

AI算法黑箱与直播电商供应链金融风险传导机制研究

张曦鹏

贵州大学经济学院, 贵州 贵阳

收稿日期: 2025年3月28日; 录用日期: 2025年4月18日; 发布日期: 2025年5月20日

摘要

本研究聚焦AI算法对直播电商供应链金融风险的传导机制,揭示了算法不可解释性引发的三种核心问题:风险控制失效、信息不对称加剧以及系统性风险累积。通过机制分析后发现,算法黑箱通过模糊决策逻辑、加剧数据垄断、放大目标冲突等途径,促使风险从局部传导至全局,形成网络化的系统性金融风险威胁。针对上述问题,研究提出系统性治理框架,包括构建可解释性算法、重构数据治理结构、实施跨平台协同监管,以增强透明度、优化验证机制并阻断风险传播。然而,研究依旧存在局限:治理策略的实证效果尚未得到验证,算法异质性对治理普适性的影响并未充分探讨,风险动态性与非线性特征的模拟分析不足。未来需结合案例模拟与跨学科方法深入研究,为直播电商平台供应链金融的智能化治理提供更为全面的理论与实践支持。

关键词

直播电商, 供应链金融, AI算法, 传导机制

Study on the Risk Transmission Mechanism of AI Algorithmic Black Box in Live-Streaming E-Commerce Supply Chain Finance

Xipeng Zhang

School of Economics, Guizhou University, Guiyang Guizhou

Received: Mar. 28th, 2025; accepted: Apr. 18th, 2025; published: May 20th, 2025

Abstract

This study focuses on the risk transmission mechanism of AI algorithms in live-streaming e-commerce supply chain finance, revealing three core issues triggered by algorithmic unexplainability: risk control failure, exacerbated information asymmetry, and systemic risk accumulation. Through mechanism analysis, it is found that algorithmic black boxes propagate risks from localized to systemic levels through pathways such as obfuscation of decision logic, intensified data monopolization, and amplification of goal conflicts, forming networked systemic financial risk threats. To address these issues, the research proposes a systemic governance framework, including building explainable algorithms, restructuring data governance frameworks, and implementing cross-platform collaborative supervision, to enhance transparency, optimize verification mechanisms, and block risk transmission. However, limitations persist: empirical validation of governance strategies remains untested, the impact of algorithmic heterogeneity on governance universality is not fully explored, and simulation analysis of risk dynamics and nonlinearity is insufficient. Future research should integrate case simulations and interdisciplinary methods to deepen investigations, providing more comprehensive theoretical and practical support for intelligent governance in live-streaming e-commerce supply chain finance.

Keywords

Live-Streaming E-Commerce, Supply Chain Finance, AI Algorithms, Transmission Mechanism

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

1.1. 研究背景

近年来,直播电商作为数字经济的新兴业态,通过“人货场”重构了传统消费场景,成为推动供应链金融创新的重要力量。数据显示,2023年中国直播电商市场规模已接近千亿级别,其即时性、互动性与高转化率的特征,催生了预付款、即时支付等新型资金流动模式,但也加剧了供应链金融的复杂性。与此同时,人工智能技术(AI)的深度应用进一步改变了直播电商的生态格局。例如,AI算法被用于智能选品、风险评估、自动化交易等环节,显著提升了运营效率,但也引入了“算法黑箱”问题——即算法决策过程的不透明性导致风险传导路径难以追溯。

在供应链金融领域,传统风险传导机制主要依赖资金流、物流与信息流的透明度。然而,AI算法黑箱通过模糊决策逻辑、加剧数据隐私风险、放大算法歧视等问题,使得风险传导从线性、可预测的模式转向复杂、不可知的动态网络。部分直播平台因算法推荐机制导致流量分配不均,间接引发供应商资金链断裂;金融机构依赖黑箱化的风险评估模型,可能误判中小企业的信用状况,加剧系统性风险。AI算法如何影响直播电商平台金融风险,金融风险又通过何种渠道加以传导,又应该通过何种手段减轻“黑箱”机制对直播电商平台的风险冲击,成为学界亟待研究和解决的问题。

