

人工智能在财务舞弊识别中的应用

张丽静

黑龙江职业学院，黑龙江哈尔滨，150080；

摘要：随着全球资本市场快速发展，财务舞弊事件频发，给投资者造成损失，侵蚀市场信心和金融体系稳定性。传统财务舞弊识别方法依赖审计师经验和基于规则的静态分析，面对复杂、隐蔽且大规模数据时力不从心。近年来，人工智能技术发展为财务舞弊识别带来机遇。本文系统探讨人工智能在该领域的应用。首先，分析传统舞弊识别方法局限性，阐述人工智能应用的必要性与优势；其次，论述机器学习、自然语言处理、复杂网络分析等核心 AI 技术在识别各类舞弊方面的应用模式与案例；再次，剖析当前 AI 技术应用面临的数据质量、模型可解释性等主要挑战；最后，展望未来发展趋势，提出构建人机协同智能审计生态系统、加强跨学科人才培养、完善法规标准等建议，推动 AI 技术提升资本市场透明度和诚信水平。

关键词：人工智能；财务舞弊；机器学习

DOI：10.64216/3104-9702.25.03.053

引言

财务舞弊是困扰全球资本市场的顽疾。从早期的安然事件、世通公司丑闻，到近期的瑞幸咖啡财务造假，每一次重大的舞弊案都引发了市场巨震，并对审计行业、监管机构和社会公众提出了严峻挑战。传统的财务舞弊识别手段高度依赖于审计师的职业判断和对抽样数据的分析。然而，在“大数据时代”，企业交易数据量呈指数级增长，业务模式日趋复杂，舞弊手段也更加隐蔽和高科技化，这使得传统方法的效率和效果都面临瓶颈。审计师难以仅凭人力从海量、多维度数据中精准定位异常模式。

在此背景下，人工智能技术凭借其强大的数据处理、模式识别和预测分析能力，为财务舞弊识别提供了全新的范式。AI 能够 7x24 小时不间断地分析整个数据集（而非抽样），从结构化财务数据和非结构化文本信息（如管理层讨论、新闻舆情、合同文件）中挖掘出人难以察觉的细微异常和复杂关联，从而实现对舞弊风险的前瞻性预警和精准打击。将 AI 技术引入财务舞弊识别，不仅是技术发展的必然趋势，更是提升审计质量、强化公司治理、保护投资者利益的迫切需求。

本文旨在系统性地梳理和探讨人工智能在财务舞弊识别中的应用全景。论文将首先概述 AI 技术的核心分支，然后深入分析其在具体舞弊识别场景中的应用逻辑与价值，客观评估当前存在的挑战，并最终对未来发展提出战略性思考。

1 人工智能技术概述及其在舞弊识别中的适用性

人工智能领域宽泛，其多个子技术在财务舞弊识别中各有优势，构成强大技术工具箱。机器学习是 AI 在财务舞弊识别中应用最广的技术，核心是让计算机从历史数据学规律，用模型对新数据预测或分类。监督学习是主流应用方式，前提是有已标记的“舞弊”和“非舞弊”公司样本数据，算法学习样本构建分类模型预测新公司舞弊风险，如用已知舞弊公司财务比率训练模型。无监督学习因已标记舞弊样本稀少而重要，它无需预先标记，直接在数据中找内在结构或异常点，聚类分析可分组相似行为公司，孤立点分析能识别异常观测值。深度学习是机器学习分支，通过构建多层神经网络模型，自动从原始数据学习高层次、复杂特征，如循环神经网络处理时间序列数据，卷积神经网络分析文档图像或识别交易数据空间模式。财务舞弊也藏于文字中，NLP 技术用于文本分析。情感分析分析年报、发布会实录、高管言论等判断语调，评估管理层诚信度和舞弊动机；主题建模从大量文档提取关键主题，识别公司相关风险话题；欺骗性语言检测研究发现舞弊公司语言风格有特定模式，NLP 模型可据此给文本欺骗性评分。现代企业舞弊涉及复杂关联方交易网络，复杂网络分析将企业、高管等视为节点，交易等关系视为边构建关系网络，分析其拓扑结构可识别隐藏关联方、异常资金闭环和潜在操纵核心，揭露结构性舞弊。

