据可以客观地评估模型在真实环境中的表现，可以作为验证和校准模型的重要依据。

可以结合图 9，可以观察到，模型在预测企业未来是否会逾期方面的准确率，相较于信用等级的预测，展现出了更高的精准度。这一优势得益于输入参数与逾期风险的强关联度，这些参数在经过 Logistic 的精细筛选后，显现出了高度的相关性。反观信用评价的预测，虽然它是基于 A 公司原有指标体系计算得出的等级，但其与实际情况的关

表 16 信用评价模型验证数据

序号	是否有逾期记录	multirow2* 是否上市	经营 负荷情况	现金比率	应收账款 周转率	销售 净利率	是否逾期 -真实值	信用等级 -真实值	是否逾期 -预测值	信用等级 -预测值
1	是	否	否	1.21	6.64	0.0662	否	A	否	A
2	是	否	是	0.22	0.95	0.2426	否	C	否	C
3	否	否	是	0.9	2.46	0.084	否	B	是	B
4	否	否	是	0.24	7.08	0.1078	否	AA	否	AA
5	否	是	是	1.37	7.82	0.2319	否	AA	否	AA
6	是	是	是	0.09	5.13	0.0098	是	B	是	B
7	是	否	是	0.89	10.32	0.0504	否	A	否	A
8	是	否	是	1.4	32.42	0.2361	否	AA	否	AAA
9	否	否	是	0.19	6.06	0.0339	否	B	是	B
10	是	是	是	0.39	12.31	0.0499	是	A	否	A
11	否	否	是	0.24	26.77	0.1089	否	AA	否	AAA
12	否	否	是	4.34	8.74	0.0144	否	A	否	AAA
13	否	否	是	0.21	0.75	0.0009	否	C	是	C
14	否	否	是	0.4	8.76	0.0334	否	A	否	A
15	否	否	是	0.54	62.34	0.0914	是	AAA	否	AAA
16	否	否	是	0.26	1.52	0.2105	否	C	否	C
17	否	是	是	0.28	17.26	0.2588	否	AA	否	AA
18	否	否	否	1.06	9.26	0.1799	否	AA	否	AA
19	否	否	是	0.16	1.81	0.2886	否	A	是	B
20	否	否	否	1.19	8.05	0.0905	否	A	否	AA
21	否	是	是	0.2	10.53	0.0922	否	A	否	A
22	是	否	是	1.08	17.53	0.1197	否	AA	否	AA
23	否	是	是	0.75	12.72	0.151	否	AAA	否	AAA
24	是	否	是	0.02	1.93	0.0059	是	C	是	B
25	否	否	是	0.22	5.36	0.0717	否	B	是	B
26	否	是	是	0.31	4.22	0.028	是	AA	否	AA
27	否	否	是	0.19	3.63	0.0303	否	A	否	B
28	否	是	否	0.07	11.33	0.1709	否	AA	否	AA
29	否	否	否	0.3	2.43	0.1045	是	B	否	B
30	否	是	是	0.95	11.28	0.2124	否	AA	否	AA
31	否	是	是	0.58	41.94	0.1814	否	AAA	否	AAA
32	是	否	是	0.19	13.34	0.2941	否	AA	否	AA
33	否	是	是	0.47	5.35	0.6895	否	AAA	否	AAA
34	是	否	是	0.63	5.3	0.0063	是	C	是	C
35	否	是	是	0.56	61.94	0.1214	否	AAA	否	AA
36	否	否	是	1.89	5.75	0.3916	否	AA	否	A
37	否	否	是	0.31	3.76	0.1686	否	A	否	A
38	否	是	是	1.25	21.63	0.0926	否	AAA	否	AAA
39	是	否	是	0.51	11.83	0.0743	否	AA	否	AA
40	是	否	否	0.69	13.31	0.2252	否	A	否	A
41	否	是	是	0.52	5.58	0.5486	否	AAA	否	AAA
42	否	否	是	0.4	9.95	0.027	否	AA	否	AA

