

# 人工智能赋能金融强国建设： 理论逻辑、现实路径与风险挑战\*

汪 勇 范云朋 王 瑶

〔摘要〕人工智能已超越单纯的技术工具范畴,逐步演化为重塑金融资源配置、风险识别及价格形成机制的关键力量。在加快建设金融强国的背景下,人工智能深刻重塑了金融运行和发展的内在逻辑。但是这种变革具有明显的双重属性,在赋能金融“五篇大文章”的同时,也伴随着人工智能投资泡沫、模型幻觉、算法同质化、数据隐私泄露等新问题。以人工智能赋能金融强国建设,要坚持发展与安全并重,突破人工智能基础理论、方法体系与关键工具,统筹推进算力基础设施与数据资源体系建设,构建规范透明的法律法规与伦理治理框架,健全人工智能技术监测与风险应急体系,引导人工智能更好发挥作用。

**关键词:**人工智能 金融发展 金融监管 金融强国

**JEL 分类号:**E20 E60 G00

## 一、引言

习近平总书记指出:“加快发展新一代人工智能是事关我国能否抓住新一轮科技革命和产业变革机遇的战略问题。”《中共中央关于制定国民经济和社会发展第十五个五年规划的建议》提出全面实施“人工智能+”行动,并要求加强人工智能治理和完善法律法规。2025年12月召开的中央经济工作会议强调,深化拓展“人工智能+”,完善人工智能治理。在加快建设金融强国和推动高质量发展的背景下,围绕人工智能如何赋能金融发展,有必要展开更加系统深入的研究,引导人工智能走向善治,促使先进技术在金融强国建设过程中发挥积极作用。

本文所讨论的人工智能(Artificial Intelligence,简称AI),并非指具体算法或单一技术工具,而是指一种以数据要素为基础、以算法模型为核心、以辅助乃至参与决策为主要功能的通用性技术形态。其关键特征不在于技术形式本身,而在于对信息处理能力的显著扩展。人工智能通过整合分散于不同场景和环节的信息,改变了金融决策主要依赖少量财务指标和静态规则的传统模式,使金融机构能够在

\* 汪勇,中国社会科学院金融研究所,副研究员,经济学博士;范云朋,中国社会科学院金融研究所,副研究员,经济学博士;王瑶,中国社会科学院金融研究所,助理研究员,经济学博士。

更细致的层面识别经济主体的经营状况、风险特征和行为变化,从而对金融运行方式和发展模式产生持续影响。

人工智能在金融领域的深入应用,是其渗透整个经济社会体系的一个典型切片。金融作为现代经济的核心,其运行机制始终与社会经济结构紧密关联。当前,社会经济正在经历数字化和智能化转型,人工智能逐步融入生产、消费、治理各环节,重塑了社会信息结构和行为模式。在这一背景下,金融领域应用人工智能既有其内因,即人工智能能够极大提升金融运行的信息处理效率、风险识别精度和服务覆盖范围;也有其外因,即金融体系必须与基于人工智能的社会经济系统有效对接,以适应全社会数据驱动、智能互联的发展趋势。同时,经济社会运行的信息化为金融提供了前所未有的海量、多元、实时数据资源,进一步加速了人工智能在金融领域的融合与创新。

人工智能在金融领域的应用所带来的问题,也同时反映出其在全社会场景中面临的普遍性挑战和金融领域的特别约束。普遍性挑战包括算法偏见、数据隐私、AI幻觉、伦理责任以及新型欺诈等,这是人工智能在各行业应用中面临共同的问题。金融领域因其对信息真实性、决策精确性和系统稳定性的极高要求,以及其巨大的外部性和强监管属性,使得上述问题表现得尤为突出和敏感,要求人工智能输出必须可靠、可解释、可追溯,同时必须在合规和风控的框架下运行。因此,探讨人工智能如何重塑金融发展,不仅需要关注技术本身的能力提升,也需深入分析在金融这一特殊领域内,如何平衡创新激励和风险防范、效率提升和公平保障、技术自主和监管适配等多重关系。

我国高度重视人工智能的发展。2025年以来,“人工智能+金融”的概念被明确提出和广泛使用。国家金融监督管理总局2025年12月印发《银行业保险业数字金融高质量发展实施方案》,首次对“人工智能+金融”提出了明确的发展要求,并指出要防范人工智能带来的金融风险。在国家发展战略的指引下,我国金融业在发展中积极拥抱人工智能,不仅使得后者成为重构金融服务成本结构、风险模式和用户体验的主要驱动力,也有力促进了人工智能尤其是大模型和金融学研究的结合(苏治等,2017;李斌等,2019;范小云等,2022;姜富伟等,2024)。本文基于人工智能在金融领域的应用和发展趋势,从理论上厘清人工智能重塑金融功能的内在逻辑,从实践上论述其作用于以“五篇大文章”为重要内容的金融强国建设的机制和路径,并系统评估其在推动金融强国建设过程中的潜在风险与治理挑战,深入探讨技术变革背景下如何以人工智能更好赋能金融强国建设。

## 二、人工智能应用于金融领域的发展趋势

历史上,每一代信息技术都对金融体系产生了重大影响,并极大促进了效率的提高和金融创新的进步。从1943年开始,人工智能相关的技术发展至今已有80多年的历史,在经历了1956~1974年的第一次浪潮和1982~1995年的第二次浪潮后,

目前正处于从2006年开始的第三次发展浪潮中。从最初的简单规则策略发展到在实际场景中的应用,进而催生出算法模型,接着又从简单的算法模型升级成复杂模型,同时伴随着应用数据量的大幅增多。基于人工智能逐渐从规则系统到机器学习,再到深度学习的演进,其在金融领域的应用也经历了从“浅层智能”到“深层智能”、从“辅助决策”到“驱动核心业务”的革命性变化,增强了金融系统处理信息、分析数据、识别模型和预测的能力(Aldasoro et al., 2024)。尤其是2022年以来大语言模型的生成和发展,催生了一场人工智能领域从数据处理、特征工程到洞察解读的全链条变革,以智能体为代表的人工智能成为最新趋势。

### (一)专家系统是人工智能应用于金融领域的初步表现

20世纪80年代以前,人工智能技术以专家系统为主,专注于开发基于规则的系统,以模拟人类智能。具体而言,专家系统指的是一些行业专家提供专业的知识,将知识存储到数据库中备用,当用户提问时,系统可通过一些规则在数据库中匹配答案并返回给用户。这一时期的人工智能主要扮演“数字员工”角色,应用于最基本的金融功能,如风险管理、基本的算法交易规则、信用评分和欺诈检测等方面,主要价值在于降本增效,将人力从繁琐的劳动中解放出来,并未触及金融业务的核心决策逻辑。此外,这一阶段的人工智能技术应用过于僵化,在模型识别、不确定性处理和复杂推理等方面远未达到人类水平,其不能从新数据中学习和适应不断变化情境的特点使其无法适用于动态的金融领域,且需要人工监督。

### (二)机器学习促进人工智能在金融领域的应用拓展

20世纪80年代到21世纪初,对更具适应性和响应性系统的需求使人工智能发展到机器学习阶段,多层感知机、决策树、Boosting算法、Bagging算法和Svm算法等得到发展,人工智能开始可自主学习经验并执行任务,而无需明确阐述背后的规则,以实现自动优化、预测和分类,开始具备数据分析和模式识别能力,成为决策的智能助手。早期的机器学习依赖于大量结构化数据<sup>①</sup>,但由于这些模型会动态地适应新数据,通常无需人工干预,导致其在决策过程中的透明度较低。

除应用于金融业的基本功能外,机器学习的进步开启了人工智能在金融领域的一系列新应用,尤其值得关注的是其在资产定价领域的角色定位。价格预测是金融市场分析的核心,但由于用于预测分析的相关信息集通常非常大,很难将其纳入传统模型,叠加其严重依赖于缺乏共识的功能形式的基本假设,使金融市场的价格预测往往具有较低的效率和准确性。由于机器学习具有可以合并大量数据和基于灵活的非参数函数形式的优势,其被广泛用于预测收益、权衡风险回报和配置最佳投资组合,扩大了人工智能在金融领域的应用。

### (三)深度学习下的生成式人工智能在金融领域被广泛应用

21世纪以来,机器学习的成熟度大大提高,发展到深度学习阶段,使用多层人工神经网络模拟人脑学习方式,自动提取复杂特征。与传统机器学习相比,深度

<sup>①</sup> 结构化数据指的是以有序、定量形式存储于关系型数据库中且易于搜索的信息,如数字信息。

学习基于深度神经网络,具有无监督学习、泛化能力强、对大数据具有稳健训练能力等优点(Huang et al., 2020),不仅可以利用更多数据,还能够使用非结构化数据<sup>①</sup>进行运作,训练出更为复杂的模型,提高准确度。这一时期的人工智能具备了自主学习和复杂决策的能力,成为驱动业务创新的核心引擎。这一特点使金融领域的收益预测(Frankel et al., 2022; 马甜等, 2022)、股票交易(Li et al., 2019)、投资组合管理(Almahdi and Yang, 2017)和宏观经济预测(Chatzis et al., 2018)得到了进一步发展。

随着深度学习的快速发展,用于训练最新前沿人工智能模型的计算能力大大提升,更新换代也较为迅速,推动了以生成式人工智能为主线的人工智能的快速发展,尤其值得关注的是大语言模型(简称大模型)。大模型经过大量参数的训练,能够生成、分析和分类文本以及生成合成数据等(张晓燕和吴辉航, 2024)。相较于传统机器学习的技术进步,生成式人工智能尤其是大模型扩大了其在金融领域的应用范围,助力金融降本增效(廖高可和李庭辉, 2023)。一方面,大模型增加了数据使用的类型,更能满足拥有大规模实时数据资源的金融领域的需求。另一方面,大模型在处理开放性问题上具有优势,有助于提高金融领域的智能化水平。大模型不仅可以改善金融领域的后端处理、客户支持、机器人咨询和监管合规,广泛应用于智能办公、智能客服、智能营销、智能投顾、智能风控等诸多场景,还能够促进自动化任务的开展(Korinek, 2023)。

#### (四)智能体成为人工智能在金融领域应用的新趋势

智能体具有自主规划和决策、使用工具和记忆与反思的特点,可适用于业务复杂、对可靠性具有高要求、以及海量多模态数据的金融领域。目前,智能体<sup>②</sup>已初步应用于金融领域中重复性高、规则相对清晰的场景,核心作用在于提升整体运行效率而非替代决策,如客服运营、营销支持、信贷辅助、保险理赔等。由此可见,当下的智能体主要活动在金融领域的业务外围,更多是在既有流程框架内担任辅助角色。

