Advances in Social Sciences 社会科学前沿, 2026, 15(9), 491-500
Published Online September 2026 in Hans. https://www.hanspub.org/journal/ass
https://doi.org/10.12677/ass.2026.159777

Hans汉斯

信贷领域中信用评分算法技术的法律规制研究

姜金良，岳雨薇*

南京信息工程大学法学与公共管理学院，江苏 南京

收稿日期：2026年7月30日；录用日期：2026年9月11日；发布日期：2026年9月21日

摘　要

信贷领域中算法技术与信用评分的深度融合，改变了传统的信用评分工作机制，在提高了信贷服务效率的同时，也衍生了算法偏见、个人信息泄露、数据滥用等风险。我国现有法律体系已初步形成算法规制框架，但在实践中仍面临针对信贷领域中信用评分算法模型审查缺乏明确标准、动态监管缺位，审查主体分散，传统侵权规则适配不足等困境。为防范、化解信用评分中风险，可建立“准确性－公平性－可解释性”的算法模型审查机制，明确以央行为主的监管主体。界定多元主体责任范围，针对信贷领域信用评分算法涉及的多元责任主体，差异化适用无过错责任原则和过错责任原则，补齐救济制度短板。为信用评分算法技术的发展提供良好的法治保障。

关键词

信用评分算法，算法偏见，算法模型审查，权利救济

Legal Regulation of Credit Scoring Algorithm Technology in the Credit Sector

Jinliang Jiang, Yuwei Yue*

School of Law and Public Administration, Nanjing University of Information Science and Technology, Nanjing Jiangsu

Received: July 30, 2026; accepted: September 11, 2026; published: September 21, 2026

Abstract

The deep integration of algorithmic technology with credit scoring in the credit sector has transformed traditional credit scoring mechanisms, enhancing the efficiency of credit services while simultaneously giving rise to risks such as algorithmic bias, personal information leakage, and data misuse. Although China's existing legal framework has preliminarily established a regulatory

*通讯作者。

文章引用: 姜金良, 岳雨薇. 信贷领域中信用评分算法技术的法律规制研究[J]. 社会科学前沿, 2026, 15(9): 491-500.
DOI: 10.12677/ass.2026.159777

structure for algorithms, practical challenges persist, including the lack of clear standards for model review of credit scoring algorithms, the absence of dynamic supervision, fragmentation among regulatory authorities, and insufficient adaptation of traditional tort rules. To prevent and mitigate these risks in credit scoring, it is advisable to establish an algorithmic model review mechanism based on the triad of "accuracy-fairness-explainability", with the central bank designated as the primary regulatory authority. It is also necessary to delineate the scope of liability for the multiple responsible parties involved, and to apply the principles of no-fault liability and fault liability differentially to these parties, thereby addressing the shortcomings of the existing redress system. These measures will provide robust legal safeguards for the development of credit scoring algorithm technologies.

Keywords

Credit Scoring Algorithms, Algorithmic Bias, Algorithmic Model Review, Rights Redress

1. 问题的提出

随着科学技术的发展，信贷领域使用算法技术进行信用评分已逐步常态化、规模化。《金融科技发展规划（2022-2025 年）》把“完善现代化治理结构，强化金融科技治理顶层设计，建立健全稳妥发展金融科技，加快金融机构数字化转型”列为重点工程[1]，这也表明智能信用评分已成为我国金融领域的重要组成部分。在未来社会发展中，我国会逐步向智慧国家、智慧城市的方向进行转型升级，征信是其中重要一环。将算法技术运用于信用评分工作，可以在短时间内完成评分提高工作效率，有利于缩短贷款业务办理周期；并且算法模型更具有客观性，摒除了人为的主观判断因素，可以避免因为工作人员主观评判带来的不公平问题。

信用评分算法的应用虽提高了评分效率和公平度，但也衍生出了新的风险。其一，信用评分算法的精准性受数据质量、算法技术、运行逻辑等因素影响，任何环节出现偏差都有可能导致结果错误。其二，在算法模型规模化应用下，算法歧视和算法偏见会强化社会不公[2]。其三，算法模型依赖大量数据，其中存在数据采集不规范，数据泄漏等风险。其四，算法的不透明性和不可解释性使用户难以获取关键信息，进而导致用户维权困难[2]。基于算法的专业性，以及征信机构和用户能力悬殊等问题，对于信用评分算法的规制需要健全法律规范体系，实现对算法技术的有效的规制，推动算法技术在法治轨道上发展和运行，构建高效、公平、安全的征信体制。