联度相对较低。尽管如此，模型所预测出的信用等级结果，却更能真实地反映企业的还款能力。图 9 中明显呈现出一种趋势：当模型预测某企业存在逾期风险时，其信用等级预测往往低于 A 公司原有的评级，这进一步印证了模型在揭示企业未来还款情况方面的有效性。从图 9 中可以看出，模型在信用等级预测上准确判断了 33 个样本，准确率达到 78.5%。而在未来是否逾期的预测上，模型更是精准地判断了 34 个样本，准确率高达 80.95%。尽管数据量相对有限，但这一结果足以证明模型在信用等级预测方面具备一定的预测能力。

#### 4.4.5 实证分析结论

信用等级预测方面，模型在信用等级预测方面表现出了较高的准确性。通过对比测试集及验证集上的预测结果与实际信用等级，本研究发现模型能够准确地区分不同信用等级的客户，为业务提供有力的决策支持。

在预测客户未来是否会逾期方面，模型同样展现出了良好的性能。通过对客户的历史还款记录和其他相关指标的学习，模型能够有效地识别出具有逾期风险的客户，为业务度对客户的风险管理提供了重要依据。该模型在信用等级预测和未来是否逾期预测方面表现良好。虽然两个预测结果的准确率都较高，但未来是否逾期的预测准确率略高于信用等级预测。这可能是因为逾期与否的预测与实际的还款行为更为直接相关，模型在捕捉这种直接关联上可能更具优势。

虽然该模型在当前的数据集上表现出了较好的性能，但还需要注意，模型的性能可能会受到多种因素的影响，如数据集的分布、指标的选择、模型的参数设置等。因此，在未来的研究中，可以进一步探索如何优化这些因素，以提高模型的性能。

## 4.5 本章小结

本章基于第三章构建的信用评价模型，利用 A 公司的企业客户数据进行了进行了模型训练及检验。首先将收集到的数据进行处理，利用处理后的数据对模型进行了反复的训练。通过不断调整模型的参数，努力寻找最优的模型配置，以最大程度地提高模型的预测准确性。在训练过程中，采用了交叉验证等方法，对模型进行了充分的验证和评估，确保了模型的稳定性和泛化能力。完成训练后，使用验证数据将模型的预测结果与实际数据进行了对比和分析。通过计算模型的准确率、拟合度等关键指标，全面评估了模型的预测性能。

## 结 论

信用等级是对企业信用能力和还款意愿的综合评估，等级高低可以反映企业的信用状况和履约能力。信用等级高并不意味着企业客户一定不会逾期。尽管信用等级高的企业客户在信用评估中表现出色，具有较低的信用风险和良好的还款能力，但这并不意味着他们完全没有违约或逾期的可能性。在实际业务中，各种不可预见的情况和突发事件可能导致企业客户无法按时偿还债务。因此，高信用等级并不绝对保证企业不会出现逾期的情况。

### 1. 研究结论

对于以上问题，本研究在构建信用评价模型时，分别在信用等级和未来是否会逾期这两个方面进行了预测。研究表明：基于经营负荷情况、现金比率、应收账款周转率、是否有逾期记录、是否上市以及销售净利率这六个指标构建的 BP 神经网络方法可以有效预测企业客户未来是否会发生逾期，使用 Matlab 编程语言，借助计算机工具可以方便地完成 BP 网络模型的算法设计以及数据运算，建立起预测企业客户信用的模型，更加准确地评估客户的信用风险，为公司的业务决策和风险管理提供了有力支持。

### 2. 不足及改进方向

在数据量方面，目前本研究数据来源于 A 公司的企业客户，收集的数据量较为有限，对于构建和验证 BP 神经网络模型而言，这样的数据量确实存在一定的局限性。在数据量不够充分的情况下，神经网络可能无法充分捕捉到数据的内在规律和特征，从而导致模型的泛化能力受限，预测结果不够精准。

模型输出中使用的信用等级数据来源具有局限性。信用等级采用 A 公司信用评价系统得出的结果。该信用等级仅反映了 A 公司内部的业务逻辑和评价标准，而这些标准和逻辑可能并不适用于其他公司或行业。使用特定公司的业务数据作为研究基础，其代表性和普遍性难免受到质疑。A 公司的业务模式和评价标准可能与其他公司存在显著差异，因此其数据可能无法全面反映整个行业或市场的真实状况。这使得研究所得出的评价模型在推广和应用时存在一定的局限性，难以直接应用于其他公司或行业。