1.2. 研究意义

1.2.1. 理论意义

现有供应链金融风险研究多聚焦于传统业态,对AI算法介入后的风险传导机制缺乏系统性分析。本

研究通过揭示算法黑箱对风险传导路径的重构效应，可弥补传统理论在技术伦理与复杂系统动力学层面的不足，推动供应链金融理论向智能化时代演进。

1.2.2. 实践意义

在实践层面，直播电商平台和金融机构面临着日益复杂的市场环境和风险挑战，AI 算法黑箱的存在使得风险识别和防控难度进一步加大。通过本研究，可以为直播电商平台和金融机构提供应对 AI 算法黑箱风险的理论支持和实践指导，帮助其优化风险防控机制，提升供应链金融的稳定性和安全性，促进直播电商行业的健康发展。

2. 文献综述

传统供应链金融的核心问题聚焦于风险防控方面，信息不对称导致的信用风险是传统供应链金融的主要挑战，常表现为供应链节点企业之间的信任缺失，而通过区块链技术，实现供应链资金的透明化追踪和自动化合约交易，能有效降低供应链金融面临的操作风险(林美吟, 2024) [1]，且在物联网时代下，如果供应链断裂引发流动性风险，物联网监控设备也能实时监测物流状态，结合区块链技术实货物信息的不可篡改记录，从而提升监管的透明性与安全性(葛文博, 2024) [2]。因此，供应链全流程风险评估研究意义重大。在未来，基于大数据采集整合、完善互联网环境下风险评估指标体系，运用深度学习构建高性能且可解释的模型，可以推动供应链金融的稳健运行(祝由等, 2023) [3]。

直播电商供应链金融在传统供应链金融的基础上，融合了直播运营的即时性与互动性，其风险特征和协同机制呈现出新特点。由于直播电商风险具有动态性和网络性，涉及物流储运、信息技术、直播运营等多环节，部分学者利用 DEMATEL 方法构建风险指数模型，识别出第三方物流外包、信息追溯、配送环节等关键风险因素(张敏, 2024) [4]。此外，供应链金融可缓解中小企业融资困境，部分学者从信用质押物、平台性质和资金来源分析电商供应链金融模式，以多案例研究剖析 B2B、B2C、C2C 电商平台典型案例，发现电子信用模式受青睐、电子仓单模式应用少，不同模式应相互借鉴(谢文静等, 2021) [5]。这都向与之配套的供应链提出了更高的要求，尤其是在建设信息共享平台、优化配送路径、引入第三方物流合作等方面，并在跨主体协同，加强与物流企业的协作方面持续发力(金灏, 2025) [6]。

现有文献多从大数据、云计算赋能下的电商转型角度进行研究，缺乏对直播电商在 AI 算法影响下风险传导的系统性研究。本研究拟对 AI 算法黑箱引发的潜在供应链金融风险问题进行机制梳理，揭示 AI 技术如何影响供应链金融风险的传导路径，并提出相关的建议以供学术界参考。

3. AI 算法黑箱引发的供应链金融风险问题

3.1. 风险控制失效

AI 算法黑箱的不可解释性使得金融机构难以理解风险评估的具体逻辑。不同电商平台之间的 AI 算法模型不同，对于不同金融变量赋予的权重不尽相同，在风险评估的环节给金融机构带来困难。且黑箱模型的输出结果也缺乏透明性，使得金融机构既无法验证金融变量的具体权重，也无法识别突发情况对于直播电商供应链的真实影响，造成金融机构与电商平台的“信息鸿沟”，最终引发授信偏差，给金融机构不良贷款率造成影响[7]。