2. 信贷信用评分算法技术的构建及侵权风险

2.1. 信用评分算法技术构建的基本逻辑

信用评分算法以用户大量的信用相关数据为基础，例如消费记录，公共缴费记录，就业经历、还款记录等，通过深入的数据分析，对用户的还款能力以及偿还意愿进行评估[2]，为是否放贷以及贷款数额、贷款利率提供依据[3]。信用评分算法模型的构建主要包括数据采集－数据处理－模型训练三大核心环节。各环节层层递进，互相支持，共同组成信用评分算法的完整技术链。

第一部分，采集表征个人信用的数据，这是算法模型构建的第一步[4]。采集数据的类型应围绕着与信贷方面个人信用评分有关的内容，不同类型的征信主体采集方式也有所不同。央行作为公共征信系统

要求金融机构提供征信数据[5]；征信机构通过当事人授权、数据合作及互联网数据采集的方式获取数据信息[5]；互联网公司通过采集用户在其平台的信息进行信用评价[5]。

第二部分，数据处理，这是将原始数据整合为便于模型训练数据的重要环节。信用评分算法与传统信用评分机制相比，其数据具有多源异构性，导致数据质量高低不均、数据缺失值、噪声数据等难题，数据处理也成为算法模型搭建过程中的重要一步[6]。数据处理主要包括数据清洗，数据缺失处理和噪声过滤。其一，数据清洗是通过自动化识别与修复数据当中的错误提高数据的完整性与一致性[7]。其二，对数据缺失进行处理，处理方式包括删除法和插值法，数据缺失少时使用删除法，对数据缺失较多的使用插值法[8]。其三，噪声过滤是指通过识别和清除噪音提高数据整体质量[7]。

第三部分，模型训练，这是数据和算法结合生成可用模型的重要步骤。该部分主要包括三部分：第一，根据数据特点与风险评估目标，选择合适的模型。第二，通过机器学习算法模型能不断自我学习，机器学习的过程便是筛选变量和变量加权不断优化的过程[9]。第三，最终确定算法模型还需经过模型验证，利用测试集评估模型性能，并不足部分进一步优化，通过反复模型训练，验证与优化，不断提升模型各方面特性[10]，最终达到能够投入市场使用的标准。

2.2. 信用评分算法技术的侵权风险

1) 信用评分算法偏见侵害用户平等权

信贷领域中，信用评分算法偏见已成为侵害用户平等权的突出问题，信用评分算法偏见是指运用算法进行信用评分时，因算法技术存在缺陷，导致信用评分输出结果系统性地偏向特定群体、个体，形成不合理差别对待[3]，侵害用户平等贷款权。信用评分结果直接影响其信贷情况，可能导致贷款被拒或者影响贷款数额等，进而侵害用户的财产权。信用评分算法偏见既可能源于算法设计者的偏见植入，也可能源于带有偏见性的训练数据[3]。算法偏见会导致信用评分结果缺乏客观性，可能出现部分用户事实上具备还款能力，由于信用评分结果有失偏颇，仍面临拒绝放贷，贷款额度受限等情况，美国一家非营利机构进行了一次测试，该机构工作人员在一家名为 UPSTART 的“算法放贷人”平台以毕业院校为变量多次申请贷款，发现在其他条件相同的情况下，霍华德大学毕业生要比纽约大学毕业生负担更多的利息和其他费用[11]，这表明霍华德大学毕业生受到不公平对待，这是对用户信贷平等权的侵害，同样也会损害金融市场的公平性[9]。

2) 信用评分数据采集程序不规范侵犯用户知情权

根据《个人信息保护法》[1]第 13 条规定，个人信息采集应需取得个人同意。但实践中，数据采集存在不规范的问题，例如部分平台将电商推荐功能授权采集的用户购物频率、个人喜好等数据用于内部信用评分模型，对于这些数据的二次使用并未告知数据，未获授权[5]，这直接导致对用户知情权的侵害。并且由于算法模型的训练数据规模化、匿名化特征，传统告知同意机制难以实质落地[12]，具体表现为涉及数据庞大，数据主体分散，逐一获取单独同意成本较高，传统的告知方式难以获得主体的有效同意[13]，这进一步增加了对侵犯用户知情权的风险。

3) 信用评分数据滥用侵犯用户隐私权

算法模型涉及大量的数据，其中不乏数据滥用行为，这对用户隐私权构成极大威胁。一方面，算法设计者将于信用评分的数据纳入采集范围，违背数据使用的目的，充分必要原则，突破法律对隐私权的保护边界，另一方面，算法设计者违反法律要求，采集用户的隐私信息。数据滥用不仅侵害用户的隐私权，并且可能进一步侵害用户的财产权，名誉权等合法权益。

[1] http://www.npc.gov.cn/npc/c2/c30834/202108/t20210820_313088.html

3. 信贷信用评分算法技术的法律规制现状及实践困境

我国有关信贷信用评分算法的规制已初步形成以《征信业务管理办法》[2]《个人信息保护法》《中华人民共和国数据安全法》[3]为核心，《商业银行资本管理办法》[4]《信用评级业管理暂行办法》[5]等专项立法进行专门规范的法律体系，对监管主体，模型审查，数据使用等关键环节作出了规定，但由于现有法律只是初步进行了规定，尚有不足和空白，未能充分回应信贷领域中信用评分算法的风险。

