

Research on credit risk identification and evaluation model for commercial bank group customers

Jiaqi Li

Ningxia Bank Co., Ltd., Yinchuan, Ningxia, 750299, China

Abstract

Against the backdrop of strengthened financial regulation and economic restructuring, commercial banks face complex challenges in managing credit risks for corporate clients. These clients exhibit concentrated capital flows and frequent related-party transactions, with risks characterized by concealment and systemic nature. Building on risk transmission theory and multidimensional data analysis, this study develops a credit risk identification and assessment model integrating Analytic Hierarchy Process (AHP), logistic regression, and random forest algorithms. The model comprehensively quantifies financial, structural, and market risks for corporate clients. Empirical results demonstrate that the model improves risk identification accuracy by approximately 22%, significantly enhancing early warning capabilities for potential default clients. The research recommends establishing a dynamic monitoring and tiered evaluation system supported by big data and machine learning, providing decision-making foundations for commercial banks to scientifically prevent and control corporate client credit risks.

Keywords

commercial bank; group customer; credit risk; risk identification; evaluation model

商业银行集团客户信用风险识别与评估模型研究

李嘉琪

宁夏银行股份有限公司, 中国·宁夏 银川 750299

摘要

在金融监管强化与经济结构转型的背景下, 商业银行集团客户信用风险管理面临复杂挑战。集团客户资金集中、关联交易频繁, 风险具有隐蔽性与系统性。本文基于风险传导理论与多维数据分析, 构建融合层次分析法(AHP)、逻辑回归与随机森林算法的信用风险识别与评估模型, 对集团客户的财务、结构与市场风险进行综合量化。实证结果显示, 模型可将风险识别准确率提高约22%, 显著增强潜在违约客户的预警能力。研究建议以大数据与机器学习为支撑, 建立动态监测与分级评估体系, 为商业银行科学防控集团客户信用风险提供决策依据。

关键词

商业银行; 集团客户; 信用风险; 风险识别; 评估模型

1 引言

集团客户是商业银行信贷资产的重要组成部分, 其业务规模大、结构复杂、内部资金往来频繁, 一旦发生信用风险, 极易引发系统性金融波动。近年来, 我国多起大型企业集团的债务违约事件暴露了传统信用管理模式的局限性。集团客户的信用风险不仅来源于母公司经营状况, 还受制于子公司偿债能力、交叉担保链条以及产业周期波动。如何通过科学建模手段识别风险信号、建立分层评估体系, 成为商业银行风险管理转型的关键命题。目前, 国内外研究多聚焦于单一客户信用评级或行业风险评估, 而对集团客户的整体风险传导与结构识别研究相对不足。传统模型以财务指标为核

心, 缺乏对集团内部关联交易、隐性担保及外部市场风险的量化描述。本文在风险系统性视角下, 提出基于多源数据融合与智能算法的信用风险识别与评估模型, 以期提升商业银行集团客户风险管理的科学性、前瞻性与可操作性。

2 集团客户信用风险的特征与识别逻辑

2.1 集团客户信用风险的内涵与特征

集团客户信用风险是指企业集团整体及其成员单位在履约过程中发生违约或损失的可能性。与一般单体客户不同, 集团客户的风险具有三大特征: 一是传导性强, 集团内部资金与担保链结构复杂, 一家子公司风险可快速扩散至母公司及其他关联企业; 二是隐蔽性高, 部分风险通过股权投资、关联交易、表外融资等方式潜伏在财务报表之外; 三是系统性突出, 集团企业往往涉足多个行业领域, 宏观经济波动或政策调整易引发多层次风险叠加。

【作者简介】李嘉琪(1994-), 女, 中国河北沧州人, 硕士, 从事金融统计学研究。

2.2 风险识别的关键维度

集团客户的信用风险识别应覆盖财务、治理、产业、市场与外部环境五个维度。财务风险关注偿债能力、盈利质量与现金流稳定性；治理风险侧重集团结构透明度、内部控制有效性与决策集中度；产业风险体现在行业周期性与上下游依存度；市场风险则受外部利率、汇率及大宗商品价格波动影响。外部环境风险包括政策变化、信用评级调整与地缘政治因素。多维度风险指标的综合分析，是识别集团信用风险的基础。

2.3 风险传导路径与系统模型逻辑

集团信用风险的形成可分为“内部传导—外部放大—系统扩散”三个阶段。内部传导源于成员企业之间的债务联动与担保链条；外部放大发生在市场与金融环境的应激反应中；系统扩散则表现为集团整体融资能力与偿债预期下降。基于这一逻辑，本文建立风险识别模型框架，以“核心企业—成员企业—外部环境”三层嵌套结构刻画风险传播机制，实现对信用风险的动态监测与分级评估。