3.2. 信息不对称加剧

直播电商平台通过 AI 算法黑箱机制垄断数据入口，往往以“保护用户隐私”为由，仅仅向金融机构输出结论性指标(如商家健康度评分、用户信用质量评分等)，但遮蔽底层的计算规则。金融机构在获得数据后无法验证交叉数据的逻辑，可能误判商家的具体经营状况，对其财务指标和供应链指数进行错误估

计，引发融资歧视或过度授信问题[8]。

3.3. 系统性风险累积

AI 算法的隐蔽性、不可透明性也会带来潜在的风险。不同直播电商平台的 AI 算法模型往往只兼容自身结构，关注平台自身而忽略了与之相关联的其他供应链企业。通俗来说，就是 AI 算法可能更关注平台的某个主播的自身情况，而忽略了其背后供应商和其相关产业链的协同效应。一旦这些关键环节出现错误，这些主播的头部风险很可能由于“聚合效应”而不断放大，对其背后的代工厂、物流商等关联方造成违约，引发行业性或区域性的金融风险[9]。

4. AI 算法引发的问题机制分析

4.1. 风险控制失效的机制

4.1.1. 模型逻辑不可验证

算法黑箱通过隐藏式特征提取(如直播间互动频率、用户停留时长等非结构化数据)作为前因变量，生成风险评估结果，本质上是分析各个前因变量对结果变量(如风险评估结果、平台信用能力等)的线性关系，但是，对于不同前因变量之间的内在影响往往被忽略，而恰恰是这种“交互作用”对供应链金融的影响最为显著，由于 AI 算法的隐蔽性与不可知性，这种关联的“交互作用”无法被金融机构所解析，从而形成了“信息鸿沟”。且不同平台算法基于差异化目标(如 GMV 最大化、用户留存率最大化)等设计变量权重，导致跨平台风险评估的标准冲突，金融机构的授信逻辑无法确定，难以针对不同的平台建立统一的评估体系。

4.1.2. 动态环境误判

黑箱模型对突发事件的响应机制不可观测，尤其是“6.18”“双十一”等促销活动会对 AI 算法的估计造成一定冲击。这种短期订单量或者退货量上升的信息如果忽略了电商平台经营的时间效应，很有可能被 AI 算法误判为长期经营能力的提升，或是信用能力的下降，在进行模型估计和预测时，这种短期的波动很有可能因为黑箱机制而演变成长期的趋势走向，进而对供应链企业的决策与经营带来不利的影响。

4.1.3. 授信偏差放大

现有的银行等金融机构往往依据资产负债率、流动比率等偿债指标和存货周转率等运营指标来评价企业的经营效率，这些“硬信息”虽然在很大程度上决定了企业质量，但是也无形中增加了对电商平台实际经营效果的误判。直播电商以其虚拟化、轻资产化的特点，很可能不满足银行贷款审批的“硬条件”，而电商平台的 AI 算法如果过度加权“直播间观众地域分布”“直播间观众停留时长”等，而忽略“直播间销售数额”和“直播带货行业”这些指标，很有可能造成银行授信和直播电商实际经营状况不符，导致金融资源错配的严重问题。

因此，这一传导路径的结论可以概括为：

算法的不可解释性→模型逻辑不可验证→直播电商供应链风险的误判→金融机构授信偏差→资源严重错配。

4.2. 信息不对称的传导机制

4.2.1. 数据权力集中化

作为反映平台流量和经营的重要指标，大部分直播电商平台选择将用户行为数据(如直播间点击流、回购率等)、供应链实时数据(如存货周转率)进行保密，仅仅输出加工后指标，并没有向公众公开这些指