3.1. 监管主体分散缺乏协同治理机制

我国现有法律体系针对信用评分规定了多元主体监管权限，不同的监管规定散落于不同的法律文件中，例如，《征信业管理条例》规定中国人民银行有权监督管理征信业。《商业银行资本管理办法》规定银行自行估计违约概率的零售评分模型须报银保监会验收。《信用评级业管理暂行办法》规定，发改委、财政部、证监会在各自债券监管领域依法行使评级模型检查、处罚权。这些规定赋予了多元主体权利，却未构建起有效协同治理机制，这导致监管边界模糊，易引发重复监管、责任互相推诿等问题。

3.2. 算法模型审查多为框架性规定，动态监管缺位

从现有法律体系来看，《征信业务管理办法》规定征信机构要对算法模型进行内部测试和评估验证程序，并且初步规定了向中国人民银行或其省会(首府)城市中心支行以上分支机构报告的事项。但现有法律仅规定了审查备案事项，未建立起明确的审核标准，这也就导致备案易流于表面，实效性有限。实践中，算法模型问题往往在投入使用后，造成侵权后才暴露，这种事后修改的方式不仅成本更高，而且对用户权益、金融秩序造成的影响更恶劣。此外，我国也缺乏对评分模型投入使用后的动态审查规定，亟待构建信用评分算法模型的全周期动态监管制度。

3.3. 信用评分数据管理规定模糊，权利保护失衡

在数据采集环节，我国法律作出了当事人授权的核心规定，并确立了数据采集的“最小、必要”基本原则。然而，传统的“告知–同意”机制在算法场景下难以落地[12]，“二次使用”等超出原授权范围的使用也违背了当事人授权的规定[5]，这对法律如何有效保障用户的实质知情权提出了挑战。此外，现有规定多为原则性表述，未明确“最小”“必要”的界限，这就导致数据采集易超过必要限度，侵犯用户权益，还可能由于数据关联性不足影响结果准确性。

3.4. 征信异议实效性不足，算法侵权责任认定规定滞后

在我国现有法律规定下，信贷主体常通过对征信机构提出异议和提起诉讼的手段寻求救济。其中，向征信机构提出异议是最直接的方式，但在实践中征信机构回复质量参差不齐，实效性也存在争议。此外，诉讼是信用评分算法错误的最有力的救济手段。但由于信用评分算法技术涉及的主体众多，算法技术具有的隐蔽性等特性，传统的侵权责任主体和归责原则无法满足司法需求，法律应进一步完善算法侵权责任主体认定和明确归责原则。

4. 构建算法模型的“准确性–公平性–可解释性”审查机制

4.1. 建立以央行为中心的协同审查主体

完善信用评分的法律规制首要任务是厘清有权主体权责边界。从我国国情出发，我国征信业务监管

[2]https://www.gov.cn/zhengce/zhengceku/2021-10/01/content_5640685.htm
[3]http://www.npc.gov.cn/c2/c30834/202106/t20210610_311888.html
[4]https://www.gov.cn/zhengce/202311/content_6913410.htm
[5]https://www.gov.cn/zhengce/2019-11/26/content_5712729.htm

以政府为主，第三方机构作为权力主体负责审查在我国并不实际。针对当下多元主体权责边界模糊的问题，建议构建以央行为主导的协同监管体系。首先，将审查权限统一收回中国人民银行，由央行进行统一管理，可以实现法律上的协调统一。通过立法明确央行在个人信用评分算法模型审查中的主导地位，并细化其职责范围。制定专门针对个人信用评分算法模型审查的规定，规定央行对算法模型的准确性、公平性、可解释性进行全面专业的审查。

在具体的组织架构方面，央行内部可设立专业的审查组，审查组又可细分为法律审查小组、技术审查小组、金融评估小组，成员包括金融专家、技术专家、法律学者等，对算法模型进行全面专业的技术评估，伦理评估，法律合规性评估等。在专家遴选方面，建议设立专家库，包括金融领域、计算机领域、法学领域等相关领域的专家。在具体项目的审查中，应根据具体的项目类型和特点，回避存在利益关系的专家，最终确定审查组成员，以保证审查的专业性和中立性。专家应对审查结论负责，若因过错导致未达到标准的算法模型通过审查，需承担行政责任，乃至民事责任。而在资金方面主要依靠政府财政保障，同时可向信用评分机构(被审查方)收取适当费用。综上，最终形成专业、高效的审查机制。