目前的智能体已具有有限的规划能力,未来的技术发展可能会进一步促进人工智能发展范式跃迁,产生更高水平的智能体系统,并促进人工智能从技术想象回归到业务现实、从能力展示走向价值交付。金融智能体可能是未来的发展方向,其应用将从辅助行为扩展到金融领域的更深层次,从而重塑金融科技的竞争逻辑。

### 三、人工智能重塑金融功能的理论逻辑与实践场景

随着人工智能技术的快速迭代,金融体系的运行逻辑正在经历深刻的重塑。

<sup>①</sup> 非结构化数据是指未依据预先设定的模型进行组织的信息,如文本、音频和视频信息。

<sup>②</sup> 智能体指的是建立在大模型上的人工智能系统,具有规划能力,拥有长期记忆,通常还可以访问外部工具,从而使之达到自主交易水平,不需或较少地需要人为干预或规范。

在数字化转型与智能化升级持续推进的背景下, AI不仅改变了金融机构的业务流程、服务模式与风险管理框架, 也以更为根本的方式触及金融体系的核心功能, 引发学界与监管部门对“AI是否会改变金融功能”这一核心问题的重新讨论。

回顾金融学的发展脉络, 金融功能理论为理解这一问题提供了系统且具有解释力的分析框架。早期学者如 Gurley and Shaw (1956)、Goldsmith (1969) 等从金融中介、风险转换和制度演进角度探讨金融体系的基本作用, 而 Levine (1993) 则进一步强调金融深化、信息效率与经济增长之间的内在关联。Merton and Bodie (1995)、Merton and Bodie (2004) 提出的金融功能观认为, 金融体系的核心不在于机构形态、组织结构或工具类型的表层差异, 而在于其长期稳定存在的六大基本功能, 即支付清算、资源归集与股份分割、跨时空资源转移、风险管理、提供信息与价格发现以及激励约束。正因金融功能具有“第一性原理”特征, 金融机构的演化、市场结构的变迁和技术创新的应用皆围绕这些功能展开。金融体系的形态会发生改变, 但金融功能却具有长期稳定性。

人工智能对金融的影响需从金融功能的重塑而非替代视角来理解。AI技术的核心优势在于其在大规模数据处理、非结构化信息解读、复杂模式识别和动态预测上的显著边际收益, 这些能力改变了金融体系的运作方式, 却并未动摇金融功能的本质结构。AI并不能消除跨期风险、终结信息不对称或替代金融中介的治理职能, 但可以通过机器学习、自然语言处理、图网络模型等技术路径, 降低信息处理成本、提升风险识别精度、扩大金融服务边界, 从而在深度与广度方面增强既有金融功能。随着AI在信贷审批、资产定价、智能投顾、交易监控、合规审查与金融基础设施等领域的全面渗透, 它正在以一种渐进而系统的方式重塑金融功能, 使传统功能在效率、精度、速度与覆盖范围上均显著提升。

### (一) 支付清算

在现代金融体系中, 支付清算被普遍视为最基础、最具公共产品属性的核心金融功能之一。广义而言, 支付清算功能是指在商品交换、资产交易和债权债务关系中, 实现货币与资金在不同经济主体之间安全、有序和最终结算的过程。在我国, 它主要依托由银行账户体系、大额实时支付系统、小额批量清算系统以及各类证券和衍生品中央对手方在内的金融市场基础设施共同构成的网络结构。一旦支付系统在技术、流动性或信用上出现中断, 支付清算体系风险极易通过“多米诺效应”被放大为跨市场、跨机构的系统性危机。

在理论层面, AI通过提升对高维支付数据的建模能力、对复杂网络结构的刻画能力以及对异常模式的识别能力, 强化了支付清算功能的安全性、稳定性与效率。一方面, 人工智能对支付清算功能的影响体现在其重塑信息处理和风险识别机理。支付清算本质上是对海量高频交易记录的实时验证、路由选择与最终结算过程。近年来, 机器学习与网络分析等AI工具在支付系统监管和运行维护中, 主要用于对高风险交易、异常模式及潜在脆弱性进行实时监测与评估, 从而提升系统稳定性与安全性。另一方面, 在高度不平衡的欺诈数据环境下, 深度学习与集成学习方法在

捕捉稀有异常模式、降低误报率和提高可解释性方面具有相对优势,这一结论可推广至更广义的支付清算风控场景。

在实践层面,人工智能对支付清算功能的重塑已经在多个典型场景中体现出来。其一,在零售支付领域,银行与第三方支付机构广泛采用基于机器学习的实时反欺诈系统,对交易金额、时间分布、地理位置、设备指纹、历史行为等多维信息进行联合建模,在毫秒级别输出风险评分,实现“交易-风控-授权”一体化的嵌入式安全机制。相较于传统规则引擎,AI模型在降低误报率、提高识别新型欺诈模式方面具有明显优势,从而在不显著增加用户摩擦的前提下,提高了零售支付系统的整体安全边际。其二,在大额实时支付系统与批量清算系统中,AI被用于流动性预测、清算顺序优化和异常行为监测。其三,在跨境支付与多币种结算场景中,AI与合规科技结合,用于反洗钱、制裁筛查和复杂网络结构中的可疑交易识别。BIS创新枢纽的“Project Aurora”等项目通过将网络分析与机器学习结合,对跨境支付网络进行图结构建模和异常子图挖掘,提高了对跨境洗钱网络的识别效率。此外,在央行数字货币和“可编程支付”探索中,AI可以在交易监控、智能合约风控以及匿名性与可追溯性平衡等方面发挥关键作用,使支付清算功能在充分利用技术优势的同时,兼顾金融稳定与隐私保护等公共目标。

## (二)资源归集与股份分割

资源归集与股份分割功能是金融作为“信用组织者”和“资本形成机制”的基本体现。其内在逻辑在于:金融体系通过中介机构、市场结构和契约设计,将分散在家庭与企业部门中规模较小、期限各异、风险偏好差异化的资金资源集中起来,并依托基金化、证券化和公司制等制度安排,将整体融资需求切割成风险可分担、期限可管理、流动性可选择的“份额”,从而促进资本形成与经济增长。

在理论层面,人工智能主要通过三个机制重塑资源归集与股份分割功能。首先,在资产定价与风险测度领域,AI通过处理高维、多源与非结构化数据,将归集而来的资金基于更丰富的信息结构用于资产配置,从而提高“分割后的股份”被有效持有的可能性。大型资产管理机构正在采用深度神经网络和Transformer模型构建“资产嵌入空间”,通过学习跨资产类别的时间序列、文本披露、供应链图谱以及市场微观结构数据,捕捉传统因子模型难以识别的风险因子与隐含关联。其次,在投资者异质性建模方面,AI的突破显著提升了“谁应当持有哪些份额”这一功能的配置效率。传统模型依赖静态问卷和风险等级划分,难以捕捉投资者在行为、收入、生命周期、职业结构及宏观情绪变化中的高频动态特征,而机器学习的引入使投资者画像能够基于历史交易数据、设备习惯、文本内容和社会网络信息等多类型数据源进行实时更新。AI已经使金融机构能够构建更精细的投资者分类体系,实现从标准化产品到“个性化份额结构”的过渡(Gu et al., 2020)。第三,在市场微观结构与交易成本维度,AI通过算法交易、智能订单路由与流动性预测,显著降低股份在投资者之间交换的摩擦成本,改善分割后份额的可交易性与流动性结构。

在实践层面,人工智能对资源归集与股份分割的影响已在多个领域显现。其一,在财富管理与资管领域,智能投顾已经成为归集中小投资者资金的重要载体。AI模型能够对宏观变量、因子暴露、资产相关性和投资者行为进行综合建模,将大量分散的资金纳入标准化或半定制化投资组合中,再通过动态再平衡与智能风控实现“股份的实时分割与调整”。其二,在公募基金、ETF与指数化投资领域,AI已被用于指数权重优化、主题指数构建和“智能Beta”策略开发,使归集资金以更复杂、更具适应性的份额形态进行分割。其三,在股权众筹、数字证券发行与资产代币化领域,AI的引入扩大了可归集与可分割的资产边界。众筹平台通过AI对企业质量、投前评价、市场舆情和投资者偏好进行综合分析,降低了逆向选择风险,也提高了小微投资者的项目选择效率。在代币化场景中,AI则被用于资产定价、欺诈识别、自动化合规以及智能合约执行,使原本难以分割的资产(如不动产、艺术品、基础设施项目)可以以高频流通的小额份额形式参与市场。其四,在大型机构投资者的多资产配置领域,AI模型帮助养老金、主权财富基金等机构在庞大的资产集空间中寻找最优的配置组合,使归集的大规模长期资金能够以更具耐心和更符合负债特征的份额结构进行分配。

### (三)跨时空资源转移

跨时空资源转移功能是通过金融契约、信用机制与期限结构安排,使经济主体能够在不同时间、不同经济状态之间配置资源,从而平滑消费、优化投资决策并更好应对不确定性。其本质是“将资源在时间维度上以更低的成本、安全地进行转移”。

在理论层面,人工智能通过增强金融机构的信息处理能力、改善风险评估与定价机制、优化期限结构管理和预测能力,强化了跨时空资源转移功能的效率。首先,在贷款定价与违约预测方面,AI对高维信用数据、交易行为数据、社交网络数据及非结构化数据(文本、图像等)的处理能力显著提高了信用资源跨期配置的准确性。机器学习能够显著降低违约预测误差,使贷款期限结构与利率定价更能反映借款人在未来时期的收入风险,从而提高跨期信贷配置的资源效率(Berg et al., 2020)。其次,在储蓄-投资行为方面,AI在资产收益预测、风险因子识别与资产生命周期管理中的应用,使长期投资(如养老金、保险资金)能够更准确预测未来回报与风险,实现跨时间的资源优化配置。最后,在期限结构管理方面,AI的动态预测能力、场景生成能力和压力测试功能,使金融机构能够更有效地识别未来的利率风险、流动性风险与系统性压力点。

