

基于 AI 的金融客户信用风险评估体系构建

吴怡阳

江苏银行无锡分行

DOI:10.12238/acair.v2i4.10321

[摘要] 传统的金融机构客户信用风险评估方法过于依赖某一特定指标,如信用评分、收入水平等,而忽略了其他重要因素,使得评估不够全面。因此,本研究旨在构建基于AI技术的金融机构客户信用风险评估体系。通过对金融机构客户数据的清洗、建立信用风险评估指标、人工智能模型训练以及实验验证等步骤,实现了对客户信用风险的准确评估。该体系整合了多种AI技术,包括机器学习算法和深度学习模型等,能够有效地处理大规模数据并挖掘潜在的风险特征。实验结果表明,该体系在准确性、时效性和稳定性方面表现出色,为金融机构的风险管理提供了有力的支持。

[关键词] AI技术; 金融机构; 客户信用; 风险评估; 体系构建

中图分类号: F830 文献标识码: A

Construction of financial customer credit risk assessment system based on AI

Yiyang Wu

Bank of Jiangsu, Wuxi Branch

[Abstract] Traditional financial institutions' customer credit risk assessment methods overly rely on specific indices, such as credit scores and income levels, while neglecting other significant factors, resulting in an insufficiently comprehensive assessment. This study, therefore, aims to construct an AI-based customer credit risk assessment system for financial institutions. By cleaning customer data, establishing credit risk assessment indicators, training AI models, and experimental verification, we have achieved accurate customer credit risk assessment. The system integrates various AI technologies, including machine learning algorithms and deep learning models, to effectively process large-scale data and uncover potential risk features. Experimental results demonstrate the system's excellent performance in terms of accuracy, timeliness, and stability, offering robust support for financial institutions' risk management.

[Key words] Artificial Intelligence Technology; Financial Institutions; Customer Credit; Risk Assessment; System Construction

引言

随着金融市场的不断发展和创新,金融机构面临着日益复杂的客户信用风险^[1]。传统的信用风险评估方法往往依赖于人工经验和有限的数据库,难以满足现代金融业务的需求。AI技术的出现为金融机构客户信用风险评估带来了新的机遇^[2]。人工智能具有强大的数据处理能力和学习能力,可以从海量的数据中挖掘出有价值的信息,为信用风险评估提供更准确、更全面的依据。本研究构建的基于AI技术的金融机构客户信用风险评估体系,旨在解决传统评估方法的不足,提高信用风险评估的效率和准确性。通过整合先进的AI技术和金融领域的专业知识,该体系能够为金融机构提供更加科学、有效的信用风险评估工具,帮助金融机构更好地管理风险,保障金融市场的稳定运行。

1 金融机构客户数据清洗

从金融机构内部系统(如核心业务系统、信贷管理系统等)以及外部数据源(如征信机构、公共数据库等)收集客户相关数据^[3]。数据可能包括客户基本信息、财务数据、信用记录、交易数据等。检查数据的完整性和一致性,查看是否存在缺失值、异常值或明显错误的记录,对于缺失值可以采用不同的方法进行处理。如果缺失值比例较小,可以考虑直接删除这些记录,通过插值法可以补充缺失的数据使数据更加完整,为后续分析提供更多信息。对于异常值,可以根据情况进行删除、修正或保留并进行特殊标记。使用 3σ 原则,假设客户交易金额服从正态分布,若某个交易金额超出均值 ± 3 倍标准差的范围,则视为异常值。去除异常值可以避免其对后续分析产生不良影响,使模型更加准确和稳定。为了使不同特征的数据具有可比性,需要进行标准化处理,Min-Max标准化公式如式(1)所示:

$$X_n = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

其中, X 是原始数据, $X_{\min}$ 是数据的最小值, $X_{\max}$ 是数据的最大值。标准化后的数据分布在特定范围内, 便于模型进行处理和比较不同特征的重要性。最后, 需要对清洗后的数据进行再次审查, 确保数据的质量和准确性, 可以随机抽取一部分数据进行人工检查, 或者使用一些简单的统计方法进行验证, 清洗后的数据可以更好地用于后续的信用风险评估模型构建。