3 集团客户信用风险指标体系构建

3.1 财务风险指标层设计

财务风险是商业银行评估集团客户信用水平的核心要素，其反映企业偿债能力与经营稳定性的综合状况。本文从偿债能力、盈利能力、营运效率与现金流稳定性四个维度构建财务风险指标层，以全面捕捉企业财务健康状况的变化特征。在偿债能力方面，选取资产负债率、流动比率与利息保障倍数等指标，衡量企业资金杠杆与短期偿债压力；盈利能力维度中纳入营业利润率、净资产收益率（ROE）与息税折旧摊销前利润率（EBITDA Margin），用于反映盈利的质量与可持续性；营运效率通过存货周转率与应收账款周转天数等指标衡量资产运营效率与流动性水平；现金流稳定性则以经营活动现金净额占比、现金流与债务比率等指标表征企业现金流安全垫。为避免行业差异造成评估偏差，采用极差标准化与 Z-score 方法对各指标进行归一化处理，并通过时间序列平滑技术剔除季节性波动，使指标体系具备跨行业可比性与长期稳定性，为风险模型提供坚实的定量基础。

3.2 结构与治理风险指标层

治理结构与内部控制水平是识别集团客户信用风险的重要非财务维度，直接影响风险传导路径与防控能力。本文构建结构与治理风险指标层，从股权集中度、交叉持股率、内部交易依存度与资金池集中度四个方面量化企业内部控制强度。股权集中度反映企业权力结构的合理性，过度集中可能导致决策失衡，而过度分散又会削弱治理效率；交叉持股率与内部交易依存度揭示集团内部资金往来与资源配置风险，若内部交易比重过高，将削弱独立经营能力并放大系统性风险；资金池集中度用于衡量集团资金流动的集中与透明程度，是判断集团资金安全的关键指标。除上述定量指标

外，还引入信息披露透明度、审计意见类型与董事会独立性等定性指标，用以反映管理层的诚信度与风险偏好。通过模糊综合评价法，将定性与定量指标有机融合，使治理风险评估结果更具解释性与可操作性。

3.3 外部风险指标层设计

外部风险指标层旨在揭示宏观经济、产业周期与市场环境对集团信用状况的外部冲击效应。该层主要包括行业景气指数、产业集中度、区域政策风险指数及利率敏感度四类指标，用以评估集团所在行业与区域的系统性风险水平。行业景气指数反映行业整体发展趋势，产业集中度揭示市场竞争格局及抗风险能力；区域政策风险指数衡量地方财政状况、产业扶持政策及监管环境变化的影响。针对能源、制造与基建类集团，模型进一步引入原材料价格波动率、外汇敞口比重及国际市场依存度等指标，以捕捉周期性与国际化风险的叠加效应。为确保外部风险测度的前瞻性，采用基于大数据爬取的动态指数更新机制，将宏观经济指标与企业信用数据实时对接，实现对外部环境变化的敏感响应。该指标层的构建使模型能够从系统性与情境性双视角识别风险，提升信用评估的宏观适应性与预测精度。

4 信用风险识别与评估模型的构建

4.1 模型总体思路与方法选取

本文在构建集团客户信用风险评估模型时，综合运用统计建模与机器学习的混合策略，以实现风险识别的精确化与动态化。基于已构建的多维风险指标体系，首先采用层次分析法（AHP）确定各类指标的权重分布，建立风险权重矩阵，计算综合风险指数，为后续模型训练提供数据输入。模型选用逻辑回归与随机森林两种算法进行联合建模：逻辑回归模型用于揭示风险指标与违约概率的线性关系，明确各变量的边际影响；随机森林算法则在多特征交互条件下进行非线性识别，能够有效捕捉复杂的风险组合特征。通过双模型融合方式，将逻辑回归的可解释性与随机森林的预测精度结合，实现对集团客户信用风险的多维度识别与分层评估。该混合建模框架既具备统计稳定性，又能在动态环境下保持高识别灵敏度，为商业银行实现数据驱动的信用决策提供技术支撑。

4.2 模型数据来源与样本选取

为确保模型具有代表性与可推广性，本文选取某大型商业银行 2020—2024 年间的集团客户信贷数据作为研究样本，共计 420 组有效数据，覆盖制造业、能源业、基建业及信息技术等四大行业。数据来源包括企业年报、财务报表、央行征信系统、银企授信记录及外部评级信息，确保了多源数据的完整性与客观性。在数据预处理环节，采用 Winsorize 法对异常值进行截尾处理，以减弱极端数据对模型稳定性的影响；同时利用主成分分析法（PCA）提取主要特征因子，降低多重共线性带来的干扰，提高模型运行