央行进行准入审查之后，与其他相关部门建立跨部门的信息共享平台，例如证监会，地方有关部门等共同建立信用档案，并纳入全国信用信息共享平台。在之后的动态监管中，以央行为主体，其他部门联合，共同制定审查标准和流程，共享审查结果，避免重复审查和责任推诿，打造高效的，统一审查监管体系。

4.2. “准确性－公平性－可解释性”的审查基准确立与协同路径

我国开展智能征信较晚，对于算法模型使用之前的准入审查机制，尚处于发展阶段，针对算法模型的具体事前审查，部分国家发展较快，已经作出了较为完善的规定。以欧美为例，其中欧盟《人工智能法案》作为典型立法具有重要借鉴意义，《人工智能法案》以风险分级为核心，对不同风险的人工智能在透明度、数据安全、可解释性等方面进行差异化治理以保障人工智能安全发展[14]。其中，信贷属于高风险领域，高风险人工智能系统是法案规制的重点，在风险管理、数据治理、透明度等各方面进行了全面规制[15]。欧盟为保证算法的安全性，规定相对来说较为严苛，这一定程度上会打压算法的创新性[16]。欧盟《消费信贷指令》则规定消费者有权要求对信用评分全面了解，并有权要求解释、复查[9]，这为消费者提供了法律保障。相较之下，美国的法律规范则更偏向于鼓励技术创新[16]，美国的《公平信用报告法》限制个人征信机构对信用信息的使用，以保障用户的隐私权，对消费者报告的公平性也有所要求[9]。《公平信用机会法》则明确禁止种族、性别等与信用评分无关的歧视性因素[9]。美国联邦和州级立法机构通过了《2022 算法问责法案》，进一步推动了算法透明、公平性以及权益保护[17]。但美国的立法较为分散，不同辖区管理标准存在差异[16]，法律实效性有限。不难发现，虽然欧美对于信贷领域信用评分算法的侧重点和规制方法均有所不同，但其对于算法技术的关注点主要集中在公平性，准确性，可解释性方面。

但是信贷领域中信用评分算法所追求的“准确性”“公平性”“可解释性”三者之间也存在着价值冲突，厘清三者之间的关系是建立审查机制的基础。信用评分算法模型主要包括线性评分模型和采用机器学习的算法模型，涉及判别分析法、线性回归、逻辑回归、梯度提升决策树模型和 BP 神经网络模型等[2]。不同的算法模型有其不同的优势和风险，传统的线性评分模型透明度高，可解释性强，但其准确性不足。而依托于机器学习的复杂算法模型，可以处理替代性数据，准确度较高，但其较为复杂，可解释性较弱，并且由于纳入替代数据，引发不公平的风险也更高[2]。因此，对于技术的使用需要价值加以引导和规制。在信贷这一领域中，信用评分算法应形成“以公平性为基石、准确性为目标、可解释性为保障”的动态平衡。首先，信用评分算法的公平是重中之重。由于算法模型一旦投入将会规模化使用，与

传统信用评分模式相比，一旦算法模型带有偏见和歧视性，那么会进一步固化偏见性，社会影响恶劣，在我国发布的政策《关于加强互联网信息服务算法综合治理的指导意见》中也提出引导算法应用公平公正是未来发展的重要内容[18]，保证算法的公平性是第一要务。除此之外，信用评分的结果是贷款审批的核心依据，会直接影响到用户的贷款情况，结果的准确性是信用评分算法的核心要求。可解释性则是对前两者的保障，算法本身存在“黑箱”特性，算法技术的“黑箱”是指算法系统中输入与输出之间的因果关系和技术逻辑公众难以理解[19]，导致自动化决策缺乏透明性[20]，这种不透明性可能侵犯用户的知情权、个人信息安全等[19]，并且算法“黑箱”也会阻碍用户维权，因此对算法技术的可解释性审查是保障用户权利的要求，是用户寻求司法救济的抓手。因此，在信用评分算法模型设计时要兼顾三者，寻找平衡点。

4.3. 构建算法模型的“准确性－公平性－可解释性”审查标准

1) 算法模型的准确性审查－形成内外双重监管核验体系

对于算法模型准确性的审查需要算法开发主体的内部测试和监管主体的外部监管共同发力，研究表明，建立信用评分算法自我评估机制能够有效降低风险概率[21]。信用评分算法模型的开发者应进行内部测试，由算法模型开发者提供个人信用评分算法模型的算法模型搭建逻辑，算法正确率等数据，形成算法模型的测试报告报送监管主体[21]。然后，再由审查主体针对算法模型的各项数据进行审查。审查主要包括设立审查标准和审查程序两部分，审查标准由中国人民银行组织专业人员，结合算法模型结果影响，使用领域等方面，设置统一的市场准入标准，即算法模型的准确度要达到准入区间才可投入使用。设立统一的审查程序，即算法模型必须通过设定好的审查程序，根据指标是否达到市场准入门槛作出审查结果。