在实践层面,人工智能对跨时空资源转移功能的影响体现在多个金融应用场景。其一,在个人与家庭部门,AI支持智能投顾、生命周期投资规划工具和行为预测模型,使居民储蓄与投资行为从静态配置向动态、数据驱动的跨期规划转型。例如,智能投顾根据年龄、收入路径、风险偏好变化和宏观环境,自动在不同时间点调整股票、债券、货币基金等资产的比例,实现“生命周期跨期资源配置”。其二,在消费金融领域,AI基于行为模式识别和实时信用评分机制,使金融机构能够更准确

评估借款人在未来时期的收入稳定性与违约可能性(Gambacorta et al., 2024),从而降低逆向选择,使信用资源更有效从“未来收入”转化成“当前可用资源”,实现消费的跨期平滑。其三,在养老金、保险等长期投资领域,AI被用于寿命预测、健康风险分析、人口结构动态模拟以及长期资产收益风险预测,显著提高了长期准备金管理、保险定价与养老金投资策略的科学性。

#### (四)风险管理

风险管理功能是通过多样化的金融工具、契约结构与制度安排,使经济主体在面对不确定性时能够有效识别、分散、转移和承受风险。风险管理并不依赖金融机构的特定形态,也不因市场结构演化而改变,其本质在于通过风险分散、保险机制、衍生品市场与资本结构设计,使潜在不确定性能够在更大的时间与主体范围内被承载。

在理论层面,AI主要通过三个机制影响风险管理功能。首先,在风险识别机制上,AI有效突破了传统金融模型在特征数量、复杂度与非线性关系识别方面的约束。机器学习在捕捉企业违约概率、市场波动源以及跨市场关联结构方面具有显著优势,能够识别传统逻辑回归、VAR模型难以捕捉的潜在风险因子与尾部风险(Danielsson et al., 2022; Berg et al., 2020)。AI利用多源数据构建风险特征,不仅提升了风险因子的解释力,也改变了风险识别的理论基础。其次,在风险预测与压力测试机制上,AI能够通过复杂的非线性模型模拟未来状态路径与极端情景。AI在预测流动性枯竭、市场波动聚集、跨金融机构传染路径等方面具有明显优势,尤其是在高频交易网络中,其对异质性交易者行为的建模能力显著强于传统模型。AI支持的情景分析与压力测试,不再依赖有限的宏观冲击假设,而是通过生成模型自动生成数百种可能的情景,增强了系统性风险评估模型的稳健性。最后,AI通过复杂网络分析增强了系统性风险识别的能力。传统的系统性风险分析依赖相关性矩阵与宏观杠杆指标,而AI能够基于银行间贷款网络、跨市场交易网络和清算网络的结构特征,识别金融机构之间的隐性关联度与潜在传染路径。

在实践层面,人工智能对风险管理功能的重塑已在金融机构、金融市场与监管体系中全面展开。其一,在信用风险管理中,银行与消费金融机构普遍采用机器学习模型进行信用评分、违约预测与额度管理。AI模型利用借贷行为、现金流模式、设备信息、地理轨迹和社交网络等多种特征,对借款人未来还款能力进行更精准的估计,使信用风险定价与额度分配更加精细化。其二,在市场风险管理中,AI被用于预测波动率、识别结构性市场风险和构建实时风险预警系统。许多资产管理机构利用AI构建“波动率预测模型”与“尾部风险模型”,从期权报价、订单簿深度、涨跌幅路径结构和高频交易行为中提取风险信号,显著增强对极端风险的预测能力。其三,在流动性风险管理中,AI模型能够分析交易网络与资金流动模式,预测市场流动性枯竭的节点。其四,在反洗钱、反诈骗与合规管理中,AI的应用尤为广泛。通过自然语言处理与图网络分析,AI能够识别复杂交易模式、追踪资金

链条,并检测隐匿在跨境支付体系中的可疑关系网络。其五,在保险与精算领域, AI被用于事故预测、疾病风险预测与赔付概率模型,特别是在健康险、车险等场景中, AI模型可以分析用户行为、医疗图像和传感器数据,从而形成更精准的风险分层。

#### (五)提供信息与价格发现

提供信息与价格发现功能在于通过交易、资产定价与市场透明度机制,使经济主体能够从分散的信息结构中提取有价值的信号,从而正确评估资产的内在价值、未来收益与风险状况,实现资源配置的效率提升。无论市场结构如何演化,价格发现功能都通过“信息收集—交易互动—价格形成”的链条呈现出稳定性:信息使市场参与者能够形成预期,交易使信息得以反映到价格上,价格再引导资源流向更具生产率的领域。

在理论层面, AI对信息生产与价格发现功能的影响主要体现在三个方面。首先, AI在信息生产环节的最大贡献在于突破了传统金融模型只能处理结构化数据的限制,使信息生产从有结构、有标签的数据拓展到文本、图像、语音、社交网络、供应链图谱等多模态信息。自然语言处理技术在财报分析、监管文件解读、新闻情绪提取和社交媒体舆情监测中被广泛应用,使金融市场能够实时吸收原本难以量化的“软信息”。这些技术的引入,增强了市场对“真实价值”的感知能力,提升了价格信号的质量。其次,在价格发现机制中, AI的作用主要体现在提升预测能力、增强定价效率与捕捉非线性关联。传统资产定价模型依赖有限因子与线性框架,而 AI能够处理数百甚至上千个因子,同时识别复杂的非线性结构,使价格预测更具准确性。最后,从市场微观结构角度看, AI改变了交易行为与订单簿动态,促使价格形成更加高效。 AI驱动的算法交易与智能订单路由在全球资本市场中已成为主流,能够根据实时价格、成交量与市场深度动态调整交易策略,从而降低冲击成本并增强市场连续性。

在实践层面, AI对信息生产与价格发现功能的重塑已在股票市场、债券市场、衍生品市场乃至新兴数字资产市场中显现。其一,在股票市场, AI广泛用于财报自动解读、分析师报告生成、舆情监测与市场情绪指数构建,使投资者能够实时掌握公司价值变化与市场信号。大型机构投资者通过 NLP 模型分析上市公司电话会议纪要、公告文本与媒体报道,从而形成对公司经营状态的前瞻性判断。其二,在债券市场, AI被用于信用评级、债券违约风险预测与宏观风险分析,使债券价格能更快反映借款者信用变化。特别是在企业债与地方政府融资平台债券市场, AI对财务报表质量、隐性担保结构与关联交易的识别能力,使市场更容易定价隐藏风险。其三,在衍生品市场, AI被用于波动率曲面预测、期权隐含波动率建模与期货市场行为识别,使市场能够更好地反映风险中性预期。其四,在外汇与大宗商品市场, AI基于多源宏观数据、地缘政治信息与供需变化自动构建预测模型,使价格发现更具全球视野与实时性。其五,在数字资产与区块链市场, AI通过识别链上交易网络、异常资金流、智能合约风险和矿工行为,使高匿名性环境下的价格发现机制更

加透明。

#### (六)激励约束

激励约束功能的理论基础是代理理论、信息不对称理论与契约理论。具体而言,信息不对称导致逆向选择,隐性行为导致道德风险,不完备契约导致监督成本上升,进而影响金融资源配置效率与金融稳定。金融体系通过设计各种制度机制(股权结构、债务契约、风险分担机制、监督与治理制度、审计与披露制度等),使资金提供者与资金使用者之间的激励保持一致,降低信息不对称成本,从而确保金融交易得以安全实施;这一“激励约束功能”并不因金融机构结构或金融工具的变化而改变。

在理论层面,AI对激励机制的改善主要体现在三个互相关联的机制。其一,AI在信息处理能力上的优势显著降低了逆向选择问题。传统方法依赖财务报表、信用记录或静态指标,信息维度有限且易被操纵,而AI可以利用多源数据开展高维度的信息收集,包括交易行为、合同履行记录、生产数据、供应链网络、新闻文本、舆情信息以及社交网络数据等。机器学习在信用评分中能够捕捉借款人未披露的信息,从而降低逆向选择概率。其二,AI显著提升了监督能力,从而降低道德风险。道德风险来源于行为无法被完全观察或监督成本过高,导致管理者、借款者或金融机构可能采取对自身有利而对资金提供者不利的行为。AI的实时监测、异常检测与行为分析能力,使信息可得性由离散、滞后变为连续、动态。其三,AI通过提升契约执行效率与合规能力,强化了激励兼容安排的约束力。传统契约存在不完备性,原因包括可观察性限制、执行成本高、契约条款难以处理不确定性等。AI技术结合区块链与自动化执行机制,使契约审查与执行更为精细、准确。例如,AI用于识别契约条款间的矛盾、缺失与潜在漏洞,减少契约不完备性导致的激励偏差;AI可以辅助自动化执行契约条款、识别风险行为并触发强制措施,使契约不再依赖人工干预,从而在交易执行层面加强激励一致性。

在实践层面,人工智能对激励约束功能的影响已经体现在多个金融场景。其一,在银行信贷领域,AI驱动的贷后管理系统能够实时监测贷款企业的生产经营状况、资金流动趋势与财务健康状况,使信贷关系从“静态契约”转变为“动态监督关系”,显著减少企业在获得贷款后从事高风险行为的可能性。其二,在资本市场监管中,AI辅助的市场监控系统可实时识别异常交易行为、操纵策略与内幕交易网络,使监管成本大幅下降,也提升了违规行为的发现概率,从而对市场参与者产生更强的激励约束效果。其三,在保险市场中,AI在欺诈检测、风险定价与理赔审核中的应用,使保险产品能够在设计阶段内生地抵御道德风险。例如,通过分析医疗记录、历史理赔数据与行为数据,AI可以识别高风险投保行为并调整保费,从而减少逆向选择。其四,在智能合约与去中心化金融(Decentralized Finance,简称DeFi)中,AI辅助的自动化合规与执行机制正在推动“激励自动化”:合约违反、资金挪用等风险可通过自动化措施触发,从而减少人为操纵与道德风险。

需要注意的是, AI的应用不仅在边际上改善了单一金融功能的效率,更通过改变信息生产与交易成本结构,重塑了不同金融功能在金融体系中的相对地位与相互关系。例如,在传统金融中介理论中,支付主要被视为资金转移的通道,但在AI环境下,支付系统因其具备显著的网络外部性和高频数据特征,演变为降低全系统信息不对称的关键机制。这意味着支付功能不再仅仅是交易的终点,而是凭借其产生的高维数据积累,成为提供信息、风险管理以及资源配置等高阶功能的逻辑起点与数据基础,演变成聚合其他金融功能的一个枢纽。