效率。样本划分方面,随机抽取70%数据用于模型训练,30%用于验证与测试,确保模型在不同数据集上的适应性与稳定性。

4.3 模型验证与评价指标

为检验模型的预测能力与稳健性,本文采用K-fold交叉验证法进行模型可靠性评估,并以AUC值与KS统计量作为主要性能指标。验证结果显示,综合模型的AUC值达到0.912,明显高于单一逻辑回归模型的0.845和随机森林独立模型的0.887,说明混合模型在区分违约与非违约样本方面表现更优。KS值为0.58,表明模型具有较强的区分能力。通过混淆矩阵分析发现,模型对潜在违约客户的识别准确率提高约22%,风险漏报率下降18%,显著增强了预警的灵敏度与风险筛查精度。此外,模型在不同时间区间的滚动验证中表现稳定,说明其具有良好的可迁移性与鲁棒性。研究结果证明,AHP-逻辑回归-随机森林混合模型能有效捕捉多维风险特征,实现对集团客户信用风险的科学量化与动态评估,为商业银行构建精细化、智能化的风险管理体系提供了可靠依据。

5 模型应用效能与优化路径研究

5.1 模型应用效果分析

模型在样本企业中的验证结果显示,构建的综合信用风险指数与集团客户的实际违约率呈显著正相关(相关系数0.86),证明其能较准确地刻画企业信用状况。通过模型输出的分层评级结果,商业银行可实现“风险分级、授信分层、资源分类”的动态化管理策略,让高信用等级客户获得更多融资支持,而潜在高风险客户则被纳入重点监控名单,从而实现信贷结构的优化。研究数据表明,该模型的引入使信贷审批效率提高约15%,风险预警提前期平均延长20天,为银行争取了更充足的风险应对时间。此外,模型的量化结果具有良好的可解释性,能够为风险管理部门提供清晰的决策依据。通过风险指数与财务指标间的映射关系,银行可快速定位问题企业、识别风险源头,实现由“事后补救”向“事前识别”的战略转型,提升风险防控的科学性与前瞻性。

5.2 优化模型的技术路径

虽然模型在风险识别与违约预测方面表现优越,但在样本多样性、数据实时性及算法适应性方面仍存在改进空间。未来的优化路径可从三方面入手:一是数据层优化,引入非结构化数据资源,如企业舆情信息、新闻报道、政策文件与监管通报,结合自然语言处理(NLP)与情感分析技术,拓展风险识别边界;二是算法层创新,通过引入动态贝叶斯网络与强化学习算法,实现风险指数的时序预测与模型自学习,使模型能在经济周期变化中自动修正权重与参数;三是

平台层集成,依托云计算与区块链技术,构建银行内部统一的风险管理云平台,整合授信数据、交易行为与外部征信记录,形成实时数据流动与预警机制。该路径的实施将推动模型由静态分析工具向动态智能系统演化,为商业银行提供可持续的信用风险防控能力。

5.3 模型在风险治理中的应用前景

在数字金融与智能风控持续深化的背景下,该模型的推广应用将推动商业银行从经验型、人工化授信模式向智能化、数据驱动型管理体系转变。通过模型嵌入信贷业务流程,可实现自动化信用评级与风险分层管控,提升风险识别的精准度与响应速度。未来,随着监管科技(RegTech)和金融科技(FinTech)的融合发展,模型可与监管数据接口无缝对接,形成“数据采集—风险识别—智能预警—应对反馈”的全流程闭环管理体系。值得注意的是,模型还可与ESG(环境、社会与治理)指标体系结合,拓展绿色金融和可持续发展风险识别功能,助力银行构建环境友好型信贷结构。总体来看,该模型在智能决策支持、精准授信和合规监管等方面具有广阔应用前景,不仅有助于提升商业银行的风险治理能力,也为金融体系的稳健运行与经济高质量发展提供坚实支撑。

6 结语

集团客户信用风险的识别与评估是商业银行风险管理体系的核心环节。本文通过构建多维度指标体系与综合模型,实证验证了AHP-逻辑回归-随机森林混合模型的可行性与优越性。研究结果表明,数据驱动的风险评估体系可显著提升风险识别的准确性与决策效率,为商业银行实现差异化授信与动态管理提供技术支撑。然而,模型的优化仍需数据融合、算法更新及监管协同方面持续探索。未来,应构建以大数据分析、人工智能与知识图谱为核心的智能信用风险管理体系,实现风险识别的精准化、预测化与智能化,助力商业银行在复杂金融环境下实现稳健经营与可持续发展。

参考文献

- [1] 高敏.商业银行信用风险防范策略探究[J].金融文坛,2024,(07):68-70.
- [2] 李慧.基于大数据的商业银行企业贷前信用风险评估研究[D].东南大学,2023.
- [3] 刘鹏轩.基于大数据的N商业银行信用风险预警管理的优化研究[D].南京农业大学,2022.
- [4] 刘国全.大数据环境下商业银行信用风险监测体系构建研究[D].上海财经大学,2020.
- [5] 王文倩.绿色信贷政策背景下商业银行信用风险防控研究[D].杭州电子科技大学,2024.