2 AI技术在财务舞弊识别中的具体应用场景

2.1 识别虚构收入与虚假交易

虚构收入是较为常见的舞弊手段。人工智能（AI）技术可从多个维度对其进行识别：

财务数据异常检测：机器学习模型能够对销售收入与应收账款、现金流、库存等科目的勾稽关系予以监控。例如，若销售收入呈现大幅增长态势，但应收账款周转天数急剧上升，且经营活动现金流持续为负，模型会将其标记为高风险异常情况。无监督学习中的孤立森林算法十分适用于在此场景下发现此类“背离常态”的异常公司。

交易流水分析：深度学习模型能够对每一笔交易的元数据（如时间、金额、交易对手、地点等）进行分析。例如，模型可以识别出在期末集中出现的大量小额交易（具有“刷单”特征），或者与某些背景不明的交易对手进行的大额、异常交易。

供应商/客户网络分析：借助复杂网络分析方法，可以绘制出企业的全部供应商和客户网络。通过社区发现算法，有可能识别出一批注册地址相近、成立时间接近、主要人员重叠的“幽灵客户”或“空壳供应商”，这些主体的存在极有可能是为了虚构交易。

2.2 揭露关联方交易舞弊

关联方交易由于其复杂性和隐蔽性，成为审计工作的重点与难点。人工智能在这一领域具有较大的应用潜力：

实体关系识别：自然语言处理（NLP）技术能够自动从海量的工商信息、招股说明书、年报附注等文本中提取公司、股东、高管、子公司等信息，并构建出完整的股权结构和任职关系图谱。这相较于人工查询和核对更为高效、全面。

交易公平性评估：在识别出关联方的基础上，机器学习模型可以将关联方交易的条款（如价格、利率、信用期）与同行业、同规模的独立第三方交易进行对比，自动评估其公允性，并对定价异常的交易发出警报。

2.3 评估管理层舞弊意图与“软信息”分析

舞弊三角理论表明，舞弊机会、压力和借口是导致舞弊行为发生的三大条件。人工智能可以有效地评估后两者，即管理层面临的压力和寻找的借口。

管理层语调分析：应用自然语言处理情感分析技术，对首席执行官（CEO）和首席财务官（CFO）在业绩发布

会上的发言进行分析。若公司面临巨大的业绩压力（如对赌协议、股价下跌风险），而管理层在描述业绩时使用异常积极的语言或回避关键问题，其舞弊风险将显著升高。

分析师报告与媒体舆情监控：人工智能系统能够实时抓取和分析券商分析师报告、新闻媒体、社交媒体上关于该公司的评论。若负面舆情持续发酵，但公司官方回应乏力或前后矛盾，这同样是重要的风险信号。

2.4 增强持续性审计与监控能力

传统审计具有周期性特点，而人工智能可以实现持续性的审计监控。通过将人工智能模型部署在企业的企业资源计划（ERP）或交易系统中，系统能够实时监控每一笔新产生的交易和凭证，并与历史模式、行业基准进行实时对比。一旦发现异常模式（如某供应商的发票金额突然激增，或某员工的操作行为偏离常态），系统会立即向审计师发出预警，从而实现从“事后侦测”到“事中干预”的转变，极大地提高了审计的时效性和威慑力。

3 人工智能应用面临的主要挑战与局限性

尽管AI在财务舞弊识别领域具有广阔的应用前景，但其应用仍面临着诸多挑战。若这些挑战不能得到妥善解决，将不仅影响AI在该领域效果的发挥，甚至可能带来新的风险。

数据质量方面，“垃圾进，垃圾出”是AI领域的基本准则。模型的准确性在很大程度上依赖于输入数据的质量。企业财务数据可能存在错误、缺失或不一致的情况，特别是在进行跨系统、跨地域整合时。

数据可获得性方面，审计师获取全面、实时的内部数据可能会受到权限和合规方面的限制。此外，最具价值的舞弊样本极为稀少，且其特征多变，这使得监督学习模型容易出现过拟合或泛化能力不足的问题。

许多先进的AI模型（尤其是深度学习模型）是复杂的“黑箱”，它们虽能给出高精度的预测，但难以解释其做出判断的具体原因和逻辑。在严谨的审计和司法环境中，仅依据一个模型给出的“高风险”结论，不足以作为确凿的证据。审计师和监管机构需要了解是哪些具体因素触发了警报，以便进行职业判断和后续调查。因此，如何提升模型的可解释性，是AI应用于高风险决策领域必须克服的障碍。