2 建立信用风险评估指标

明确信用风险评估的具体目的, 例如判断客户是否可能违约、预测违约概率等。整合内部和外部数据资源, 包括客户的基本信息(年龄、职业等)、财务数据(收入、资产、负债等)、信用历史(还款记录、逾期次数等)、交易数据(交易金额、频率等)以及宏观经济数据等。假设收集到的客户年龄为35岁, 月收入8000元, 负债10万元, 过去一年有2次逾期记录, 每月交易金额平均为5000元。根据经验和相关研究, 初步确定可能与信用风险相关的指标。客户特征指标包括年龄、婚姻状况等; 财务指标包括收入、负债比等; 行为指标包括交易频率、消费习惯等。对每个潜在指标进行分析, 评估其与信用风险的相关性, 相关性分析公式如式(2)所示:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (2)$$

其中, X 和 Y 分别为两个变量(例如一个指标与违约状态), n 为样本数量, $\bar{X}$ 和 $\bar{Y}$ 分别为 X 和 Y 的平均值。通过相关性分析可以确定指标与信用风险之间的线性关系程度, 如果相关性系数接近1或-1, 表示强相关; 接近0表示弱相关。根据指标分析的结果, 选择相关性强、重要性高的指标组成信用风险评估指标体系。同时, 要考虑指标的可获取性、可解释性和稳定性。通过以上流程可以建立一个有效的信用风险评估指标体系, 为后续的信用风险评估模型构建提供基础。

3 人工智能模型训练

金融机构客户信用风险评估体系构建中, 人工智能模型训练的流程具体如图1所示。

收集经过清洗和指标确定后的金融机构客户数据, 将数据划分为训练集、验证集和测试集。通常可以按照70%、15%、15%的比例进行划分。根据问题的特点和数据的性质选择合适的人工智能模型。如果数据具有非线性关系, 可以选择神经网络模型; 如果数据量较小, 可以选择决策树或随机森林。使用训练集

对选择的模型进行训练。不同的模型有不同的训练方法和参数设置。对于神经网络, 通过反向传播算法来调整神经元之间的权重和偏置, 以最小化损失函数(如均方误差、交叉熵等)。使用验证集对训练好的模型进行评估。如果模型在验证集上的准确率较低, 可以尝试增加模型的复杂度或调整学习率; 如果模型过拟合, 可以使用正则化方法(如L1正则化、L2正则化)来减少过拟合。使用测试集对最终调整好的模型进行测试, 评估其在未见过的数据上的性能。如果模型性能满足要求, 可以将其部署到实际的金融机构信用风险评估系统中, 用于对新客户进行信用风险评估。通过以上流程可以有效地训练人工智能模型, 为金融机构客户信用风险评估提供准确的预测结果。

