

大数据驱动债券风险识别，统筹防范风险与优化金融资源配置

2025 年国家自然科学基金重点项目：大数据驱动的公司债券风险识别与资产配置

项目负责人：陆瑶，清华大学经济管理学院金融系教授，清华大学经济管理学院中国金融研究中心主任

在当今世界数字经济浪潮的推动下，大数据与人工智能（AI）正以前所未有的速度与深度重塑传统金融市场。这不仅是一场技术革新，更是一场由内而外的系统性金融模式重构。尽管我国持续出台相关政策推动数据要素市场化配置，但是金融市场仍面临一系列结构性挑战。传统债券风险识别方法在数据维度、信息时效、模型精度及主观评估等方面存在明显不足，难以满足现代风险管理的需求。同时，中小企业融资难、融资贵的问题依然突出，普惠金融虽取得一定进展，但资金供需匹配效率依旧偏低。在此背景下，大数据技术可通过整合多模态信息，构建更全面、实时的风险评估体系，从根本上推动风险识别与资产配置的变革；AI 技术则依托机器学习与深度学习算法，生成高精度信用评估模型，在提升融资可得性的同时优化风险管理平衡。由此，深入探讨大数据与 AI 驱动公司债券的风险识别与资产高效配置等方面的深度应用，已成为金融领域的重要理论前沿。这不仅具有显著的学术价值，也对推动数字经济与实体经济深度融合、促进资本市场健康稳定运行、强化金融支持经济高质量发展具有重要的现实意义。

清华大学经济管理学院金融系陆瑶教授的 2025 年度国家自然科学基金重点项目《大数据驱动的公司债券风险识别与资产配置》正式获批。该项目将通过理论与实证研究，结合典型案例与政策分析，系统评估大数据驱动下的公司债券风险识别与资产配置机制，并深入揭示其内在作用机理。项目旨在为数智技术与实体经济的高效融合提供科学参考，助力提升资本市场运行的韧性与活力，并为企业决策及政府政策制定与优化提供坚实的理论支撑与数据驱动的决策依据。



本项目旨在综合运用理论分析、实证研究、机器学习建模、案例研究与政策研究等方法，系统探讨大数据与人工智能在公司债券风险识别与资产配置中的深度应用。研究围绕五大核心方向展开，构建了完整的研究框架。首先，项目将聚焦于大数据与人工智能驱动的公司债务违约风险识别。研究将整合财务报表、供应链动态等多模态数据，并基于大语言模型生成和处理文本、图像、视频、音频等非结构化数据，构建更全面的企业风险评估体系，实现对债务违约的精准预测。在此基础上，研究将分析大数据与人工智能在优化债务融资交易结构中的作用。借助精准的信用风险评估，相关技术能够有效缓解企业与债权人之间的信息不对称，提升债券市场的定价效率与融资匹配度，支持更具针对性的信贷产品设计。第三，项目将探讨大数据与 AI 赋能普惠金融与资产配置的路径。通过构建高精度信用评分模型，相关技术可为中小企业及弱势群体提供可负担的金融服务，弥补传统金融因信息缺失带来的服务不足。

然而，技术应用的广泛拓展也伴随潜在挑战。第一，大数据与人工智能可能引发统计歧视问题。研究将探讨如何在技术赋能普惠金融的同时，避免因算法偏差导致的不公平。第二，需深入分析信贷场景下人工智能算法披露的影响，重点揭示其对放贷企业违约率、金融机构估值及普惠金融发展的作用机制，并提出优化算法披露机制的策略，以在提升市场透明度的同时，有效防范因信息不对称加剧而引发的套利风险。第三，关注信贷场景下的人机协同模式，评估其在企业债务违约风险识别与普惠金融发展中的作用，并探索优化人



机交互机制的路径，将人工智能在高效数据处理方面的优势与人类专家在直觉、经验和情境判断上的优势相结合，从而提升决策的准确性、公平性与效率。

本项目旨在围绕公司债券的风险识别与资产配置，引入大数据与多模态信息，构建一个系统性、前瞻性、多维度融合的研究框架，从而提升风险识别的及时性与准确性，优化债券市场的定价效率与资源配置水平，并为中小企业创造更加公平、可持续的融资环境。项目成果不仅有助于强化金融服务实体经济的能力，还将为监管机构提供基于数据驱动的风险预警工具与算法治理方案。本项目的创新性主要体现在以下几个方面：第一，本项目是一项交叉性研究，涵盖从公司债券风险识别到资产配置的完整链条，兼顾效率提升与风险防范的双重目标。第二，本项目力图整合财务、ESG、供应链等结构化数据与文本、图像、视频、音频等非结构化信息，构建具备高预测能力的动态风险评估体系。第三，本项目强调理论与实践的深度结合，不仅通过理论模型揭示技术作用机理，还将借助实证研究、机器学习建模与案例分析进行验证，从而为政策制定与市场实践提供可操作的参考。第四，本项目将在理论上聚焦数字技术风险治理的关键痛点，探索优化金融科技算法公平性的治理框架，从而更好发挥金融在服务实体经济和支撑高质量发展中的作用。

供稿：科研事务办公室 编辑：杨海琴 责编：吴淑媛 赵霞