与此同时, AI技术的广泛渗透显著增强了金融功能之间的互补性,使得传统上相对独立的金融功能界限日益模糊。AI等技术进步使得信息的收集与处理成本大幅下降,促使金融服务从分散走向融合(Goldstein et al., 2019)。例如,通过自然语言处理与机器学习技术,即时的信息处理能力使得风险管理不再是独立于信贷决策之外的事后监控,而是内嵌于跨时空资源转移过程中的实时机制。这种不同功能之间的深度融合意味着,信息的生产、风险的定价与资金的配置在微观层面实现了同步完成。AI推动的金融功能融合,有助于从更深层次提升金融体系配置资源的效率与金融稳定性。

#### 四、人工智能赋能金融强国建设的机制与路径

做好金融“五篇大文章”是加快建设金融强国的重要方面,也是金融服务实体经济的主要内容。以高质量数据集为基础、以不断迭代创新的算法为动力、以持续扩张的算力基础设施为支撑,人工智能应用于金融领域能够进一步改善金融服务效果,提升金融服务实体经济的深度和广度。目前,人工智能尤其是生成式人工智能在我国金融领域的应用还处于初级阶段,要聚焦科技金融、绿色金融、普惠金融、养老金融和数字金融“五篇大文章”的不同发展特点,提供针对性的技术运用,实现人工智能在重点领域的价值提升。

##### (一) 赋能科技金融

##### 1. 人工智能带来全新的科技金融模式

作为一种旨在通过创新金融工具和服务以服务于科技型企业的金融模式,科技金融源于高度不确定的科技创新与传统金融体系之间的不匹配。一方面,我国科技创新的目标与内涵具有鲜明的国家战略属性。我国的科技创新,尤其是面向“卡脖子”技术的攻关,在微观层面可保证产业链供应链安全、摆脱关键领域对外依赖,在宏观层面被视为驱动经济增长模式从要素投入向全要素生产率提升转型、塑造新发展格局的核心引擎,有助于推动科技强国的建设和新质生产力的发展(张明喜等, 2018)。

另一方面,我国金融体系以银行为主、风险投资发展程度较低的特点,难以满足科技创新的融资需求,形成了三重错配。一是风险收益错配。银行信贷与“卡脖子”技术所需的长周期、高风险、高不确定性的资本本质不匹配。二是评估体系错

配。传统金融依赖历史财务数据和抵押物,而科技型企业的核心价值在于知识产权、人力资本和未来场景,难以用传统模型量化。三是市场生态错配。科技型企业因具有研究周期长、资金投入多、风险大的特点,尤其需要耐心资本的支持和投入,且在初创期、成长期、成熟期和衰退期四个阶段均具有个性化的融资需求。但在金融体系中长期资本(耐心资本)匮乏、退出渠道单一、专业早期投资机构不足,无法形成支持科技创新的全周期资本链。

人工智能不仅为科技型企业金融服务的改善提供了可能,而且改变了科技创新的方法和内容,由此可能带来全新的金融科技模式。其一,人工智能的应用使得我国科技金融体系更加智能与便捷,从而为科技型企业提供高效的金融服务。通过利用机器学习整合企业的非财务数据,为科技型企业的知识产权和无形资产定价,构建更全面的科技型企业信用画像,为不同企业精准匹配资金需求,助力信用评估和智能风控。利用人工智能自动阅读和分析海量的专利文档和行业报告等,实时监控行业新闻、竞争对手动态与政策变化,自动预警可能影响科技型企业的风险事件,全流程自动化以降本增效和管理风险。

其二,人工智能自身也正成为科技创新的核心驱动力,改变了科技创新的方法与内涵。人工智能不仅是金融服务的辅助技术,其发展本身就是前沿的科技创新,从而催生了全新的金融科技范式,也对金融服务提出了新需求。

## 2. 加强人工智能与科技金融的深度融合

基于未来的发展方向,依托金融大数据和人工智能技术的发展,加强人工智能与科技金融的深度融合,不仅发挥人工智能在科技创新方面的引领作用,还在金融更好服务科技型企业方面提质增效,这对于中国抢占新一轮科技革命和产业变革制高点至关重要。

其一,科技产业认知更加智能化,有效缓解信息不对称,降低风险评估难度。相较于传统方法,人工智能的应用不仅有利于拓宽资金提供者的信息来源,还提升了获取维度更多元、覆盖范围更全面信息的能力(刘少波等,2021),从而助力获取“硬信息”和“软信息”(Liberti and Petersen, 2019)。人工智能要在关联既有信息和数据的前提下,更深入地理解科技型企业的技术演化逻辑和产业发展趋势,同时提高知识产权和无形资产等的估值精确度,真正判断科技型企业的价值和风险所在,从而攻克“技术价值难衡量”和“专业壁垒高”的问题,减少科技金融发展面临的信息不对称问题。

其二,金融产品自适应与生成式,提高科技金融服务的针对性。科技金融要有针对性地为科技型企业提供全链条、全生命周期的金融服务。人工智能可将“产品匹配”升级为“产品创造”,利用生成式人工智能,根据每个科技型企业的独特生命周期节点、现金流模式、风险敞口等,提供最优的、最合适的个性化、全周期融资方案。

其三,风险控制全周期和自动化,大幅降低服务成本。科技的运用将减少机构服务单一客户的成本。借助机器学习算法实现全过程自动化管理,并自动触发决

策机制,降低人力的服务成本。

## (二)助力绿色金融

### 1.人工智能提高绿色金融的“市场驱动”性

在“双碳”目标和经济高质量发展的背景下,传统金融服务难以满足经济向绿色低碳的转型升级。一是减排活动在金融实践中难以计量和验证。一方面,减排数据多来源于企业主动披露和第三方人工抽查,对投资者而言缺乏实时、统一的绿色活动监测信息,且披露内容不规范,标准不统一,部分企业夸大和伪造环境效益,存在“洗绿”与“漂绿”风险。另一方面,由于缺乏权威、可追溯的绿色项目认定标准(陈雨露,2018),减排活动界定与标准不统一,难以横向比较和验证。二是外部性问题内部化不足。经济活动产生的环境成本或收益没有被准确计算和纳入到市场交易价格中,导致金融体系对绿色领域的投融资功能难以实现。三是缺乏有效的投融资渠道。绿色资金供需之间“桥梁”不足或不通畅,资金与项目并未高效、精准、可信地匹配,如投资回报周期与激励机制错配,金融产品和服务仍需进一步丰富。

绿色金融的提出源于国家致力于引导资本流向绿色低碳领域,防止“洗绿”,确保金融真正能为环保、节能、清洁能源等绿色项目提供充分高效的投融资服务,通过政策刺激提高资金的导向性和精准性,从而促进企业绿色转型和发展(张明和陈骁,2024)。而人工智能与绿色金融的结合,可通过解决数据可信、风险可量化、效益可定价等根本问题,为绿色金融从“政策驱动”转向“市场驱动与创新驱动”提供核心支撑。第一,环境数据处理智能化。人工智能技术不仅可以监测森林覆盖、土地利用变化、工厂排放等情况,还能整合传感器网络数据,实时监控企业能耗和排放信息。公共数据和企业数据的结合提高了大数据分析能力,使投资者能够建立可计算和可验证的绿色数据流,从而较准确地识别和验证真正的绿色项目。第二,动态风险定价模型优化。人工智能通过融合多维数据,动态量化绿色项目综合风险,将环境因子嵌入信用评价模型,使绿色表现与资产定价、融资成本动态挂钩,提供更精准的利率定价和信贷额度建议,提升市场机制在绿色金融领域的功能发挥。如中国工商银行建立的气候风险数据库,将气候风险纳入智能化风控系统。第三,资产标准化和流动性提升,拓宽投融资市场。借助人工智能技术,有利于实现对小微、分散绿色资产的聚合与证券化,促进对成千上万底层资产的穿透式、自动化管理,降低证券化后的管理风险和成本,增强投资者信心,从而拓宽绿色金融体系中的投融资渠道。蚂蚁数科将能源基础设施及实体电力资产接入其区块链平台“蚂蚁链”,通过“区块链+物联网”技术追踪新能源设备并探索代币化融资,就是人工智能技术应用于绿色金融投融资领域的典型案例。

### 2.驱动人工智能应用于绿色金融的核心环节

基于人工智能技术在绿色金融领域的辅助角色,未来应继续驱动人工智能技术应用于绿色金融领域的核心环节。可从两方面入手增强人工智能技术的嵌

入深度。一方面,促使人工智能驱动验证机制直接嵌入绿色金融交易的核心。其根本路径是将验证流程从交易外部的独立审核环节,深度内化为交易结构本身的智能执行条件与自动化决策基础,通过提供实时、可审计的减排数据流,生成专业的绿色数据自动验证即服务平台。另一方面,将数据资产融入绿色金融定价的核心。金融科技有助于丰富绿色金融产品数量与种类,构建多元化投融资主体。其不仅可以企业实时环境表现数据纳入银行信贷模型和保险精算模型,将可信的减排量数据作为碳市场交易、绿色债券利率浮动的直接依据,同时还可以开发可持续发展挂钩贷款和环境效益证券化等复杂金融产品和服务。此外,推动绿色金融标准演进也是重要路径之一。借助于人工智能监测的环境效益数据,可以动态优化绿色分类标准,并自动比对不同司法管辖区的绿色标准,促进国际互认。

### (三)优化普惠金融

#### 1.人工智能解决普惠金融的核心矛盾

作为实现金融强国的重要力量,普惠金融源于金融机构的“商业可持续性”和服务对象的“高风险、低利润”特性之间的结构性矛盾。首先,金融机构的运行基于风险和收益相匹配,其信贷决策依赖于稳定的现金流、充足的抵押品和清晰的信用历史。然而,部分群体如“三农”、小微企业和低收入人群,普遍缺乏传统的“硬信息”和合格抵押物,使得金融机构在为其提供服务时,面临信息高度不对称和风险难以量化的双重困境,倾向于采取更为审慎甚至回避的态度,以保障自身的资产安全与商业可持续。其次,为“小而散”的客户群体提供金融服务通常意味着高昂的运营成本。每一笔小额贷款的贷前调查、风险评估、贷后管理和催收成本导致业务的单位成本高昂。如果以市场化的风险溢价来覆盖成本和风险,则会推高利率,从而导致传统金融体系在纯市场驱动下,难以形成可持续的商业模式。最后,传统金融体系(尤其是大型商业银行)是在服务大企业、大项目和大城市的过程中建立和发展起来的,其组织架构、业务流程、风控模型和技术系统都是为标准化、大批量的业务所设计,难以适用于服务异质性极强、非标准化的客群。

