

基于人工智能的中小金融机构内部审计数据处理方法

姚宇

辽宁农村商业银行股份有限公司 辽宁沈阳 110032

摘要: 人工智能在大数据处理、模式识别和预测分析等领域的应用,为金融机构内部审计提供强有力的技术支持。本文探讨基于人工智能的中小金融机构内部审计数据处理方法,分析人工智能如何提升数据整合、风险监测、非结构化数据分析及合规检查的能力。重点阐述智能数据预处理、风险预警系统的构建以及合规性检测的智能化实施,提出利用人工智能技术优化审计流程的方法,并探索其在提升审计效率与风险管控能力方面的创新应用。

关键词: 人工智能; 内部审计; 风险监测

作者简介: 姚宇(1988-), 硕士研究生, 中级审计师、中级经济师, 研究方向: 银行内部审计理论与实务

引言

随着金融科技的迅猛发展,中小金融机构面临着数据量激增与审计复杂度增加的双重挑战^[1]。内部审计作为金融机构风险管理的关键环节,必须实现精准、高效的风险评估与合规审查^[2]。随着人工智能技术的不断进步,尤其是在大数据分析、机器学习与自然语言处理等领域的突破,人工智能为内部审计带来新的机遇^[3]。利用智能化的数据处理,人工智能能够提升审计工作效率、降低风险漏报率、优化合规性检查流程^[4]。本文目的在于探讨基于人工智能的中小金融机构内部审计数据处理方法,分析其应用潜力与实现路径。

1 中小金融机构内部审计数据处理需求

1.1 数据整合需求

中小金融机构审计数据来源丰富,包含核心业务系统、交易记录等多维度数据。内部审计人员需从不同数据库、表单等提取数据,整合至统一平台分析。其核心需求是高效准确获取数据并统一格式标准。在金融行业,审计数据整合不仅有结构化数据,还涉及诸多非结构化数据,因混合数据源特性,设计时要兼顾数据格式多样与不同数据源的连接性。数据整合中,安全性与合规性不可或缺,数据传输和存储须满足监管对金融数据保护的严格要求。这不仅是技术问题,更是合规要求,涉及敏感信息加密传输、访问控制等安全措施。只有保障数据安全合规,才能完成整合,为风险识别等提供优质数据基础。

1.2 风险监测需求

在中小金融机构内部审计中,风险监测是实时识别、评估潜在审计、操作、市场等风险的关键。机构面临信用、流动性等多种风险,常以数据异常、交易模式变化等形式呈现。因数据分散于多系统,具有高维度、复杂特性,传统方法难以有效识别风险。人工智能,尤其是机器学习算法,能从历史数据构建风险预测模型,实时监控交易、账户等数据,自动识别大额资金频繁流动等异常模式,快速锁定操作风险与欺诈行为。同时,监测系统需具备多维度风险评估能力,人工智能可处理海量数据,构建综合模型,分析金融产品、客户行为等因素,输出量化预警,助力审计人员及时应对。此外,高度自动化与智能化的风险监测系统,能让中小金融机构在资源有限时高效把控风险,提升审计效率,减少人工审查疏漏。

1.3 数据处理需求

审计数据处理涉及数据的清洗和转换,还需快速、准确地进行分析,为决策提供支持。数据处理的首要需求是清洗和标准化,由于审计数据来源广泛,格式多样,必须借助自动化手段清洗数据,去除冗余、填补缺失并统一格式,减少人工错误。数据的转化和聚合同样关键,中小金融机构审计数据结构复杂,需用自动化工具将其转化为统一格式并有效聚合,以实现统一分析,而人工智能可识别数据源关系,助力数据整合,提升处理效率。此外,数据分析和建模需求迫切,传统人工分析效率低且易疏漏,人工智能通过自动化数据挖掘与建模,能快速识别潜在规律、异常数据与风险,既加快数据处理速度,又精准定位审计风险。