2) 算法模型的公平性审查——“数据－技术”全面把控

在信贷领域中，公平性不仅是算法技术的内在要求，更是包含贷款主体可以平等的享受贷款机会的内涵[9]。在信贷领域的公平是指所有贷款主体，不论其经济实力、种族、地域、行业等，都能在申请贷款时获得平等的机会[22]，以寻求更好的发展。在个人信用评分中，算法不公的成因主要在训练数据存在历史性歧视，数据范围小和算法设计不当这三方面，基于此，算法模型的公平性审查应从历史数据选用，算法设计逻辑两大方面展开。

第一，由算法设计者提供数据采集报告，包括数据采集程序，采集范围，采集来源等，组织专家对数据采集的合法合规性进行全面审查。

有效审查的前提是明确的审查标准，我国现有的法律规定过于宽泛，需要进一步细化保证监管的有效性。第一，细化数据采集原则中“最小、必要”的范围。我国《征信业务管理办法》规定征信领域中信息采集的应当采取合法、正当的方式，遵循最小、必要的原则，不得过度采集。但随着算法技术的发展，最小、必要原则的具体界限也越来越模糊，例如杨帆学者在《个人信用分平台企业大数据应用的法律规制》一书中所说随着技术的更新迭代，一些看似与个人信用评分无关的信息也可以体现出用户的信用情况[23]。因此，法律应明确“最小、必要”的边界，应进一步强调“必要”要求数据与信用评估具有稳定的逻辑关联，“最小”则强调采集的范围应是“实现目的”的最小范围[24]，不得过度侵害隐私、敏感信息[24]。随着技术的发展，当下的信用评分算法更为复杂，涉及到大量替代性数据的使用[24]，对于此在正面规制外，可从反面列举限制采用、禁止采用的无关联数据类型，并进行动态变更。例如吴晓晨从公平性角度出发，指出应排除使用用户个性信息、社交网络关系信息，限制采用反映个人品格的信息[24]，对于涉及种族、性别、地域等信息也应直接排除使用。第二，通过数据分级实施差异化管理。数据的分级管理根据数据使用的风险性对一般数据、重要数据、核心数据、敏感数据作出差异化的使用限制，从

源头上避免数据使用不当影响算法模型的公平性。当前我国数据分级管理多为一般性规定，但中国人民银行发布的《金融数据安全 数据安全分级指南》[6]较为全面列举了各种数据的分类，信贷亦属于金融领域，可以此为参照，结合信贷领域信用评分算法模型的特殊性进一步评估和完善，并将其确定为具有强制效力的行业统一标准。在管理上可借鉴欧盟 GDPR 规定在数据处理前进行数据保护影响评估以识别其风险性[25]，并据此进行管理。具体使用规则建议如下：核心数据，有危害国家安全风险的数据以及《个人信息保护法》明确列举的敏感信息原则上禁止使用，特别使用向中国人民银行进行申请；重要数据及其他敏感信息限制使用，应充分告知用户其权益和后果，获得其同意并且进行报备。一般信息限制相对宽松，获得用户同意，并一同报备即可。第三，数据公平性审查除了采集范围及程序合法合规外，还应重点审查数据集的公平性，例如是否可代表社会群体结构，是否带有歧视性信息等，对有疑问的数据算法模型的设计者应进行充分解释以通过公平性审查。

第二，算法技术的公平性审查除数据层面的审查外，应组织技术专家，法学专家对信用评分算法模型技术层面展开审查，对算法架构、运行逻辑、关键技术、特征参数等进行全方位评估，判断其是否符合算法模型的公平性要求，其中可使用专业技术手段加以辅助审查，最后形成审查报告。

3) 算法模型的可解释性审查－从备案到强制规制

算法模型的可解释性是指算法的决策过程能够进行清晰有效的解释，用户充分理解算法输出结果的依据和逻辑[26]。算法模型的可解释性是避免“黑箱算法”的重要手段，也是用户申诉时，能够进行追溯的基础。

算法模型的可解释性是各国共同关注的议题，域外立法主要集中在设定事前准入标准和事后解释义务两方面，例如，欧盟《人工智能法案》规定对于高风险人工智能系统的设计和开发的透明度管理较为严格，信贷领域的信用评分算法在进入市场前，需满足透明度、测试与报告要求[16]。美国的《平等信用机会法》则是要求信贷机构对用户的“不利行动”做出实质解释，例如影响信用评分的因素和权重等[27]。我国现有规定则倾向于算法设计者的内部检测，多以备案形式进行管理。例如，中国人民银行发布的《人工智能算法金融应用评价规范》对算法模型可解释性的作出了规范，具体包括特征选择可解释，算法可解释，参数可解释[28]，但缺乏硬性审查标准。但这种方式依靠从业者的自觉性，规制效果有限，完善审查机制应将对算法模型的备案要求转为对算法模型的审查内容的硬性标准，核心在于审查特征选择、权重设置、算法技术等关键信息，并向用户公开，旨在帮助公众理解算法的工作原理及其运行方式。