普惠金融强调金融资源的优化配置,侧重于解决金融资源分配不均的问题,以保证金融服务能够覆盖更多的地区和人群,尤其是在传统金融服务中被边缘化的群体(孙玉环等,2021)。在普惠金融的发展中,科技尤其是人工智能成为重要推手,为解决普惠金融的核心矛盾提供了新的动力,推动普惠金融从解决“有没有”的基础覆盖问题,向实现“好不好”的高质量服务阶段升级,通过技术赋能弥补传统金融服务在触达范围、服务效率与可持续性方面的短板,不仅推动金融服务边界的持续拓展,也促进金融服务场景的深度融合与创新,使普惠金融更加可负担、可触及和可持续。一是赋能风险定价。通过将人工智能应用于多维度数据挖掘,构建“数字信用”,使其能够替代传统抵押物。二是驱动降本增效。使用人工智能实现业务流程的自动化,大幅降低运营效率,以在保证利润下扩大金融服务的规模覆盖。三

是提供个性化服务。智能分析客户需求,达到从提供基础金融服务到形成金融伙伴关系的升级。

## 2. 推动“数字普惠”到“数智普惠”

现阶段的人工智能应用,仅发挥统计测算功能,将模型分析结果作为传统人工风控判断的辅助工具,并未在模型的修正或预测方面提供更有效的帮助。突破人工智能在普惠金融领域的初步应用,以数智之力攻克普惠金融“最后100米”挑战是人工智能未来的重要应用场景之一。第一,发挥人工智能对资源配置的智能化功能。综合运用移动互联网、智能物联网、区块链等技术依法依规采集多元数据,加强金融基础设施建设,补足普惠群体全链条“数字足迹”,消解传统金融服务的边界与壁垒,推动金融从服务少数群体更深入地走向赋能整个社会。第二,优化算法模型。未来的“数智普惠”不仅是流程线上化,更是“AI+数据+业务”的深度重构。通过多模态历史经营和市场数据的挖掘,探究金融业务规律,进一步释放人工智能在风险评估、信用判断中的价值,为普惠金融信任构建提供技术支撑,助力纾解“信用白户”难题(曾燕等,2024)。如通过人工智能与数字人民币的结合,精准追踪贷款用途,降低特定领域融资成本。第三,利用人工智能技术实现产品定价的差异化 and 个性化。基于生成式人工智能和推荐算法,构建个性化信用服务模式,提高金融产品个性化和差异化的可行性,让普惠群体能用、会用、敢用金融工具和服务。

### (四) 支持养老金融

#### 1. 人工智能融入养老金融全链条

在人口老龄化的现实挑战下,发展养老金融是完善和健全我国金融市场、有效应对人口老龄化现象和推动养老事业高质量发展的重要举措。现阶段,我国的养老金体系存在过度依赖第一支柱、第二支柱覆盖面窄和第三支柱实际缴存率低的问题,导致三大支柱发展不平衡。同时,金融体系对资金供需双方而言都存在服务质效不高的问题。一方面,养老产业作为资金需求方难以得到有效融资,传统金融供给难以满足资金需求大、风险高、回报周期长和抵押物少的养老产业发展。另一方面,养老金作为资金供给方面临投资渠道有限、效率不高和风险管理不足的困境,金融产品的低创新性既无法满足养老金的投资需求,也未能覆盖居民多样化的财富管理和保障需求。

基于以上现实问题,我国养老金融发展有三个方面的目标:在宏观层面,发展养老金三支柱,将个人养老责任从国家主导向“政府+市场+个人”多元共担模式转变;在中观层面,鼓励金融资源更多向养老产业倾斜,培育养老服务经营主体和产业;在微观层面,满足居民的全生命周期多样化资金需求,保障老年人的金融安全和权益。因此,人工智能与养老金融的结合可分别赋能养老金金融、养老产业金融和养老服务金融,以全链条融入促进养老金融体系的发展。一方面,人工智能渗透于养老金融的产业与服务中,并与其他要素(如金融软件、金融实体机构、养老服务人才等)相协同,利用人工智能技术生成与交互特性提高用户的技术接受度。另一

方面,人工智能通过机器学习,利用大数据与算法分析用户行为,实现部分金融与养老人力资源的替代,通过对传统养老金融行业的变革,促进新兴养老金融行业的创新发展。

具体而言,一是从数据整合和场景落地两方面夯实养老金融的可得性。通过智慧销售、智慧投资、智慧风控、智慧运营和智慧客服等多个产品的应用,人工智能大幅提升金融服务的效率、降低成本,从而降低养老金融的服务门槛,并利用大数据分析为客户提供“一人一策”的资产配置与养老规划建议,提升养老金融产品的透明度与适配性。二是赋能养老产业和照护服务。人工智能除了通过应用于信用评价和风险控制等领域为养老产业提供资金供给外,还能直接赋能养老产业金融,展示了人工智能助力金融机构在养老金融领域形成新突破的能力。

## 2. 人工智能与养老金融深度结合

未来的人工智能技术要与养老金融深度结合,真正为解决其结构性难题提供系统性方案,支持其实现跨越式发展。针对养老金三大支柱不平衡的问题,人工智能要通过精算模型增强第一支柱的可持续性预测,利用智能匹配降低中小企业建立年金的门槛以扩大第二支柱覆盖,通过个性化“养老画像”和智能投顾提升第三支柱的参与度和缴存率,并开发更贴合居民需求的养老金融产品。在破解养老金融供需错配方面,人工智能需构建基于多维度运营数据的信用模型,破解养老产业因缺乏抵押物面临的融资困境(李金金和曾圣钧,2025);同时,作为资金供给方的养老金也可借助人工智能技术优化资产配置,使用机器学习处理宏观长期变量,实现与负债匹配的投资组合,并通过另类资产智能筛选拓宽投资渠道。

### (五)推进数字金融

#### 1. 人工智能全面重构数字金融

作为金融科技运用于金融业务中实现的最终形态,数字金融主要指的是利用数字技术在支付、贷款、保险、投资等方面提供新的金融工具、金融模式和金融基础设施建设(黄益平和黄卓,2018),其重点在于使用数字技术手段为实体经济提供金融服务,而不在于划分具体的服务对象。在建设“数字中国”战略引导下,数字金融的提出不仅源于传统银行面临的数字化转型压力,还在于用户对智能、普惠和个性化金融服务的期待,从而致力于构建起现代数字金融体系。人工智能应用于数字金融领域,不仅是作为效率工具,更是其核心的驱动力和定义者,推动数字金融从渠道的数字化转向以智能为核心的全面重构,尤其表现在助力金融机构数字化转型方面。具体而言,人工智能是当前金融机构降本增效、优化风险管理及提升客户体验的重要引擎,其应用贯穿于获客、风控、运营、合规等全流程,依赖交易、行为、社交、物联网等多维度数据实现决策优化。如通过开设智慧网点、开发手机 App、推动信贷业务线上化、加快推动数字供应链金融业务等方式,人工智能尤其是生成式人工智能,广泛应用于智能客服、数字信贷、量化投资和智能投顾等方面,助力金融机构不断拓展业务范围、降低服务成本、创新服务模式、改善客

用户体验、创新金融产品供给以及提高风险管理能力,助推全流程、全场景的数字化转型。

## 2. 推进人工智能全面服务于数字金融

现阶段人工智能的应用聚焦于金融机构的数字化转型,局限于辅助决策功能,且在算法和数据的局限性以及算法的安全性方面存在风险。人工智能在数字金融领域的未来应用,除聚焦于金融数字化转型之外,还应增强金融服务实体经济质效,以及提高金融基础设施的服务能级。

其一,将人工智能作为技术推动金融数字化转型的主动动力。通过整合金融机构的技术与数据投入,优化现有资源配置,改善数字化的投入产出,从而做到以数据为驱动、以算法为支撑,打造全链条数智化运营模式,健全多样化、专业化金融产品体系,强化技术赋能的风险预警与处置能力,着力提高金融服务的便利性和竞争力。

其二,借助人工智能准确、全面把握实体经济的金融需求,在实现金融服务实体经济的国家战略方面发挥重要功能。其中最为重要是,将人工智能技术与科技金融、绿色金融、普惠金融和养老金融深度融合,在深入探究这四篇金融大文章发展特性的基础上,使人工智能更好服务于实体经济发展。

其三,发挥人工智能全面提升金融基础设施服务能力的功能。一方面,将人工智能技术应用于跨境支付领域中,通过人脸识别、语音识别、生物识别、智能投资顾问等方式缓解其存在的成本高、速度慢、渠道有限、透明度不足等问题,提升金融机构的服务能力。另一方面,人工智能基于海量的结构和非结构数据,可替代数据的训练,也可借助提高关联分析能力,打造更加复杂的评估评价模型,并进行动态迭代,不断优化参数,进一步提升信用评估的科学性和准确性,弥补传统征信体系的不足之处。

## 五、以人工智能赋能金融强国建设的风险挑战

人工智能在金融领域的深度应用,在提升效率的同时,也因技术与金融目标之间的内在错配,引发了新的风险与挑战。人工智能的运行逻辑重在效率与标准化,而金融体系则需兼顾公平、稳定与服务实体等多重目标,二者之间的差异是各类新风险产生的根源。这些挑战集中体现在资源配置结构、服务本质属性、系统稳定性以及跨境治理等多个层面,对现有监管框架形成了全面考验。

### (一)资源配置的结构性可能失衡

人工智能融入金融资源配置环节后,决策依据将更倚重数据基础与模型判断。不同主体的数据条件和模型适配程度,会映射为风险评估与定价的差异,进而对融资获取和竞争地位产生影响,并在配置进程中持续累积。这种影响将沿主体与区域、产业投向、机构竞争的路径推进,最终体现为资源配置的结构性失衡。

### 1. 服务主体与区域失衡

人工智能的金融应用使数据在资源配置中的权重上升,可能加剧服务主体与区域层面的失衡。从服务主体看,人工智能在信用评估和风险定价中的应用,使金融决策更加依赖信息的可量化程度、历史记录完整性以及风险特征的可预测性。经营状况相对稳定、交易行为连续、数据基础较为完善的主体,更易进入基于算法的标准化评估流程,其融资渠道相对稳定。相对而言,中小微企业、创业团队等主体,由于数据积累不足或风险特征波动较大,其信用状况难以被模型充分刻画,融资可得性更易受到约束。这种差异主要由算法评估对数据的内生要求决定。当金融资源配置越来越依赖模型输出时,缺乏可持续数据积累的主体,其风险溢价易被系统性高估(Berg et al., 2020; Fuster et al., 2022)。