针对以上不足可以结合行业内的其他公司或引入其他信用评价机构的信用评价数据，以获取更广泛、更全面的信用信息。通过扩大数据集的规模，为 BP 神经网络提供



更多的学习样本，从而提高模型的泛化能力。

## 参考文献

- [1] 霍思彤, 郑新毅. 供应链服务业务中企业信用评级方法研究[J]. 一重技术, 2022: 69–72
- [2] Boughaci D, Alkhawaldeh A A K, Jaber J J, et al. Classification with segmentation for credit scoring and bankruptcy prediction[J]. Empirical Economics, 2021, 61
- [3] Baser F, Koc O, Selcuk-Kestel A S. Credit risk evaluation using clustering based fuzzy classification method[J/OL]. Expert Systems with Applications, 2023: 119882. <http://dx.doi.org/10.1016/j.eswa.2023.119882>
- [4] Golbayani P, Florescu I, Chatterjee R. A comparative study of forecasting corporate credit ratings using neural networks, support vector machines, and decision trees[J]. RePEc: Research Papers in Economics - RePEc, RePEc: Research Papers in Economics - RePEc, 2020
- [5] Chen Y, Calabrese R, Martin-Barragan B. Interpretable machine learning for imbalanced credit scoring datasets[J/OL]. European Journal of Operational Research, 2024: 357–372. <http://dx.doi.org/10.1016/j.ejor.2023.06.036>
- [6] 宋永明, 胡君. 基于 topsis 方法的中小企业信用风险评价——以山东省制造业中小企业为例[J]. 现代商业, 2021: 142–144. DOI: [10.14097/j.cnki.5392/2021.05.045](https://doi.org/10.14097/j.cnki.5392/2021.05.045)
- [7] 李晓寒. 信用评级对非金融企业资本结构的影响——基于中国企业的表现[J]. 统计理论与实践, 2021: 58–66
- [8] 罗勇, 陈治亚. 基于模糊综合法构建供应链金融客户信用评价指标体系[J]. 财会月刊, 2015: 43–46. DOI: [10.19641/j.cnki.42-1290/f.2015.14.008](https://doi.org/10.19641/j.cnki.42-1290/f.2015.14.008)
- [9] Saha S, Saha M, Mukherjee K, et al. Predicting the deforestation probability using the binary logistic regression, random forest, ensemble rotational forest, reptime: A case study at the gumani river basin, india[J/OL]. Science of The Total Environment, 2020, 730: 139197. <http://dx.doi.org/10.1016/j.scitotenv.2020.139197>
- [10] 阳彩霞, 阳平华, 何杰. 基于 bp 神经网络的供应商信用评价分析——以造价通为例[J]. 轻工科技, 2022, 38: 50–52