综上所述，审计数据处理核心需求涵盖高效的数据清洗、标准化、转化聚合，以及基于人工智能的建模能力。

1.4 合规检查需求

在合规检查中，首要需求是自动识别与更新合规标准。金融机构面临的法规政策繁多且变动频繁，人工审计难以实时追踪，而人工智能借助自然语言处理技术，能自动解析法规，及时更新标准，保障审计符合最新要求。同时，合规检查需高效处理数据并验证一致性。金融业务数据层面复杂，人工审核效率低，人工智能可自动化筛查交易、账户等数据，快速揪出利率不符等不合规行为，提升审计效能。合规性分析要覆盖多维度，人工智能综合业务类型、客户信息等因素，进行全面评估，并依据实时数据和法规动态调整检查规则。合规检查系统必须具备实时性与灵活性。鉴于金融市场和监管环境瞬息万变，人工智能可持续更新合规规则，动态监控风险，确保审计各环节紧跟最新规定。

2 基于人工智能中小金融机构内部审计数据处理方法设计

2.1 实现智能数据预处理

人工智能可以自动清洗、标准化、整合和转换数据，提高处理速度和质量。对于数据清洗，人工智能采用 KNN 算法填补缺失值，利用聚类分析识别并剔除异常数据，避免人工干预。重复数据可采用相似度计算自动合并。数据标准化和归一化也能利用人工智能自动完成，将不同量纲的数据转化为统一标准，确保后续分析的准确性。

将数据调整为零均值、单位方差，消除量纲差异，归一化将数据映射到 $[0, 1]$ 区间。对于数据整合，人工智能采用自编码器或图神经网络等深度学习技术自动融合来自不同系统的数据，消除格式差异，提供一致的分析基础。特征选择和降维也是人工智能处理的关键环节，Lasso 回归、决策树等算法可自动筛选出有用特征，减少计算量并提高精度。利用主成分分析等降维技术，人工智能能降低数据维度，减少处理时间。

2.2 构建风险识别预警系统

针对风险监测需求，人工智能可以实时分析多维数据，自动检测异常行为，并为审计人员提供风险预警信号。构建风险识别预警系统的核心方法是数据建模和风险预测，利用历史数据和实时交易数据，人工

智能可以利用监督学习算法训练逻辑回归、决策树、随机森林、支持向量机等风险预测模型。这些模型能够根据历史案例和当前数据，识别出潜在的风险模式并进行分类。如在识别信用风险时，人工智能模型可以根据客户的信用历史、还款能力、交易行为等多个维度的数据，判断客户违约的可能性。公式如下：

$$P(\text{风险}) = \sigma(w_1x_1 + w_2x_2 + \dots + w_nx_n + b) \quad (2)$$

$P(\text{风险})$ 式中， $x_1, x_2, \dots, x_n$ 表示发生风险的概率， $w_1, w_2, \dots, w_n$ 为对应特征的权重， b 为偏置项， σ 为激活函数（如 Sigmoid 函数），用于计算风险的概率值。利用训练过程，人工智能能够动态调整各特征的权重，实时识别风险并进行预测。除了使用监督学习模型，人工智能还可以应用无监督学习方法识别未标注的异常数据。在金融机构内部审计中，异常交易、资金流动异常、客户信息不一致等现象常常预示着潜在的操作风险或欺诈行为。人工智能利用自适应算法，可以对大量未标注的数据进行分析，自动识别出可能存在风险的模式。如使用 K-means 聚类算法将交易行为划分为正常与异常两类，再利用距离度量自动识别出与大多数数据点差异较大的异常行为。在构建风险预警系统时，实时数据监控和反馈机制尤为关键，金融市场和客户行为快速变化，传统的风险评估模型无法适应实时变动的市场条件。人工智能利用深度学习和流数据处理技术，可以持续更新模型，实时分析新产生的数据。例如，人工智能可以采用 LSTM（长短期记忆网络）模型对金融时间序列数据进行预测，评估未来短期内潜在的市场波动或信用风险。时间序列预测模型可表示为：

$$y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-n}; \theta) \quad (3)$$

式中， y_t 表示时间 t 时刻的预测值， f 为 LSTM 模型的函数， θ 为模型的参数， $t-1, t-2, \dots, t-n$ 为历史数据点。该模型能够捕捉数据中的时序关系，精准预测市场波动，为审计人员提供前瞻性的风险警示。人工智能在风险识别中的另一个关键优势是其自动化和自我学习能力。随着数据量的增大，人工审计的效能逐渐下降，而人工智能可以利用增量学习和在线学习算法，持续优化模型，减少人为干预的需求。人工智能模型的自我调整能力使得系统能够适应新类型的风险，从而提高风险识别的准确度。

