

Construction of big data-driven early-warning index system for systemic financial risks

Yujie Zhang¹ Hairui Yan²

1. Wind Information Technology Co., Ltd., Beijing, 100010, China

2. Beijing Zhongming Guocheng Accounting Firm, Beijing, 100010, China

Abstract

In recent years, the financial system has grown increasingly complex, with systemic financial risks exhibiting distinct characteristics. These risks demonstrate strong contagion capabilities, where a single outbreak can rapidly spread throughout the entire system. They also exhibit high volatility, making risk levels difficult to accurately predict, and rapid diffusion capabilities that can cause widespread impacts within short periods. These features pose significant challenges to financial stability, making timely early warning systems for systemic financial risks particularly crucial. As the financial system continues to grow increasingly complex, systemic financial risks have demonstrated notable characteristics including high contagion potential, significant volatility, and rapid diffusion. This study conducts systematic research on early warning mechanisms for systemic financial risks. Leveraging big data technology advantages, we developed a risk early-warning indicator system through multiple approaches such as data integration, feature extraction, indicator screening, and model construction. The framework is built on multidimensional data foundations and oriented toward financial stability.

Keywords

big data; systemic financial risk; early warning index system; financial stability; financial supervision

大数据驱动的系统性金融风险预警指标体系构建

张钰洁¹ 闫海瑞²

1. 万得信息技术股份有限公司, 中国·北京 100010

2. 北京中名国成会计师事务所, 中国·北京 100010

摘要

近年来, 金融体系愈发复杂, 系统性金融风险的特征也日益凸显, 其传染性强, 一处风险爆发可能迅速蔓延至整个体系; 波动性大, 风险程度难以准确预估; 扩散速度快, 能在短时间内造成广泛影响。这些特征给金融稳定带来巨大挑战, 因此, 对系统性金融风险进行及时预警十分必要。随着金融体系复杂程度的持续上升, 系统性金融风险呈现出传染性强、波动性大、扩散速度快等显著特征。本文围绕系统性金融风险预警展开系统研究, 基于大数据技术优势, 从数据整合、特征提取、指标筛选、模型构建等多个角度出发, 构建了以多维数据为基础、以金融稳定性为导向的风险预警指标体系。

关键词

大数据; 系统性金融风险; 预警指标体系; 金融稳定; 金融监管

1 引言

在全球金融一体化背景下, 各国金融市场联系日益紧密, 金融体系的脆弱性愈发显现, 系统性金融风险呈现高频爆发和跨市场蔓延态势。本文基于大数据视角, 探讨系统性金融风险的识别路径与预警机制, 力图构建一套具备动态调整能力、指标科学严谨、模型稳定性强的预警指标体系, 为我国金融监管现代化、精准化发展提供理论支撑与实践方向。

2 系统性金融风险的特征与识别需求

2.1 系统性金融风险的定义与演化趋势

系统性金融风险, 指的是金融体系里某一或几个关键节点遭受冲击后, 风险迅速扩散至其他金融机构、市场或行业, 进而致使整个金融系统功能紊乱甚至全面瘫痪。它和单一机构或市场的个体性风险不同, 有着显著的外溢与扩散特性, 形成和演化过程充满高度的非线性、不确定性和反馈效应, 具体呈现出突发性、扩散性、放大性以及强烈的市场恐慌情绪等特点。

传统上, 系统性金融风险多聚焦于流动性枯竭、信用违约、金融机构破产等方面。但随着信息技术、金融创新和市场开放程度的提升, 风险成因与传导机制愈发复杂。金融

【作者简介】张钰洁(1997-), 女, 中国山西临汾人, 本科, 从事财务研究。

机构业务和资产负债结构复杂化,增加了风险隐蔽性,像资产错配、期限错配、高杠杆成为常见诱因^[1]。同时,影子银行、表外业务、金融衍生品交易等非传统金融活动扩张,打破了传统风险边界。

近年来,数字金融、平台金融、区块链技术和加密资产兴起,金融与科技界限模糊,平台型企业混业经营风险受关注,跨境金融活动增多也让本土市场更易受外部冲击。所以,构建能反映宏观微观动态变化、捕捉交叉性风险传导路径的指标体系很有必要,要涵盖多维度内容,结合大数据技术实现实时监控与动态分析。