- [11] Bao W, Lianju N, Yue K. Integration of unsupervised and supervised machine learning algorithms for credit risk assessment[J/OL]. *Expert Systems with Applications*, 2019, 128: 301–315. <http://dx.doi.org/10.1016/j.eswa.2019.02.033>
- [12] Baser F, Koc O, Selcuk-Kestel A S. Credit risk evaluation using clustering based fuzzy classification method[J/OL]. *Expert Systems with Applications*, 2023: 119882. <http://dx.doi.org/10.1016/j.eswa.2023.119882>
- [13] Sun J, Li J, Fujita H. Multi-class imbalanced enterprise credit evaluation based on asymmetric bagging combined with light gradient boosting machine[J/OL]. *Applied Soft Computing*, 2022, 130: 109637. <http://dx.doi.org/10.1016/j.asoc.2022.109637>
- [14] Serrano-Cinca C, Gutiérrez-Nieto B. The use of profit scoring as an alternative to credit scoring systems in peer-to-peer (p2p) lending[J]. *Decision Support Systems*, 2016, 89(C): 113–122
- [15] 徐德顺, 马凡慧, 寇明珠. 基于三重积分原理的企业三维信用评价模型研究[J]. *征信*, 2021, 39: 22–28
- [16] 姚定俊, 顾越, 陈威. 基于 rf-lsma-svm 模型的中小微企业信用风险评价研究[J]. *工业技术经济*, 2023, 42: 85–94
- [17] 周颖, 张志鹏. 基于违约企业分布约束的上市公司信用等级划分模型[J]. *管理评论*, 2023, 35: 3–18. DOI: [10.14120/j.cnki.cn11-5057/f.2023.05.011](https://doi.org/10.14120/j.cnki.cn11-5057/f.2023.05.011)
- [18] Harris T. Quantitative credit risk assessment using support vector machines: Broad versus narrow default definitions[J]. *Expert Systems with Applications*, 2013, 40(11): 4404–4413
- [19] Mileris R. Macroeconomic determinants of loan portfolio credit risk in banks[J]. *Inzinerine Ekonomika Engineering Economics*, 2012, 23(5): 496–504
- [20] 华泰证券固定收益部课题组, 王磊. 基于大数据和人工智能技术的数据治理与信用评价模型构建[J]. *金融纵横*, 2023: 51–59
- [21] 董秉坤, 郑陈柔雨, 原源, 等. 基于 ahp-熵值组合赋权法的卷烟零售户信用评价模型研究[J]. *中国烟草学报*, 2023, 29: 104–111. DOI: [10.16472/j.chinatobacco.2022.010](https://doi.org/10.16472/j.chinatobacco.2022.010)
- [22] Kumar Roy P, Shaw K, Ishizaka A. Developing an integrated fuzzy credit rating system for smes using fuzzy-bwm and fuzzy-topsis-sort-c[J/OL]. *Annals of Operations Research*,

- 2023: 1197–1229. <http://dx.doi.org/10.1007/s10479-022-04704-5>
- [23] 王灿. 基于不平衡数据与集成学习算法的信用评价模型[J]. 中国新技术新产品, 2022: 10–12. DOI: [10.13612/j.cnki.cntp.2022.18.034](https://doi.org/10.13612/j.cnki.cntp.2022.18.034)
- [24] 任向英. 基于灰聚类的中小企业信用等级评价及应用探究[J]. 佳木斯大学学报 (自然科学版), 2022, 40: 156–159
- [25] 刘英杰, 李光辉, 马锐鑫. 基于改进灰色关联模型的工程监理企业信用评价等级[J]. 项目管理技术, 2022, 20: 22–27
- [26] Chen Y S, Cheng C H. Hybrid models based on rough set classifiers for setting credit rating decision rules in the global banking industry[J/OL]. Knowledge-Based Systems, 2013, 39: 224–239. <http://dx.doi.org/10.1016/j.knosys.2012.11.004>
- [27] Duan J C, Li S. Pd-implied ratings via referencing a credit rating/scoring pool' s default experience[M/OL]. 2021: 105–115. [http://dx.doi.org/10.1007/978-3-030-49728-6\\_6](http://dx.doi.org/10.1007/978-3-030-49728-6_6)
- [28] 刘翠玲, 胡聪, 王鹏, 等. 基于营销大数据的电力客户多维度信用评价模型研究[J]. 西南大学学报 (自然科学版), 2022, 44: 198–208. DOI: [10.13718/j.cnki.xdzk.2022.06.020](https://doi.org/10.13718/j.cnki.xdzk.2022.06.020)
- [29] 杨莲, 石宝峰, 董轶哲. 基于 class balanced loss 修正交叉熵的非均衡样本信用风险评价模型[J]. 系统管理学报, 2022, 31: 255–269+289
- [30] 郭鑫. 基于层次分析法的科研信用评价模型优化[J]. 自动化技术与应用, 2022, 41: 169–172
- [31] 范柏乃朱文斌. 中小企业信用评价指标的理论遴选与实证分析[J]. 科研管理, 2003 (06): 83–88. DOI: [10.19571/j.cnki.1000-2995.2003.06.014](https://doi.org/10.19571/j.cnki.1000-2995.2003.06.014)
- [32] 曹轲, 谭冲, 刘洪, 等. 基于改进灰狼算法优化 bp 神经网络的无线传感器网络数据融合算法[J]. 中国科学院大学学报, 2022, 39(02): 232–239
- [33] 郑晓杰, 王双. M 商业银行对中小型企业客户信用风险评价体系研究[J]. 现代商业, 2019: 128–129. DOI: [10.14097/j.cnki.5392/2019.32.055](https://doi.org/10.14097/j.cnki.5392/2019.32.055)
- [34] Shen F, Zhao X, Li Z, et al. A novel ensemble classification model based on neural networks and a classifier optimisation technique for imbalanced credit risk evaluation [J/OL]. Physica A: Statistical Mechanics and its Applications, 2019: 121073. <http://dx.doi.org/10.1016/j.physa.2019.121073>
- [35] 程砚秋, 徐占东. 基于泰尔指数修正的 electre III 小企业信用评价模型[J]. 中国管理