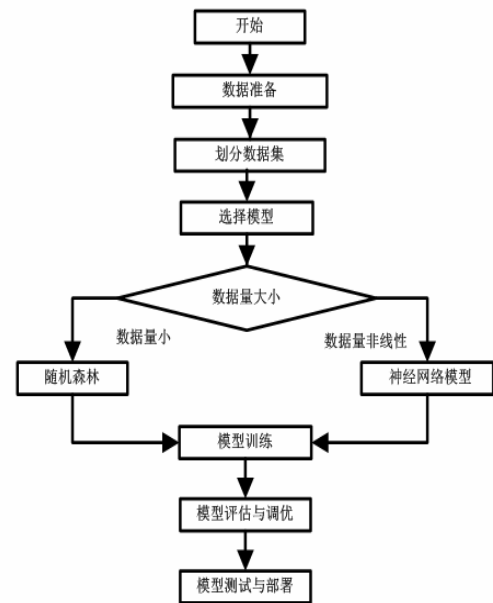


图1人工智能模型训练流程

4 应用实验

4.1 实验准备

在基于AI技术的金融机构客户信用风险评估体系构建的实验中, 某银行在客户申请贷款时, 利用该评估体系快速、准确地评估客户的信用风险, 为审批决策提供依据。审批人员可以根据评估结果确定贷款额度、利率、期限等条件, 降低违约风险。操作系统选择Windows 10; 编程语言采用Python, 使用Anaconda等集成开发环境进行管理; 机器学习框架为Scikit-learn, 用于构建和训练人工智能模型; 数据库选用MySQL, 以存储客户数据。硬件环境方面, 处理器为Intel Core i7, 保证快速的数据处理和模型训练速度。内存16GB, 较大的内存对于处理大规模数据和复杂模型而言, 可以提高运行效率。硬盘采用SSD硬盘, 其具有更快的读写速度, 能够减少数据加载时间。配备NVIDIA GeForce RTX 3060显卡, 以便加速模型训练。

4.2 实验结果

随机挑选5个客户进行实验, 具体实验结果如表1所示。

表1 实验结果

客户编号	传统评估风险等级	人工智能体系评估风险等级	收入(万元)	负债(万元)	逾期次数	交易频率(次/月)	主要差异因素分析
客户一	中等风险	低风险	50	20	1	15	近期收入稳定增长,负债比例下降,逾期次数少且交易频率稳定,体系综合多因素判断风险降低。
客户二	高风险	中等风险	40	30	3	10	新收入渠道增加收入,部分债务得到清偿,逾期情况改善,体系精准评估风险下降。
客户三	低风险	低风险	35	15	0	20	收入、负债、逾期及交易频率等指标稳定良好,体系准确判断维持低风险。
客户四	中等风险	高风险	45	35	2	5	行业环境变化影响收入,负债较高且交易频率降低,逾期风险增加,体系及时调整风险等级。
客户五	高风险	高风险	30	25	5	8	负债高企难以缓解,逾期次数多且交易频率不高,体系确认高风险。

传统评估方法对客户一、二、四的风险判断与实际情况有一定偏差,而人工智能体系通过综合考虑收入、负债、逾期次数和交易频率等多方面因素,能够更准确地评估客户的信用风险等级。该体系能够实时监测客户的各项指标变化,并及时调整风险等级。以客户四为例,当行业环境变化导致收入受影响、负债较高且交易频率降低、逾期风险增加时,人工智能体系能够迅速捕捉到这些变化,将其风险等级从中等风险调整为高风险,为金融机构提供及时的预警,以便采取相应的风险管理措施。传统评估方法可能过于依赖单一因素或主观判断,而人工智能体系能够全面整合客户的多维度数据进行分析。对于客户三,传统评估和人工智能体系评估结果一致,都为低风险,这表明该体系在稳定情况下能够保持准确的判断,具有较高的稳定性。同时,对于客户五这种高风险客户,体系也能够准确确认其风险等级,不会出现误判,保证了评估结果的一致性。

综上所述,基于AI技术的金融机构客户信用风险评估体系具有准确性高、动态监测能力强、全面客观、稳定性好等优势,能够为金融机构提供更可靠的信用风险评估,有助于金融机构更好地管理风险和做出决策

5 结束语

基于AI技术的金融机构客户信用风险评估体系的构建是金

融领域的一次重要创新。通过本研究的努力,成功地构建了一个具有较高准确性、时效性和稳定性的信用风险评估体系。该体系的应用将有助于金融机构更好地识别和管理客户信用风险,提高金融机构的风险管理水平和竞争力。然而也认识到,AI技术在金融领域的应用仍处于不断发展和完善的过程中。在未来的研究中,将继续深入探索AI技术在信用风险评估中的应用,不断优化评估体系的性能和功能。同时,也将加强与金融机构的合作,将研究成果更好地应用于实际业务中,为金融市场的稳定和发展做出更大的贡献。

[参考文献]

- [1]木强,秦琳琳.金融风险管理中的统计模型与方法[J].商讯,2024,(13):88-91.
- [2]海仁措毛.金融科技背景下网上业务风险与对策研究——以商业银行为例[J].北方经贸,2024,(05):98-100.
- [3]阮一凡.AI技术在金融风控中的应用研究[J].商展经济,2024,(07):89-92.

作者简介:

吴怡阳(2001--),男,汉族,江苏省江阴市人,江苏银行无锡新区支行,本科,研究方向:大数据分析,信用风险评估。