算法可解释性要求并不意味将算法全部披露。征信机构投入大量的人力财力用于信用评分算法开发，算法是其核心竞争力，企业往往将算法视为高度机密的商业资产[26]，我国法律也支持将算法作为商业秘密进行保护。若将算法全部披露不仅可能侵犯商业秘密，损害征信机构利益，还可能由于算法透明公开，增加算法被攻击被破坏的风险[29]。因此增强算法可解释性的同时，应权衡算法公开对商业秘密、征信机构利益、算法安全的影响[26]。对此，可采取分级披露制度，对于数据来源、特征选择、影响逻辑、潜在利益冲突[30]等与算法功能相关的关键信息，应向公众披露，以保障公众知情权，提升算法透明度和公众理解度[21]。而对于具体的算法技术、源代码、详细的参数设置等涉及商业秘密的部分不予披露，参与审查的专家应签署保密条款，以确保算法核心技术仅被必要人员知悉。但在侵权救济时，不得以保护商业秘密为由拒绝向特定主体披露必要部分。另外，在技术审查中，针对不同类型的算法模型实施差异化审查，线性评分模型其透明度本身较高，应重点审查其解释公开内容是否真实有效。而对于较为依托于机器学习的复杂算法模型，除了审查其解释公开内容的真实性之外，应对算法技术和模型构建进行严格审查，避免因为技术壁垒使审查流于表面。

[6]https://std.samr.gov.cn/hb/search/stdHBDetailed?id=B081D125A6762DB8E05397BE0A0A5EA7

经过审查之后应当生成审查报告，一方面，审查报告进行存档，保证事后可追溯，并注明后续动态监管的重点关注内容，便于工作展开。另一方面，审查报告应向市场公开，保障公众知情权。其一，在保护商业秘密的基础上，详细阐述算法的运作机制，增强公众对算法模型的信任度。其二，由专家评估算法的潜在风险，并向公众说明，确保公众的知情权。

5. 信用评分算法侵权的责任主体认定与归责原则优化

健全的救济渠道是保障公民权利的重要部分，无救济则无权利。但现有的提出异议和提起诉讼两大救济手段在实践中都存在着一定困境。一方面，异议救济由于缺乏明确的异议回复标准而流于表面。另一方面，信用评分算法技术涉及主体众多，具有“黑箱”算法，技术壁垒等特性，导致侵权主体认定困难，归责原则难以适用，对司法救济提出了新的挑战，亟需法律作出回应。

5.1. 细化异议回复标准，提高救济效率

贷款人向信用评分主体提出异议是最直接的救济方式，《征信业管理条例》第四章异议和投诉中有所规定，但没有明确回复标准，征信机构可能模糊回应，贷款人难以获得有效信息，导致异议救济流于表面。

对此，立法主体需进一步明确征信机构的回复标准，包括核查方式、核查结果、评分依据、补偿措施等核心要素。如核查无误，征信机构应告知信用评分流程，针对异议进行合理说明。如核查有误，应及时告知错误原因，原因应具体，不能简单用操作失误等表述回复[27]，并积极实施补偿措施。明确回复标准有助于统一管理征信机构行为，避免异议回复流于形式，增加异议救济的实效性。

5.2. 构建信用评分算法责任主体链，界定多元主体责任范围

对于算法的责任主体一直是学界关注的焦点问题，主流观点还是否认人工智能的独立地位，主张算法背后的主体承担责任，欧盟《人工智能法案》提出以算法价值链为中心的众多主体责任的配置[31]，在信用评分算法应用中，包括算法控制者(信用评分机构、平台)，算法设计者、数据提供者三类主体[32]，对于多元主体责任的界定需要对各主体在侵权链条上的角色、能力、收益等方面进行全方位衡量。

首先，需要明确的是信用评分算法应用中，算法控制者对信用评分算法的开发、设计、应用与优化等方面都具有强支配力，应承担主要责任[30]。算法控制者在算法技术投入使用之前需进行风险测试等工作[31]，在运行过程中应持续监测，维护算法的正常运行[33]。其次，算法设计者应保证算法模型全面符合“准确性–公平性–可解释性”标准，对由于算法设计缺陷导致的侵权承担责任[33]。最后，数据提供者应当保证数据来源合规合法且真实，并保护数据安全[33]，对于采集数据时侵害个人隐私、泄露个人敏感信息的行为负责。