从服务区域看,在金融活动集中、交易记录连续、数据基础相对完善的地区,人工智能更易发挥其在风险控制和资源配置中的效率优势。相较之下,在县域和农村地区,由于市场主体规模偏小、经济活动分散、数据连续性和标准化程度较低,人工智能在信用评估和风险识别中的适用性受到一定制约,其对金融资源配置效率的改善作用相对有限,在客观上强化了金融资源向经济较为发达、数据条件较好区域集聚的趋势。

### 2. 投资过热与产能过剩

在人工智能主导投资筛选的情况下,资金配置更偏向回报路径清晰的领域,可能诱发局部投资过热。需要明确指出的是,上述倾向并非否定人工智能在科技创新融资方面的支持作用。在现实情况中,主流模型通常更依赖可量化的历史绩效以及已得到验证的商业模式来进行学习,故而在降低成熟项目的筛选成本方面更为擅长;而对于样本稀缺、技术路径尚不明确的前沿创新,其前瞻性识别能力依然相对有限。人工智能在投资决策和资金配置中的应用,显著提升了项目筛选和风险识别的效率,但其基于数据和模型的评估方式,也使资金配置更倾向于回报路径清晰、商业模式成熟、业绩表现可被量化的领域。在算法驱动的投资环境下,市场资金更容易围绕少数被反复验证的“优质赛道”集中,并带动相关领域投资规模快速扩张。中长期来看,若投资扩张节奏快于有效需求增长,相关行业可能出现阶段性产能过剩。在算法主导的投资环境下,这类错配具有一定的同步性,一旦市场预期发生反转,调整过程可能较为剧烈,并对上下游产业链产生连锁影响。

相比之下,研发周期较长、技术不确定性高、商业化路径尚不清晰的关键核心技术攻关和创新项目,由于其风险特征和回报前景难以通过既有模型充分刻画,往往难以获得与其战略价值相匹配的资金支持。这种资金配置取向与国家推动科技自立自强和培育新质生产力的政策目标之间存在鸿沟。资本缺乏对长期创新和前沿领域的耐心投入,可能使部分国家战略重点方向在融资支持上承压。

### 3. 市场竞争格局失衡

人工智能重塑了金融竞争方式,强化数据和技术优势的累积效应,进而形成

“马太效应”，可能推动市场集中度上升。人工智能在风险识别、定价测算和业务流程中的广泛应用，使金融机构在成本控制和业务扩展方面对数据积累和技术投入的依赖明显增强。头部金融机构和大型金融科技平台，依托广泛的客户基础和持续的数据积累，更容易开展模型迭代和算法优化，其规模效应和学习效应通过降低单位业务成本、提升风险管理效率逐步显现。相较之下，中小金融机构在数据获取、模型训练和系统维护方面面临较高的成本，其人工智能应用多集中于通用化工具，对其核心竞争力的提升作用相对有限。

数据积累、模型表现和业务扩张之间形成自我强化的正反馈机制，使先发头部机构在竞争中不断放大优势，而后进入者和中小机构的追赶成本持续上升。由此，金融业务呈现向少数金融机构集中的趋势，市场集中度随之上升。在竞争格局趋于集中的情况下，金融体系的多样性和风险分散功能受到抑制，整体韧性面临一定挑战。

## （二）服务实体经济功能可能弱化

人工智能深度融入金融决策流程，不仅变革了风险识别与业务执行模式，更在深层次上对金融决策的价值导向和运行逻辑产生了影响。相关风险并非单一环节失效所致，而是源于技术运行特性，并在金融场景中不断放大。其影响首先作用于数据运用与模型判定，继而逐步传导至服务对象筛选和资源配置模式，最终对实现金融服务实体经济目标形成阻碍。

### 1. 数据隐私泄露与模型幻觉

人工智能深度介入金融决策可能引发数据隐私泄露与模型幻觉等新型风险。随着人工智能在授信评估、风险控制和客户管理等环节中的广泛应用，多源数据被持续整合和交叉使用，个人信息和敏感数据在采集、处理、共享与调用过程中的暴露范围明显扩大（Acquisti et al., 2016）。在数据用途限定、最小必要原则和内部授权机制落实不充分的情况下，隐私信息被不当获取、滥用或外泄的风险随之上升，可能削弱金融服务赖以维系的信任基础。与此同时，生成式人工智能和复杂模型在金融场景中的应用，可能带来模型幻觉等新型风险。部分模型在缺乏充分约束或验证机制的情况下，可能基于不完整信息或统计相关性生成看似合理但缺乏事实依据的判断结果。当相关输出在信贷决策、投资建议和风险提示等环节中被直接用于业务判断时，可能引致决策偏差，影响客户权益保障，扰乱市场秩序。

从金融科技伦理视角看，数据隐私保护与模型可靠性不仅属于技术治理范畴，更直接关系金融机构的责任边界与信任基础。金融体系以信用为核心运行，对模型输出结果的过度依赖，一旦出现系统性偏差或判断失误，其影响不局限于单笔业务或个别客户，可能侵蚀公众对金融机构专业性的整体认知。同时，当金融机构在决策过程中弱化人工复核、责任追溯和内部问责安排，将判断权过度让渡于算法系统时，数据滥用、模型失效或错误决策所引致的风险，可能沿业务链条和市场预期迅速扩散，加剧风险累积并放大系统性影响。

## 2. 算法歧视与社会公平

金融服务向算法化、标准化转型,可能对金融公平性和普惠金融目标的实现形成挑战。算法决策基于特定的数据结构和规则假设,其运行逻辑难以充分反映不同社会群体在数字接入条件、金融参与方式和制度嵌入程度方面长期存在的差异。这类差异在算法框架下易被设定为直接的服务筛选标准,并持续影响金融服务的覆盖范围和配置方式。老年人、农村人口、低收入人群等群体,由于其金融行为特征与标准化模型设定的评估框架契合度较低,在授信准入、定价条件和服务连续性等方面面临更为严格的约束。

从普惠金融视角看,算法歧视风险不仅表现为个体层面的机会不均等,更可能削弱金融体系向弱势群体和欠发达地区提供基础金融服务的能力。当标准化算法逐步成为金融服务的主要入口时,部分原本依赖关系型金融、情境判断和人工干预获得服务的群体,可能被排除在主流金融服务之外(Bartlett et al., 2022)。这可能使普惠金融由“主动覆盖”转向“被动筛选”,削弱其在缩小区域差距和群体差距中的政策工具属性,致使其在缓解融资约束、促进社会包容性发展方面的功能受到制约。当缺乏有效的校正机制和伦理约束时,算法驱动的金融服务模式将加剧金融体系内部的分化程度。

## 3. 决策偏离金融本源

人工智能深度嵌入金融决策流程,可能导致金融资源配置的根本目标发生偏移,逐渐脱离服务实体经济的核心要求。当授信、投资与风险管理等关键环节日益依赖算法模型和数据指标时,金融决策将更强调可量化的表现和短期绩效信号,对产业运行逻辑、企业经营实质以及长期发展能力判断的权重相对下调。在该框架下,金融业务的重心容易由支持实体经济的价值创造,转向围绕模型表现和收益指标的局部优化,滋生脱实向虚风险。

如果算法逻辑在金融决策中持续强化,而相应的价值引导与伦理约束未能同步建立,金融活动的目标函数就可能趋于简化,甚至忽视其应承担的社会责任。实践中,金融机构更可能遵循技术效率优先和风险最小化的路径,而非主动统筹经济发展质量、产业结构优化与国家长期战略目标。长此以往,金融体系在服务国家发展战略、引导资源向关键领域流动的功能定位或将受到削弱。

### (三)对金融稳定带来新挑战

人工智能深度融入金融运行体系后,金融稳定所面临的挑战呈现出更为显著的隐蔽性、突发性以及处置困难等特征。具体而言,相关风险可能在算法和系统内部长期累积,却不易以明显的形式暴露出来;而当风险被触发时,又可能由于模型同质化的原因,导致风险迅速扩散。此外,模型可解释性的欠缺以及责任划分的不明确,会削弱合规审查和问责约束的效力,进而加大风险处置难度。

#### 1. 风险隐蔽性

金融体系对人工智能的深度依赖,可能使得传统金融监管工具效能降低,易构成系统性冲击。随着智能模型嵌入信贷、投资、定价和风险管理等关键环节,金融

风险的生成方式呈现出更强的内生性和隐蔽性。一方面,基于模型和算法的决策主要依赖统计关系和历史数据,风险积累不再直接体现为资产质量恶化或价格异常波动等可观测结果。以账面指标和事后表现为核心的监管手段,对潜在风险的识别灵敏度随之下降。另一方面,模型参数设定、数据口径选择以及算法迭代规则等关键因素,多嵌入于金融机构内部技术流程,其风险影响具有间接性和滞后性,难以及时反映在常规监管监测框架之中。相关风险在决策逻辑层面逐步累积,而非通过单一指标异常集中暴露,早期阶段缺乏清晰触发信号,增加了识别和应对难度,构成影响金融稳定的重要隐患。

进一步看,监管部门自身对人工智能工具的应用,也可能使风险更为隐蔽。随着监管科技与合规科技的发展,模型化工具与自动化分析框架被逐步用于风险监测、压力测试和行为分析。这在提升监管效率与覆盖面的同时,也可能使监管判断在方法论和数据结构上与被监管机构趋于同源。若监管模型与金融机构的风险模型在数据口径、特征选取或假设设定上高度相似,那么一些系统性偏差和模型盲区就难以被监管视角有效识别;反而可能在监管与市场间形成共同误判。此时,风险并非因信息缺失而被忽视,而是因模型共识而被低估,其隐蔽性与累积性将进一步加剧。

## 2. 新型风险传导路径

当金融机构普遍采用趋同的人工智能模型时,可能在市场波动中形成以算法决策为核心的新型风险传导路径,并放大金融体系的整体波动性。不同于传统依赖价格变化、杠杆调整和流动性收缩逐步传导的机制,基于模型和算法的决策规则更容易在机构之间形成高度一致的行为反应。在市场微观结构层面,算法交易在一定条件下可改善流动性并提升报价信息含量,但也使交易反应速度更快、对共同信号的敏感度上升(Hendershott et al., 2011)。当多家金融机构使用相近的数据来源、模型结构和参数设定时,其资产配置、风险对冲和仓位调整决策往往呈现同步特征。一旦市场环境发生变化或关键变量触及模型阈值,相关调整行为可能在短时间内集中触发,引发“羊群效应”,放大价格波动和市场冲击(Kirilenko et al., 2017)。