标的具体定义和计算方式。这就使得金融机构必须依靠自身所掌握的数据来衡量企业输出的结果指标，关键环节的因果逻辑无法及时得到验证。

4.2.2. 动态环境误判问题

随着大数据日益成为生产环节的重要资源，金融机构也逐渐将企业大数据作为对电商等新兴产业授信的重要依据。但目前来说，AI 模型算法的监管依旧不成熟，金融机构往往只能依赖平台提供的单一信号(如评分等)进行金融决策，但数据遮蔽导致平台传递的信号存疑。倘若平台想进行融资，要求所属线上门店通过线上刷单、批量评分等途径获得高分，AI 算法模型便极有可能出现“动态环境误判”的情况，真实的经营数据被算法忽略，违约风险也就此埋下。

因此，这一传导路径的结论可以概括为：

数据垄断→规则遮蔽→验证失效→融资约束/过度授信→市场效率损失。

4.3. 系统性风险累积的传导机制

4.3.1. 局部 - 全局目标冲突

作为企业，平台经营的目标也遵循“利润最大化”或“股东财富最大化”，反映在 AI 算法的编制上则体现为特定财务指标的高权重和对供应链企业评分的高要求。对于受季节、上下游企业影响较大的企业而言，这样的算法实际上是将供应商的资金、原材料压力转移给二级的供应商和与之合作的金融机构，变相造成了金融风险的跨行业、跨产业链传播，造成局部目标(可能是维持存货周转率稳定和客户稳定)与总体目标(利润最大化)之间的冲突。

4.3.2. 网络化风险扩散

随着网络主播“个人 IP”效应的商业化价值日益凸显，越来越多的厂家选择直接与网络主播的强“IP”绑定，但网络主播的个人 IP 曝光度在一定程度上也是由 AI 智能算法进行计算得出的。所以，主播的一些个人行为很有可能引发 AI 算法的误判(例如直播带货时的不当言论)，造成其流量推送减少、曝光度降低等问题，这种行为也会影响与其强绑定的厂家，倘若由于该不确定性事件导致销量降低，触发应收帐款违约，风险也会经供应链网络传播扩散至金融机构，形成类似网状的风险传播。

4.3.3. 尾部风险漠视

在激烈的市场竞争下，电商平台在选择供应商时，往往更看重在市场表现良好的头部供应商(比如行业前 20%的企业)，算法为其赋予的权重相对较高。但这种算法忽略了产业链之间的关联性与区域间的协同风险。一旦区域性、行业性的风险集中爆发，即便是处于头部的供应商也会受到冲击，最后可能引发区域间的供应链金融崩盘。

因此，这一传导路径可以概括为：

算法孤岛→风险隔离失效→行业风险共振→跨链传染→系统性崩溃。

5. AI 算法黑箱的系统性治理路径

5.1. 风险控制失效的应对策略

5.1.1. 可解释性算法治理框架的构建

模型逻辑的“不可验证”问题凸显了算法透明度不足的缺陷，需要构建“技术透明性”与“监管穿透性”并重的双重治理机制。在技术层面上，可推广可解释人工智能技术(Explainable AI)，逐步开放平台风险评估中的可视化解释模块，通过知识图谱技术实现不同前因变量权重的动态解析。在监管层面上，参照行业标准要求构建算法透明度评估指标体系，明确要求平台披露模型架构、变量交互关系和前因变量

权重，以确保算法决策过程中的可追溯性[10]。

5.1.2. 动态风险的监测与压力测试的优化

动态模型的误判问题源于对突发事件响应机制的不可观测性，亟待构建基于场景化压力测试的实时预警体系。金融机构可引入智能数据采集技术，通过自动化系统实时获取电商平台经营数据流，并结合时间序列方法建立促销周期(或者困难周期)预测模型。此外，进行压力测试时模拟的极端场景下的风险场景，也可检验 AI 算法模型在非稳定状态下的稳健性[11]。

5.1.3. 动态风险的监测与压力测试的优化

传统财务指标与直播电商的轻资产化的特征加剧了授信偏差，为此，需要建立一套由“硬性财务指标 + 软性运营数据”的复合评估框架，在保留核心财务指标的基础上，将“直播间互动频率”、“用户停留时长”等非结构化数据纳入评估维度。同时，行业协会应牵头制定跨行业授信标准指南，统一核心指标的定义和权重分配原则，缓解因算法目标差异导致的评估标准冲突[12]。