如果训练数据本身存在偏见（例如，历史上被发现的舞弊公司多集中于特定行业或特定规模），那么训练

出的模型也会继承这种偏见,从而导致其对某些类型的公司产生误判。同时,由于舞弊样本稀少,模型很容易对有限的舞弊样本特征形成“记忆”,而无法学习到舞弊的本质规律,进而出现模型在训练集上表现优异,但在真实场景中效果不佳的情况。

合规性方面,现有的审计准则和监管法规是基于传统审计模式制定的,对于AI模型作为审计证据的可靠性、审计轨迹的留存以及审计师的责任界定等方面,尚未有明确的规范。

伦理与隐私方面,对员工通信、邮件等进行分析以识别舞弊行为,可能会触及员工的隐私权。利用外部舆情数据也可能引发数据来源合法性的争议。因此,如何在有效识别舞弊和保护个人隐私之间取得平衡,是一个重要的伦理问题。

成功实施AI驱动的舞弊识别项目,需要既精通审计实务和会计准则,又掌握数据科学和AI技术的复合型人才。目前,这类跨界人才在市场上极为稀缺,这已成为制约该技术落地应用的一大瓶颈。

4 未来展望与发展建议

面对挑战,未来的发展宜聚焦于构建一个健康且可持续的人机协同生态系统。从单一工具迈向集成平台:人工智能(AI)应用将从孤立的风险预警工具,发展成为集数据集成、特征工程、模型训练、可视化分析以及 workflow 管理于一体的智能审计平台。可解释性人工智能的深度融合:未来的模型将更注重可解释性,例如运用局部可解释模型无关解释(LIME)、沙普利可加解释模型(SHAP)等解释性工具,为人工智能的决策提供直观且可信的依据,增强审计师的信任。多模态学习成为主流趋势:对财务数据、文本数据、网络关系数据乃至语音、图像信息进行融合分析的多模态学习,能够更全面地刻画企业行为,提高识别的精准度。联邦学习保障隐私:联邦学习技术允许在不交换原始数据的情况下共同训练模型,这为解决数据孤岛和隐私保护问题提供了可行途径,便于会计师事务所或监管机构在跨客户/跨公司层面构建更强大的风控模型。

对于会计师事务所和企业而言,应加大在审计科技领域的投入,积极与科技公司开展合作,从开展小规模试点项目着手,逐步积累数据与经验。同时,加强对现有审计人员人工智能技能的培训,并引入数据科学家,

组建混合作战团队。对于监管机构而言,应积极跟踪技术发展动态,适时修订审计准则和监管指引,明确人工智能在审计工作中的地位、使用规范以及责任边界。鼓励在脱敏和合规的前提下进行数据共享,为行业研发提供高质量的基准数据集。对于学术界与教育界而言,应大力推动会计、审计与计算机科学的跨学科研究,并改革课程体系,培养面向未来的“智能审计师”和“会计数据科学家”。

5 结论

人工智能技术正深刻地重塑财务舞弊识别的模式,使其从一项依赖个人经验与抽样测试的技艺,转变为一门基于全量数据分析与智能算法的科学。借助机器学习、自然语言处理和复杂网络分析等工具,人工智能能够以空前的广度、深度和速度挖掘舞弊线索,大幅提高识别的准确性与时效性。我们需清醒地认识到,人工智能并非万能之策。其应用效果受数据质量、模型可解释性、算法偏见以及伦理法规等多种因素的限制。技术的最终目标并非取代审计师的专业判断,而是赋能审计师,成为其强有力的“超级助手”。未来的发展方向必定是构建一个人机协同的智能审计生态系统:人工智能负责处理海量数据、发现异常模式,而审计师则凭借其深厚的行业知识、商业逻辑和职业怀疑精神,对人工智能的发现进行洞察、核实与判断。展望未来,随着技术的持续成熟和行业生态的不断完善,人工智能必将在维护资本市场诚信、提升资源配置效率、构建透明可信的商业环境中发挥愈发关键的作用。拥抱这一变革,积极应对其带来的挑战,是审计行业、监管机构乃至整个商业社会不可回避的问题。

参考文献

- [1]朱晓薇.财务报表舞弊识别与防范机制[J].城市情报,2024(19).
- [2]叶钦华黄世忠.财务舞弊的新特征与新应对[J].财会月刊,2024,45(19):3-8.
- [3]黎楚君.人工智能下的反舞弊进程及优化策略探究[J].品牌研究,2024(32):0288-0290.

作者简介:张丽静,1975年10月19日,女,黑龙江职业学院教授。