这类风险传导并非源于单一机构的风险暴露,而是通过算法决策的同质化在系统内部快速扩散。由于相关决策由嵌入式算法在技术系统内部自动触发并执行,风险调整不再依赖人工判断、内部协商或跨主体的信息博弈,其反应速度显著快于传统决策机制。机构之间原本通过价格变化、交易行为和市场预期逐步形成的反馈与修正过程被明显压缩,风险调整在短时间内集中完成。在此过程中,市场信号尚未充分显现,其他经济主体亦难以及时作出反应或提供流动性支持,依赖价格发现和行为博弈形成的缓冲机制难以发挥作用。由此,人工智能背景下的风险传导更易呈现出突发性、非线性和跨市场联动特征,对金融稳定构成新的挑战。

## 3. 合规与责任认定

随着人工智能被广泛引入授信审批、投资决策和风险管理等关键环节,现有合

规审查机制和责任认定框架在适配性上逐步显现不足。一方面,人工智能模型,尤其是深度学习模型,其决策过程依赖高维参数和复杂的非线性结构,内部逻辑难以通过传统审计和监管方式加以解释。尽管监管层面普遍要求模型具备一定的可解释性,但在实践中,如何对模型训练过程、特征选取、参数调整以及输出结果进行系统性审查,仍缺乏成熟方法和统一标准。合规审查因此更多停留在结果合规和事后评估层面,难以及时识别模型设定中的偏差、失效风险或潜在伦理问题,削弱了合规治理的前置约束功能。

另一方面,当人工智能支持的金融决策发生偏差并造成损失时,责任主体的界定面临现实困境。在数据来源、算法设计、系统部署和决策执行等环节中,金融机构、技术提供方和模型开发者之间的职责边界并不清晰。现有法律制度主要围绕人工决策构建,难以对算法驱动决策中各方的责任承担作出明确划分。一旦出现系统性偏差或模型失效,责任认定不清不仅增加事后问责难度,也削弱了事前合规激励和风险约束。总体来看,人工智能的引入并未降低金融活动对合规与责任的要求,反而对监管制度的技术理解能力、规则设计前瞻性以及制度衔接提出更高要求。若缺乏与算法特征相匹配的审查机制和责任安排,合规治理层面可能形成新的制度空白。

#### (四)国际金融治理的复杂性

人工智能在推动金融活动实现跨境化、技术化发展的进程中,亦显著提升了国际金融治理的复杂程度。技术依赖问题与算法驱动下的跨境风险传导相互交织,可能引发金融安全与监管协调面临新的不确定性。

##### 1. 技术依赖与战略自主风险

过度依赖境外人工智能技术或模型,不仅可能影响国内金融机构服务的稳定性和安全性,也制约我国金融业的全球竞争力与话语权。从运行层面看,若关键模型、算力平台或底层算法高度依赖境外技术体系,一旦相关技术服务发生调整、受限或中断,国内金融机构在短期内难以实现有效替代,金融业务连续性和风险管理有效性面临冲击。在高频交易、自动化风控和实时定价等对技术稳定性和响应速度要求较高的场景中,这类依赖更容易放大系统脆弱性。从发展层面看,长期依赖外部人工智能技术,可能削弱国内金融机构在模型构建、数据治理和算法优化方面的自主积累能力,不利于形成与本国金融体系结构、监管要求和发展目标相适应的技术体系。从战略层面看,人工智能与金融体系的深度融合,使技术依赖问题由效率和成本选择,上升为金融体系安全性与战略自主性问题。在全球科技竞争加剧、地缘政治不确定性上升的背景下,外部技术供给的不稳定性,容易转化为金融体系运行中的结构性约束,压缩宏观调控和金融监管在关键环节的自主空间。

##### 2. 跨境风险传导

人工智能驱动的跨境金融活动,使资本流动的组织方式和风险传导路径发生变化,显著提高了异常资本流动的隐蔽性与监管识别难度。基于算法的交易、定价

和风险管理系统,能够在多市场、多币种和多账户之间实现高度联动,资金调整不再依赖单一市场的交易信号或人工决策,而是由模型阈值和参数变化自动触发。在高频交易、智能配置和算法套利等场景中,跨境资金可在不同司法管辖区同步移动,风险暴露分散于多个市场之中,难以在单一监管视角下被及时识别和有效处置。在此基础上,各国在人工智能金融应用方面的监管目标、合规要求和技术标准差异,进一步放大了跨境风险传导的复杂性。当前跨境监管体系呈现明显碎片化特征,部分金融机构和技术平台通过业务拆分、数据跨境调用、模型外包或云算力配置等方式,在不同监管规则之间进行配置选择,增加了监管套利空间。一旦相关模型在某一司法辖区受到冲击,其影响可能通过技术系统和业务链条迅速传导至其他市场,加剧跨境风险的联动性。

与此同时,部分发达经济体已将人工智能、数据与金融安全纳入“长臂管辖”的重要范围,使跨境风险传导呈现出明显的法律与制度外溢特征。欧盟《通用数据保护条例》通过域外适用规则,将对欧盟居民数据的处理活动纳入监管,即便相关金融机构或技术服务提供者未在欧盟设立实体,其跨境获客、风险评估和客户画像模型仍可能受到实质性约束。美国《云法案》则以“占有、保管或控制”为标准,要求服务提供者在特定情形下交付境外存储的数据,增加了跨境金融业务在数据合规、隐私保护和法律冲突方面的不确定性。同时,美国对高端芯片、算力服务和人工智能相关技术的出口管制,将技术供给纳入国家安全框架,使跨境金融系统对外部技术依赖的风险,可能通过技术链条直接传导至金融业务运行层面。总体来看,人工智能背景下的跨境金融风险,不仅表现为资本流动速度和规模的变化,更体现为资金联动、监管规则差异以及法律域外适用性叠加形成的复合型传导机制。这对以国别监管为基础的传统金融治理框架提出了更高要求。

## 六、推动人工智能赋能金融强国建设的政策建议

习近平总书记在二十届中央政治局第二十次集体学习时指出:“面对新一代人工智能技术快速演进的新形势,要充分发挥新型举国体制优势,坚持自立自强,突出应用导向,推动我国人工智能朝着有益、安全、公平方向健康有序发展。”面对人工智能深刻重构金融体系功能结构、运行逻辑与风险特征的趋势,构建一个既能促进创新、又能确保安全的人工智能金融发展体系,已经是金融强国建设的重要任务。当前我国处于金融数字化智能化加速发展的历史阶段,既具备丰富的金融场景、庞大的数据基础和活跃的技术创新生态,也面临金融风险跨市场、跨机构、跨技术叠加的新特征。推动人工智能在金融领域的健康有序发展,需要坚持“发展为先、监管并重、开放协同”的基本原则。具体而言,可以从以下几个方面重点发力:

第一,在人工智能基础理论、方法体系与关键工具方面实现突破,夯实金融

业智能化发展的科技底座。当前人工智能在金融场景中的广泛应用,仍高度依赖于深度学习、自然语言处理、强化学习、图神经网络等基础模型,但金融领域的复杂系统特征(如路径依赖、风险放大机制、非线性冲击传导等)对模型的稳定性与可解释性提出更高要求。因此,要进一步加强国家层面的人工智能基础研究布局,形成解释能力更强、抗噪声能力更强的“金融领域专用模型”。同时,推动金融机构在高频交易、风险管理、资产组合管理、监管科技等关键领域研发场景化模型与工具,提高人工智能的行业渗透深度。此外,还需加强“可信AI”“可解释AI”等方向的研究,确保模型行为在金融场景中具有制度可控性与监管可追溯性。

第二,统筹推进算力基础设施体系建设,深化金融数据资源的开发利用与安全共享,为人工智能应用提供可持续资源支撑。人工智能算法的运行需要持续稳定、低延迟、高安全性的算力基础,而金融场景对实时性、稳定性和极端情况下的算力保障提出更高要求。因此,要加快构建覆盖全国、供需匹配、低时延可控的金融算力体系,推动建设金融行业专用云平台、智能算力调度平台,以及应对极端情况的算力备份机制。同时,健全金融数据资源体系,推动政府部门、金融机构、行业协会和大型平台企业共同构建标准化、规范化的数据资源池,形成安全可信的数据供给能力。在确保数据安全与隐私保护的前提下,探索多方安全计算、联邦学习、隐私增强技术、数据沙箱等新路径,破解跨机构、跨行业数据共享壁垒,提升数据要素在金融智能化过程中的配置效率。建立严格的数据质量管理体系、数据标注标准体系与数据治理机制,为人工智能模型的训练、验证与监测提供高质量的数据基础。

第三,制定健全的法律法规、应用规范及伦理准则,构建人工智能在金融领域发展的制度约束框架。人工智能在金融场景的应用正在突破传统制度边界,算法责任模糊、模型黑箱、数据滥用、算法歧视、消费者权益风险等问题日益凸显。因此,需要加快制定人工智能在金融领域的专项法律法规,明确算法主体责任、数据处理规范、模型管理要求、异常风险报告机制等制度安排,将“算法可解释、数据可控、行为可追溯”作为金融人工智能应用的基本规范。同时,将人工智能风险纳入金融审慎监管框架,建立“算法风险评估—模型分类分级管理—人工兜底机制”等一整套制度体系。要推动制定关于人工智能伦理的金融行业规则,加强对算法偏见、歧视性结果、不当使用AI刺激消费者风险偏好等行为的规范,引导金融机构开展技术伦理审查。对标国际监管标准,推动我国在智能金融治理领域形成具有国际影响力的制度体系,打造规范、透明、健全的智能金融制度框架。

第四,构建人工智能技术监测、风险预警与应急响应体系,确保金融人工智能系统的安全性、可靠性与可控性。在人工智能深度嵌入支付系统、清算系统、信贷决策体系、资产交易体系等金融领域后,其潜在失灵可能引发系统性风险。因此,要建立人工智能模型的全生命周期管理制度,覆盖模型设计、训练、验证、上线、监