- 科学, 2019, 27: 22–33. DOI: [10.16381/j.cnki.issn1003-207x.2019.10.003](https://doi.org/10.16381/j.cnki.issn1003-207x.2019.10.003)
- [36] Roy P K, Shaw K. A credit scoring model for smes using ahp and topsis[J/OL]. International Journal of Finance & Economics, 2023: 372 – 391. <http://dx.doi.org/10.1002/ijfe.2425>
- [37] 赵志冲, 迟国泰, 潘明道. 基于信用差异度最大的信用等级划分优化方法[J]. 系统工程理论与实践, 2017, 37: 2539–2554
- [38] Bhatore S, Mohan L, Reddy Y R. Machine learning techniques for credit risk evaluation: a systematic literature review[J/OL]. Journal of Banking and Financial Technology, 2020, 4(1): 111–138. <http://dx.doi.org/10.1007/s42786-020-00020-3>
- [39] Shen F, Zhao X, Kou G, et al. A new deep learning ensemble credit risk evaluation model with an improved synthetic minority oversampling technique[J/OL]. Applied Soft Computing, 2021: 106852. <http://dx.doi.org/10.1016/j.asoc.2020.106852>
- [40] Zhang K, Zhu J, He M, et al. Research on intelligent comprehensive evaluation of coal seam impact risk based on bp neural network model[J/OL]. Energies, 2022, 15(9): 3292. <http://dx.doi.org/10.3390/en15093292>
- [41] 于立勇, 詹捷辉. 基于 logistic 回归分析的违约概率预测研究[J]. 财经研究, 2004(09): 15–23. DOI: [10.16538/j.cnki.jfe.2004.09.002](https://doi.org/10.16538/j.cnki.jfe.2004.09.002)

## 致 谢

两年的读研究生涯，一转眼就过去了，还记得报到时的场景，仿佛就在不久前，就这样一瞬间，仿佛时空穿梭一般，到了说再见的时刻。这两年收获良多，收获了成长，收获了友情。在我的学习和论文写作中，有许多人给予了我悉心的指导和帮助，在此我要向他们表示最诚挚的感谢。

我要衷心感谢我的指导老师张军欢教授。张教授在整个论文写作过程中给予了我耐心细致的指导和无私的帮助，在我遇到模型数据的采集和分析过程中出现的难题，给予了我很多的帮助和支持，使得我的研究工作能够顺利进行，他的悉心指导让我更深入地了解了管理学的理论知识，提升了我的分析和解决问题的能力，真正领略到了学术研究的乐趣。

我要衷心感谢我的同事们。在我论文写作的过程中遇到的业务问题，给予了我耐心的解答和帮助，使我能够更好地理解和分析相关业务知识。他们的专业知识和帮助对我完成论文起到了关键作用。

最后，我要感谢我的家人，他们做我最坚实的后盾，让我在读书的过程中无后顾之忧。在我需要心理安慰时，他们始终在我身边，给予我足够的鼓励和关怀，让我从困扰的情绪中及时解脱。正是他们的支持和鼓励，才让我能顺利读研和毕业。

由衷地感谢以上所有给予我帮助和支持的人，正是有了你们的支持和帮助，我才能够顺利完成我的 MBA 学习和论文写作。再次向你们表示最诚挚的感谢！