2.2 传统预警机制的瓶颈分析

传统金融风险预警机制主要借助经济周期变量或单一市场指标构建早期预警模型,历史上发挥过一定作用,但如今金融体系复杂化,其局限性愈发明显。

其一,指标覆盖范围有限。传统模型侧重宏观经济与金融市场基本面指标,对金融体系内部机构间复杂的业务关系及潜在风险反映不足,影子银行体系、表外业务、跨市场业务等未纳入常规指标体系,存在风险识别盲区。

其二,数据更新滞后。传统监测依赖季报或年报数据,难以及时捕捉市场异动。在突发性金融事件发生时,缺乏实时数据支持,监管反应滞后,错过干预窗口期。

其三,静态数据处理模式。传统模型多采用经典计量经济方法,缺乏对非结构性数据(如新闻舆情、社交媒体信息)的利用能力,无法深度挖掘潜在风险特征。

其四,模型缺乏自适应能力。传统模型参数和结构设计后难以实时调整,面对市场环境变化缺乏灵活性和应对能力。如2023年某大型跨国银行暴雷事件,传统模型未能及时反映隐性负债和平台金融业务风险,导致监管滞后、市场恐慌扩大。因此,需借助大数据技术重构金融风险识别机制。

2.3 新技术驱动下的风险识别新路径

大数据技术为系统性金融风险识别带来了全新的识别与预警方式,优势明显。

在数据获取上,通过网络爬虫、API接口和大数据平台,能实时采集多源异构数据,突破传统数据局限,拓展了风险监测维度^[2]。舆情分析系统捕捉金融机构负面新闻,可提前察觉市场信心波动风险。

数据处理环节,结合机器学习等算法和技术,能挖掘隐藏特征与非线性关系。聚类分析等方法可动态分析金融体系内各节点风险关联与潜在脆弱性,形成动态风险图谱。

风险动态识别方面,大数据分析平台构建的预测模型能对历史风险事件复盘学习,结合实时数据预判未来走势,实现前瞻性识别。如利用相关数据预测潜在熔断风险,助监管提前干预。

此外,大数据风险识别机制还有自动化智能化、个性化定制化、协同化共享化等优势,推动金融监管从被动应对转向主动防控,提升金融体系韧性与稳定性。未来,需完

善相关方面,推动金融大数据风险预警体系广泛应用与深度融合。

3 大数据驱动的指标体系构建逻辑

3.1 指标体系构建的总体思路

系统性金融风险预警指标体系的搭建,要遵循“宏观—中观—微观”多层结构,全面覆盖经济运行、市场波动、机构稳健以及外部冲击等关键维度^[3]。构建时,科学性、系统性与动态性是核心准则,只有如此,才能保证指标体系精准、全面且能与时俱进。

在宏观层面,货币政策走向、财政状况好坏以及国际资本流动情况等宏观经济指标至关重要,它们是金融体系运行的宏观背景,直接影响着金融市场的稳定。中观层面,要聚焦银行、证券、保险等行业的运行态势,把握各行业发展的脉搏。微观层面,金融机构的杠杆水平、资产质量优劣以及流动性管理状况等核心财务指标不容忽视^[4]。此外,为提升预警体系的敏感性与适应性,还应引入市场心理、媒体关注度、舆情风险等非结构性变量,从多角度捕捉潜在风险,为金融市场的稳定运行筑牢防线。

3.2 大数据指标的采集与处理技术

在大数据驱动框架下,指标数据的采集与处理成为核心环节。数据来源应包括结构化数据(如中央银行公布的宏观经济数据、金融报表、市场行情)与非结构化数据(如新闻舆情、社交媒体评论、网络搜索指数等)。针对非结构化数据,需借助文本挖掘、情感分析等自然语言处理工具进行语义解析与量化处理,以构建反映市场预期与情绪变化的舆情指数。数据预处理阶段包括异常值识别、缺失值填补、数据平滑、标准化转换等步骤,以确保模型输入的稳定性与可比性。随后,利用主成分分析、随机森林特征排序、Lasso回归等方法对初步指标进行筛选与降维,确保最终指标体系具备代表性与可操作性。