测、复核与退役的完整链条,并强化模型日志记录、模型漂移监控、算法偏差检测等机制安排。构建人工智能算法的压力测试体系,通过模拟极端市场波动、数据突变、对抗攻击等多维冲击,检验模型的稳健性。要建立国家级人工智能金融风险监测平台,开展跨机构、跨系统模型风险扫描,监测可能的“算法共振”“模型趋同”“系统性依赖”等风险形态。同时,构建应急响应机制,包括模型故障快速隔离、人工接管机制、跨机构协同处理机制等,确保关键金融基础设施在出现算法异常时仍能够保持连续稳定运行。

第五,全面加强人工智能国际合作,推动构建具有广泛共识的全球治理框架,同时帮助全球南方国家提升技术能力,弥合“智能鸿沟”。随着人工智能在跨境支付、全球资本流动、跨国银行监管、反洗钱与反恐融资、供应链金融等领域的广泛应用,国际协调成为金融智能化治理的关键。因此,要积极参与国际组织关于人工智能与金融监管的讨论,推动在数据治理、模型评估、算法安全、监管合作等方面形成统一规范。加强与欧美、亚太等地区在技术标准、风控体系、监管科技方面的对接,参与人工智能金融标准体系的制定,提高中国在全球智能金融治理中的话语权和影响力。同时,面向全球南方国家提供智能金融基础设施建设支持,包括数字支付系统、反洗钱技术、监管科技工具等,帮助其提升数字金融与智能金融能力,为弥合全球智能鸿沟作出中国贡献。

## 参考文献

- 曾燕、查佳婧、杨海生、杨存奕(2024):《中国数字普惠金融的发展模式探究:经济与民生的视角》,《经济研究》,第8期。
- 陈雨露(2018):《推动绿色金融标准体系建设》,《中国金融》,第20期。
- 范小云、王业东、王道平(2022):《基于新闻大数据与机器学习的中国银行业系统性风险研究》,《世界经济》,第4期。
- 黄益平、黄卓(2018):《中国的数字金融发展:现在与未来》,《经济学(季刊)》,第2期。
- 姜富伟、刘雨旻、孟令超(2024):《大语言模型、文本情绪与金融市场》,《管理世界》,第8期。
- 李斌、邵新月、李玥阳(2019):《机器学习驱动的基本面量化投资研究》,《中国工业经济》,第8期。
- 李金金、曾圣钧(2025):《科技赋能养老金融高质量发展的路径选择》,《国际金融》,第6期。
- 廖高可、李庭辉(2023):《人工智能在金融领域的应用研究进展》,《经济学动态》,第3期。
- 刘少波、张友泽、梁晋恒(2021):《金融科技与金融创新研究进展》,《经济学动态》,第3期。
- 马甜、姜富伟、唐国豪(2022):《深度学习与中国股票市场因子投资——基于生成式对抗网络方法》,《经济学(季刊)》,第1期。
- 苏治、卢曼、李德轩(2017):《深度学习的金融实证应用:动态、贡献与展望》,《金融研究》,第5期。
- 孙玉环、张汀昱、王雪妮、李丹阳(2021):《中国数字普惠金融发展的现状、问题及前景》,《数量经济技术经济研究》,第2期。
- 张明、陈骁(2024):《中国绿色信贷发展状况、基本特征与政策建议》,《金融理论探索》,第3期。
- 张明喜、魏世杰、朱欣乐(2018):《科技金融:从概念到理论体系构建》,《中国软科学》,第4期。
- 张晓燕、吴辉航(2024):《大语言模型在金融领域的应用》,《清华金融评论》,第5期。
- Acquisti, A., C. Taylor and L. Wagman(2016): “The Economics of Privacy”, *Journal of Economic Literature*, 54, 442-492.
- Aldasoro, I., L. Gambacorta, A. Korinek, V. Shreeti and M. Stein(2024): “Intelligent Financial System: How AI

is Transforming Finance”, BIS Working Papers, No.1194.

Almahdi, S. and S. Yang (2017): “An Adaptive Portfolio Trading System: A Risk–return Portfolio Optimization Using Recurrent Reinforcement Learning with Expected Maximum Drawdown”, *Expert Systems with Applications*, 87, 267–279.

Bartlett, R., A. Morse, R. Stanton and N. Wallace (2022): “Consumer–Lending Discrimination in the FinTech Era”, *Journal of Financial Economics*, 143, 30–56.

Berg, T., V. Burg, A. Gombović and M. Puri (2020): “On the Rise of Fintechs: Credit Scoring Using Digital Footprints”, *Review of Financial Studies*, 33, 2845–2897.

Chatzis, S., V. Siakoulis, A. Petropoulos, E. Stavroulakis and N. Vlachogiannakis (2018): “Forecasting Stock Market Crisis Events Using Deep and Statistical Machine Learning Techniques”, *Expert Systems with Applications*, 112, 353–371.

Daniélsson, J., R. Macrae and A. Uthemann (2022): “Artificial Intelligence and Systemic Risk”, *Journal of Banking and Finance*, 140, 106290.

Frankel R, J. Jennings and J. Lee (2022): “Disclosure Sentiment: Machine Learning vs. Dictionary Methods”, *Management Science*, 68, 5514–5532.

Fuster, A., P. Goldsmith–Pinkham, T. Ramadorai and A. Walther (2022): “Predictably Unequal? The Effects of Machine Learning on Credit Markets”, *Journal of Finance*, 77, 5–47.

Gambacorta, L., Y. Huang, H. Qiu and J. Wang (2024): “How Do Machine Learning and Non–traditional Data Affect Credit Scoring? New Evidence from A Chinese Fintech Firm”, *Journal of Financial Stability*, 73, 101284.

Goldsmith, R. (1969): *Financial Structure and Development*, Yale University Press.

Goldstein, I., W. Jiang and G. Karolyi (2019): “To FinTech and Beyond”, *Review of Financial Studies*, 32, 1647–1661.

Gu, S., B. Kelly and D. Xiu (2020): “Empirical Asset Pricing via Machine Learning”, *Review of Financial Studies*, 33, 2223–2273.

Gurley, J. and E. Shaw (1956): “Financial Intermediaries and the Saving–investment Process”, *Journal of Finance*, 11, 257–276.

Hendershott, T., C. Jones and M. Menkveld (2011): “Does Algorithmic Trading Improve Liquidity”?, *Journal of Finance*, 66, 1–33.

Huang, J., J. Chai and S. Cho (2020): “Deep Learning in Finance and Banking: A Literature Review and Classification”, *Frontiers of Business Research in China*, 14, 121–144.

Kirilenko, A., A. Kyle, M. Samadi and T. Tuzun (2017): “The Flash Crash: High–Frequency Trading in an Electronic Market”, *Journal of Finance*, 72, 967–998.

Korinek, A. (2023): “Generative AI for Economic Research: Use Cases and Implications for Economists”, *Journal of Economic Literature*, 61, 1281–1317.

Levine, R. (1993): “Financial Structures and Economic Development”, *Economic Analysis Review*, 8, 113–129.

Li, J., Y. Chen, Y. Shen, J. Wang and Z. Huang (2019): “Measuring China’s Stock Market Sentiment”, SSRN Working Paper, No.337–7684.

Liberti, J. and M. Petersen (2019): “Information: Hard and Soft”, *Review of Corporate Finance Studies*, 8, 1–41.

Merton, R. and Z. Bodie (1995): “A Conceptual Framework for Analyzing the Financial System”, in *The Global Financial System: A Functional Perspective*, edited by Crane, D., K. Froot, S. Mason, André Perold, R. Merton, Z. Bodie, E. Sirri and P. Tufano, 3–31, Harvard Business School Press.

Merton, R. and Z. Bodie (2004): “The Design of Financial Systems: Towards a Synthesis of Function and Structure”, NBER Working Paper, No.10620.

# Artificial Intelligence and the Construction of a “Financial Powerhouse”: Theoretical Logic, Implementation Pathways and Risk Challenges

WANG Yong<sup>[a]</sup> FAN Yunpeng<sup>[a]</sup> WANG Yao<sup>[a]</sup>

a. Institute of Finance and Banking, Chinese Academy of Social Sciences, Beijing 100710, China

**Summary:** Artificial intelligence is progressively emerging as a pivotal technological force driving the evolution of the financial system. Distinct from earlier technological phases characterized by mere digitization and automation, AI exerts a profound influence on financial operations and development paradigms by endogenously embedding itself into business processes and decision-making loops. This impact extends beyond marginal gains in operational efficiency to induce systemic shifts in resource allocation, risk identification, and price formation mechanisms. Against the backdrop of accelerating the construction of a “financial powerhouse” and promoting high-quality development, a holistic analysis of how AI reshapes the logic of financial development is imperative.

While the widespread application of AI significantly alters the depth of information acquisition, the precision of risk identification, and the latency of decision responses, thereby transforming the modes of delivery and their efficacy. However, it does not alter the essential structural functions of finance. Currently, the application of artificial intelligence, particularly generative AI, within China’s financial sector remains in a nascent stage. Tailoring technological applications to the distinct developmental characteristics of the “Five Major Articles”, namely Technology Finance, Green Finance, Inclusive Finance, Pension Finance and Digital Finance, holds promise for value creation in key areas. However, the reshaping of financial development by AI exhibits a pronounced duality. On one hand, enhanced technological capabilities offer new possibilities for improving service quality and expanding the frontier of financial inclusion. On the other hand, issues such as data concentration, algorithmic homogeneity, and model convergence may exacerbate structural imbalances in resource allocation, posing novel challenges to market fairness and systemic financial stability.

To effectively leverage AI in the construction of a “financial powerhouse”, this paper proposes several strategic recommendations. First, it is essential to achieve breakthroughs in foundational AI theory, methodological systems, and key tools to solidify the technological base for the intelligent evolution of the financial industry. Second, policymakers should orchestrate the construction of computing power infrastructure and deepen the exploitation, utilization, and secure sharing of financial data resources to provide sustainable support for AI applications. Third, there is a need to formulate robust laws, regulations, operational standards, and ethical guidelines to construct a binding institutional framework for AI development in the financial domain. Fourth, comprehensive systems for AI monitoring, risk early-warning, and emergency response must be established to ensure the safety, reliability, and controllability of financial AI systems. Finally, international cooperation on AI should be strengthened to promote a global governance framework based on broad consensus, while simultaneously assisting the Global South in enhancing technological capabilities to bridge the “intelligence divide”.

**Keywords:** Artificial Intelligence; Financial Development; Financial Supervision; Financial Powerhouse

**JEL Classification:** E20; E60; G00

(责任编辑:周莉萍)(编校:范可鑫)