5.3. 信用评分算法归责原则的差异化适用与正当性证成

对于信用评分算法的归责原则主要有“产品责任说”“过错责任说”和“无过错责任说”三种主张。产品责任说的前提是将人工智能认定为产品，但信用评分算法显然不具备规模销售的特性，难以认定为《中华人民共和国产品质量法》[7]中的产品。过错责任说则认为将其视为一般侵权行为更符合现有法律体系，并且可以激励智能产业创新[31]。无过错责任说则主张算法技术的自主性、透明度不足等特性导致过错认定困难，传统侵权责任法中的过错责任原则难以适用[34]，适用无过错责任原则更利于救济。在信贷场景下，信用评分算法的责任主体分散且发挥的作用不同，对所有主体使用同一归责原则难以兼顾技术安全与发展。因结合多元主体的定位与能力，确立差异化的归责体系，即算法控制者作为主要责任承担

[7]https://www.samr.gov.cn/zfjcj/tzgg/art/2023/art_579118cd202a45fba28b7edfd9f6fd72.html

主体，适用无过错责任原则，对数据提供者、算法设计者适用过错归责原则。

算法控制者适用无过错责任原则的正当性从以下三个维度展开。第一，从非对称风险理论出发，非对称风险理论认为，加害方承担侵权责任的依据不是不法行为，而是风险[35]。对于高风险领域适用无过错责任原则可以矫正算法控制者制造的非对称风险[35]。《互联网信息服务算法推荐管理规定》[8]明确要求，根据算法推荐服务的舆论属性或者社会动员能力、内容类别、用户规模、算法推荐技术处理的数据重要程度、对用户行为的干预程度等，对算法推荐服务提供者实施分类分级管理。该标准为划分算法风险等级的提供了重要参考[36]。基于此，信贷领域信用评分场景下算法的应用，涉及的人群广且与社会公众的财产直接关联，应认定为高风险等级，算法控制者应对风险负责，对其适用无过错责任原则具有正当性。第二，基于报偿理论，算法控制者通过信用评分算法投入市场获取巨额经济利益，获益的主体理应承担相应的风险[35]，并且其更有能力承担责任，适用无过错责任原则更加契合这一场域特性。第三，无过错责任原则的适用有利于减轻证明责任，降低诉讼成本[31]。并且还可以增强算法控制者的责任感，使其工作更加严谨、合规。虽然对算法控制者适用严格的无过错归责原则，但并不意味其要对所有侵权后果承担责任，对于不可抗力、技术的局限性、受害人故意等导致的侵权，不承担无过错责任。

对于算法设计者和数据提供者应当适用过错责任原则，二者作为算法链接上的一环，对算法的控制程度有限，从其定位、能力出发并不适合适用过错责任原则更为合适。算法设计者和数据提供者若未尽到前文所述注意义务导致用户权益受损，可推定其存在过错[32]，应承担相应责任。

综上，归责原则的差异化适用既能保障用户的救济权利，也避免打压技术创新的积极性，实现技术安全与发展的动态平衡。

6. 结语

信贷领域的信用评分算法作为数字经济时代金融服务创新的核心引擎，提高了信贷审批的效率与模式，为金融普惠提供了技术可能。然而，算法的黑箱属性、技术更新迭代的快速性与传统法律规范的滞后性之间的矛盾，出现规范不足，监管不力的情况，本文基于对现有法律的梳理，从构建算法模型的审查机制和事后救济两大部分针对性地提出合理性建议，以保障金融消费者的合法权益和金融市场的公平秩序与公信力。

基金项目

本文系江苏省教育厅哲学社会科学一般项目(2022SJYB0169)“网络黑灰产的刑法规制研究”的阶段性成果。

参考文献

[1] 金融科技发展规划（2022-2025 年）[EB/OL]. https://www.pbc.gov.cn/zhengwugongkai/4081330/4406346/4693549/4470403/index.html, 2022-01-21.

[2] 张博文, 杨斯尧. 算法型评分工具: 优势、风险与法律规制[J]. 西南金融, 2022(9): 33-44.

[3] 宁园. 算法歧视的认定标准[J]. 武汉大学学报(哲学社会科学版), 2022, 75(6): 154-165.

[4] 李晓楠. 大数据技术下个人公平征信监管的数据治理维度[J]. 大连理工大学学报(社会科学版). 2023, 44(2): 65-74.

[5] 陈骏. 关于互联网金融背景下征信数据采集的思考[J]. 西部金融, 2018(11): 42-45.

[6] 李慧. 人工智能在信用评分模型中的技术突破与应用[N]. 企业家日报, 2025-07-22(006).

[7] 刘小松. 人工智能技术驱动下的信息处理优化方法探讨[C]//中国智慧工程研究会. 2025 工程创新与可持续发展

[8] https://www.moj.gov.cn/pub/sfbgw/flfggz/flfggzbmgz/202305/t20230509_478388.html

经验交流会论文集(下). 2025: 377-379.