5.2. 信息不对称的治理途径

5.2.1. 数据治理结构的重构与优化

数据权力集中化现象加剧了金融平台与平台间的信息壁垒，亟需建立多方协同的数据治理模式。区块链技术的分布式数据存证系统能够实现供应链数据的可追溯性与安全性。此外，相关部门和行业协会需制定专项数据治理规范，明确平台数据的采集边界与共享的合规边界，要求其公开健康度评分等衍生指标的计算逻辑，确保数据使用的透明度[13]。

5.2.2. 第三方审计与激励相融机制的创新

规则遮蔽问题暴露了现有数据验证机制的局限性，第三方独立审计的强化监督尤为必要。具体而言，可建立算法全流程的审计规范，要求平台定期提交包含特征工程、模型训练等关键环节的技术文档。同时，设计激励相融政策，对主动开放数据接口的平台予以监管评级加分等正向激励，促进数据共享环节的可持续性发展[14]。

5.2.3. 隐私计算技术的场景化应用

数据验证失败问题需要通过技术创新来实现，“联邦学习”等隐私计算技术则提供了解决方案。通过构建多方安全的计算平台，可在不共享原始数据的前提下实现金融机构模型与平台风险防控模型的联合训练，不仅能帮助直播电商平台提升风险评估的准确性，又能够有效保护用户隐私和商业机密，形成“数据价值挖掘 - 客户隐私保护”的高效平衡机制[15]。

5.3. 系统性风险的防控体系构建

5.3.1. 跨平台协同监管体制的完善

算法孤岛现象往往极易导致风险隔离失败，需要通过监管沙盒与跨链协作强化机制强化系统性金融风险防控。尤其是针对核心算法服务商实施准入管理，并要求其通过算法安全认证。同时推动区块链技术在供应链网络中的深度整合，构建跨平台数据共享系统，通过智能合约来实现风险信号的实时传递与验证[16]。

5.3.2. 网络韧性增强策略的实施

供应链金融风险的网络化扩散要求进一步提升供应链全链抗风险冲击能力，可通过社会网络分析来识别供应链网络中的核心节点，建立动态的流动性储备系统。除了可通过压力测试检验系统的稳健性外，

还可根据风险的网状传播特点, 开发风险模拟传播系统, 基于多智能体建模预测风险的传导路径, 实现风险阻断的精准施策[17]。

5.3.3. 宏观审慎监管框架的构建

供应链的“尾部风险漠视”问题反映了目前针对 AI 算法监管体系的不足。针对这一问题, 可建立“微观-宏观”的双重监管框架。在微观层面上, 应当逐步要求平台将“区域协同风险”纳入供应商评估模型。在宏观层面上, 行业协会应协同监管机构, 建立跨行业风险监测平台, 运用机器学习技术识别突发情况下的跨行业风险共振信号, 建立系统性风险的预警指标体系[18]。

6. 结论

本研究聚焦 AI 算法黑箱对直播电商供应链金融风险的传导机制, 揭示了算法不可解释性引发的核心问题及其内在逻辑。研究表明: 第一, 算法黑箱通过模糊决策逻辑, 加剧了风险评估模型的可验证性缺失, 导致金融机构授信偏差与资源错配; 第二, 数据权力集中化与规则遮蔽显著放大了信息不对称问题, 平台垄断的数据入口与衍生指标的不可溯源性, 使得金融机构难以验证信用评分的真实性, 引发融资约束或过度授信; 第三, 算法孤岛与局部目标优先性促使风险从单一节点向供应链网络扩散, 通过目标冲突、网络化传播与尾部风险漠视, 形成系统性金融威胁。上述机制表明, AI 算法的隐蔽性不仅重构了风险传导路径, 还放大了风险的动态性与非线性特征。