窗体顶端

窗体底端

2.3 开展非结构化数据分析

中小金融机构处理的数据种类繁多，除传统的结构化数据外，还涵盖大量文本、图像、音频、视频等非结构化数据。对于金融机构而言，非结构化数据分析的核心任务之一是文本数据的处理与挖掘。金融领域的非结构化文本数据通常表现为合同文件、财务报表、客户评论、市场分析报告等，人工智能利用自然语言处理（NLP）技术对这些数据进行解析和理解，识别出潜在的风险点。人工智能可先使用 TF-IDF 或 Word2Vec 等文本向量化技术将文本转化为数值特征，便于后续的分析。TF-IDF 方法能够衡量某个词在文档中的重要性，从而提取出金融审计中关键的风险信息。文本分类也是非结构化数据分析中的任务。机器学习算法使人工智能能够自动从大量文档中识别出与审计相关的文本信息，并对其进行分类。如采用支持向量机或 LSTM 神经网络等深度学习网络进行分类，判定某个文档是否属于潜在的风险类别。在图像和视频数据的分析中，人工智能技术可以帮助审计人员从图片、监控视频等中提取出有价值的信息。例如，利用卷积神经网络（CNN），人工智能能够识别出财务报表中的签名、图表、表格中的异常信息，以及分析监控视频中的异常行为。CNN 的核心运算是卷积操作，其公式为：

$$y(x,y)=(f*g)(x,y)=\sum_i\sum_jf(i,j)g(x-i,y-j) \quad (6)$$

式中， f 为输入图像， g 为滤波器（卷积核）， $y(x,y)$ 为输出特征图。这种方式使人工智能可以自动从图像或视频数据中提取出审计所需的信息，减少人工审查的工作量。音频数据分析也是中小金融机构非结构化数据分析的一部分。

2.4 实施审计合规性检测

人工智能技术采用自动化和智能化手段，提升合规性检测的效率和准确性。人工智能在合规性检测中的核心作用是自动识别潜在违规行为，采用对历史数据的学习，人工智能模型能够识别不合规的交易模式、异常资金流动等行为。如使用支持向量机等算法对交易、账户和客户数据进行实时分析，自动检测是否存在违规行为。在反洗钱（AML）审计中，人工智能能够分析资金流动、客户交易频率等数据，自动识别可疑交易并发出警报。人工智能还可应用于合规性文档的

解析。利用自然语言处理（NLP），人工智能可以将复杂的法律法规转化为机器可识别的规则，快速判断金融行为是否符合合规要求。合规性检测的实施过程主要为数据采集、预处理、模型训练、合规评估和结果反馈（表 1）。深度学习技术的引入进一步增强合规性检测的能力，深度神经网络采用持续学习，能够识别复杂的合规规则，并自动评估客户信用、交易行为等是否符合要求。

表 1 合规性检测的实施过程

步骤	任务	应用技术
数据采集	收集交易、客户及贷款数据	数据接口、API 抓取
数据预处理	清洗和格式化数据	数据清洗、归一化
模型训练	构建合规性检测模型	机器学习（SVM、决策树）
合规性评估	判断数据是否符合合规要求	规则引擎、NLP
结果反馈	输出评估报告与警报	可视化、报告生成工具

3 结语

本研究探讨基于人工智能的中小金融机构内部审计数据处理方法，重点分析数据整合、风险监测、数据处理和合规检查等需求，并提出相应的解决方案。人工智能技术在数据预处理、风险识别、非结构化数据分析及合规性检测中的应用，有效提升审计工作的效率和准确性。人工智能的引入优化审计流程，还为中小金融机构提供更为智能化的风险防控手段。

参考文献

周洋. 企业内部审计体系建设存在的问题及其对策 [J]. 中国农业会计, 2024, 34(22): 91-93.

古丹丹. 新形势下基于 WSR 的内部审计管理路径探究——以 A 集团为例 [J]. 中国农业会计, 2024, 34(22): 94-96.

傅添图. 国有企业内部审计对风险控制的作用与策略 [J]. 今日财富, 2024, (34): 80-82.

武胜芬. 内部审计下国有企业财务会计与税务会计的融合策略 [J]. 今日财富, 2024, (34): 92-94.

孔敏. 企业内部审计与财务会计协同管理模式研究 [J]. 商业文化, 2024, (21): 122-124.