[8] 陆健健. 基于集成学习算法的个人信用评分研究[D]: [硕士学位论文]. 上海: 上海工程技术大学, 2020.

[9] 何颖. 信用评分人工智能算法公平性检讨与规制[J]. 东南大学学报(哲学社会科学版), 2025, 27(3): 47-55+154-155.

[10] 傅振. 人工智能驱动的银行信贷风险经济模型构建与应用[C]//重庆市大数据和人工智能产业协会, 重庆建筑编辑部, 重庆市建筑协会. 智慧建筑与智能经济建设学术研讨会论文集(二). 2025: 278-281.

[11] Arnold, C. (2020) Graduates of Historically Black Colleges May Be Paying More for Loans: Watchdog Group. https://www.npr.org

[12] 梁远高. 论人工智能大模型训练数据风险的分层规制[J]. 郑州大学学报(哲学社会科学版), 2025, 58(3): 61-67+144.

[13] 刘国城, 潘颖轩. 数字化时代人工智能安全的风险样态与审计治理研究[J]. 兰州学刊, 2026(2): 134-149.

[14] 彭建军, 邓杰轩. 欧盟、美国人工智能治理比较分析及启示[J]. 领导科学, 2026(4): 154-160.

[15] 张麒. 欧盟《人工智能法》的实施压力与规则外溢[J]. 电子科技大学学报(社科版), 2026, 28(3): 55-69.

[16] 刘凡, 武良军. 人工智能立法的模式比较与中国的路径选择[J]. 安徽大学学报(哲社版), 2026, 50(2): 110-120.

[17] 田日胜. 欧美人工智能算法的立法比较[J]. 中国价格监管与反垄断, 2026(7): 96-98.

[18] 关于印发《关于加强互联网信息服务算法综合治理的指导意见》的通知[EB/OL]. https://www.cac.gov.cn/2021-09/29/c_1634507915623047.htm, 2021-09-17.

[19] 金雪涛. 算法治理: 体系建构与措施进路[J]. 人民论坛·学术前沿, 2022(10): 44-53.

[20] 袁曾. 算法黑箱的治理迷思与破解[J]. 中国海商法研究, 2025, 36(4): 22-31.

[21] 上官修瑾. 个人信用评分的法律规制研究[D]: [硕士学位论文]. 济南: 山东财经大学, 2025.

[22] 冯果. 金融法的“三足定理”及中国金融法制的变革[J]. 法学, 2011(9): 93-101.

[23] 杨帆. 《个人信用分平台企业大数据应用的法律规制》[M]. 上海: 上海人民出版社, 2023: 138.

[24] 吴晓晨. 个人信用信息处理的三阶规制[J]. 财经法学, 2026(1): 34-49.

[25] 刘小溪. 欧盟个人数据要素治理的制度体系: 基本架构与实施评价——基于《通用数据保护条例》的分析[J]. 价格理论与实践, 2026(5): 142-148+294.

[26] 康京涛, 陈至诚. 人工智能算法透明法治化: 底层逻辑与路径选择[J]. 武汉科技大学学报(社会科学版), 2026, 28(1): 50-61.

[27] 何新新. 算法决策解释权的法理构造与实践路径研究[D]: [博士学位论文]. 武汉: 中南财经政法大学, 2023.

[28] 人工智能算法金融应用评价规范[EB/OL]. https://std.samr.gov.cn/hb/search/stdHBDetailed?id=BF61550E44D6DEDDE05397BE0A0A63D6, 2021-03-26.

[29] 陈奇桢. 个人信用评分算法解释框架的构建[J]. 唐山学院学报, 2025, 38(5): 88-95.

[30] 杨帆. 信用评分算法治理: 算法规制与产品责任的融通[J]. 电子政务, 2022(11): 28-39.

[31] 朱紫薇. 主体与归责: 生成式人工智能侵权责任认定的困境检视与纾解路径[C]//中国智慧工程研究会. 2025 数字时代的社会结构变迁与治理创新学术交流会论文集(下). 2025: 476-479.

[32] 林洹民. 人工智能侵权的过错责任[J]. 法学研究, 2025, 47(5): 113-133.

[33] 王建群. 算法侵权的责任主体认定研究[D]: [硕士学位论文]. 上海: 上海师范大学, 2024.

[34] Araromi, M.A., Oluwabiyi, A.A., Olaleke, A.M., *et al*. (2024) Determination of Tort Liability in the Deployment of Artificial Intelligence Technology. *Journal of Law, Policy and Globalization*, **141**, 29-39.

[35] 王俐智. 论人工智能侵权的差序责任[J]. 华中科技大学学报(社会科学版), 2025, 39(6): 72-83.

[36] 陈兵, 董思琰. 生成式人工智能的算法风险及治理基点[J]. 学习与实践, 2023(10): 22-31.