研究不足主要体现在以下方面: 其一, 治理框架的可行性与实际效果缺乏实证检验, 尤其是联邦学习、区块链协同监管等技术的落地场景尚未通过案例模拟验证; 其二, 算法异质性的治理普适性的影响未被充分探讨, 可能导致结论的适用范围受限; 其三, 对风险传导的非线性动态过程模拟不足, 现有分析未能结合复杂网络理论或混沌模型揭示风险演化的深层规律。未来研究需结合跨学科方法, 深化对算法黑箱风险传导的动态建模, 为直播电商供应链金融的智能化治理提供更具操作性的理论支持。

参考文献

- [1] 林美吟. 供应链金融产品中资金流优化与风险管理对策研究[J]. 环渤海经济瞭望, 2024, (10): 154-157.
- [2] 葛文博. 基于物联网技术的供应链金融物流监管分析[J]. 新理财(政府理财), 2022(4): 22-24.
- [3] 祝由, 贾冉, 王纲金, 等. 供应链金融风险评估研究综述——基于知识图谱技术[J]. 系统工程理论与实践, 2023, 43(3): 795-812.
- [4] 张敏, 杨阁, 史一鸣. 直播电商模式下农产品直播供应链风险识别与测度研究[J]. 物流科技, 2024, 47(17): 121-126.
- [5] 谢文静, 鲍新中, 高鸽. 基于电商平台的供应链金融: 模式、典型案例与特征分析[J]. 科技促进发展, 2021, 17(1): 97-105.
- [6] 金灏. 直播电商中的物流供应链协同优化策略探讨[J]. 物流科技, 2025, 48(4): 120-122, 126.
- [7] 程雪军. 金融科技平台算法黑箱的法律规制研究[C]//上海市法学会. 《上海法学研究》集刊 2023 年第 5 卷——2023 年世界人工智能大会法治论坛论文集. 2023: 203-217.
- [8] 殷继国. 人工智能时代算法垄断行为的反垄断法规制[J]. 比较法研究, 2022(5): 185-200.
- [9] 程雪军. 生成式 AI 下超大金融服务平台滥用算法权力的风险规制[J]. 上海财经大学学报, 2024, 26(6): 122-136.
- [10] 张恩典. 算法透明度的理论反思与制度建构[J]. 华中科技大学学报(社会科学版), 2023, 37(6): 29-40.
- [11] 易靖韬, 严欢. 基于小波分解和 ARIMA-GRU 混合模型的外贸风险预测预警研究[J]. 中国管理科学, 2023, 31(6): 100-110.
- [12] 胡书逸. 基于音视频挖掘的电商主播非语言线索对直播带货绩效的影响作用研究[D]: [硕士学位论文]. 杭州: 浙江财经大学, 2024.
- [13] 刘世强, 李碧, 王伟锋. 基于大数据与区块链融合的供应链金融创新模式及应用研究[J]. 中国经贸导刊, 2025(4):

10-12.

- [14] 邓克涛, 张贵红. 算法透明度: 从理论到实践的探索与反思[J]. 科学学研究, 2024, 42(7): 1354-1360.
- [15] 肖雄, 唐卓, 肖斌, 等. 联邦学习的隐私保护与安全防御研究综述[J]. 计算机学报, 2023, 46(5): 1019-1044.
- [16] 何俊辉. 不完全契约视角下的数据跨境流动沙盒监管机制研究[J]. 数字图书馆论坛, 2024, 20(4): 58-64.
- [17] 杨康, 张仲义. 基于复杂网络理论的供应链网络风险传播机理研究[J]. 系统科学与数学, 2013, 33(10): 1224-1232.
- [18] 朱孟进. 产业互联网平台主导的供应链金融风险研究[J]. 新金融, 2020(9): 53-57.