4 预警模型构建与评估机制

4.1 模型框架设计

在搭建好系统性金融风险预警指标体系后,预警模型的构建便成为决定系统实际预警能力的关键环节。常见可供选择的模型类型丰富多样,像 Logistic 回归、支持向量机(SVM)以及长短期记忆网络(LSTM)等都有其独特优势。Logistic 模型在风险分类与概率预测方面表现突出,能清晰划分风险等级并估算其发生概率;SVM 模型在小样本、高维度的复杂场景下有着良好表现,可精准捕捉关键特征;LSTM 模型凭借其独特的记忆机制,在处理金融时间序列数据上独具优势,能有效把握数据的长期依赖关系。

模型训练时,要紧密结合实际金融数据设定标签,例如以系统性金融事件是否发生为依据,构建有监督学习框架。此外,还需引入交叉验证与滚动窗口机制,前者能充分利用数据,减少过拟合风险,后者可让模型适应数据的动态

变化,二者共同提升模型的鲁棒性与泛化能力,确保预警系统稳定可靠。

4.2 模型输出与阈值设定

预警模型输出有着明确分类,主要分为风险概率值与预警等级分类这两种形式。风险概率值意义重大,它能够以量化的方式精准呈现系统性风险程度的变化情况,让监管者清晰把握风险动态走向。而预警等级分类则更为直接,可依据不同等级迅速触发相应的监管响应机制,以便及时采取措施应对风险。

在确定预警阈值时,不能盲目设定,需综合考量多方面因素。既要参考历史数据所积累的经验,又要结合监管部门对风险的容忍程度,同时还要充分考虑模型预测的准确性。为进一步提升响应的灵敏度,避免出现对风险“高估”或“低报”的情况,还应引入自适应阈值机制。该机制能够根据市场的实时波动状况以及风险的发展趋势,动态地调整触发标准,使预警体系更加贴合市场实际,从而更有效地发挥预警作用,保障金融市场的稳定运行。

4.3 评估机制与模型迭代

为保障预警系统切实发挥作用,构建完整的模型评估机制必不可少。评估指标需涵盖准确率、召回率、F1值、ROC曲线、AUC值等多维度衡量体系。准确率体现模型预测正确的比例,召回率反映模型捕捉到实际风险的能力,F1值综合考量二者,ROC曲线和AUC值则从整体上评估模型区分风险的能力,多指标结合能全面呈现模型在不同风险场景下的表现。

同时,要设立定期回溯机制与动态更新机制。定期回溯可检验模型过去一段时间的表现,发现问题及时改进;动态更新则基于最新金融环境与数据变化,对模型进行重训练与结构调整,确保模型紧跟市场节奏,保持适应性与前瞻性,为金融风险预警提供可靠支撑。

5 实证分析与政策建议

5.1 实证数据与模型验证

为构建科学有效的系统性金融风险预警体系,本文聚焦我国金融市场,选取了2010年至2022年这一重要时间段的季度金融数据以及舆情数据。这些数据涵盖了金融市场的多个关键层面,为全面评估风险提供了丰富素材。基于这些数据,我们精心构建了一个包含60余项指标的系统性金融风险预警体系,这些指标从不同角度反映了金融体系的运行状况和潜在风险。

为了验证该预警体系的有效性,我们引入了先进的LSTM模型开展回测实验。实验结果表明,这一预警体系在应对重大金融事件时展现出了良好的前瞻性。在2015年股市异常波动以及2020年新冠疫情引发的流动性紧张等事件中,它都能提前发出预警信号。具体而言,模型对系统性风

险事件的平均提前预警期达到了3个季度,这意味着监管部门有较为充裕的时间来制定应对策略。而且,其准确率超过85%,与传统统计模型相比,优势十分显著。这一结果充分证明了本文所构建的预警体系具有较强的实用性和可靠性,能够为金融监管提供有力的支持。

5.2 优化金融监管体系的政策建议

基于上述实证结果,为进一步优化金融监管体系,有效防范系统性金融风险,我们提出以下建议。首先,金融监管机构应加快建设统一的大数据平台。当前,不同部门之间的数据存在壁垒,影响了信息的流通和共享。通过建设统一平台,能够打破这些壁垒,实现跨部门信息共享,让监管部门更全面地掌握金融市场的动态。

其次,要制定一系列规范,涵盖数据安全、算法透明度、模型责任划分等方面。在大数据和模型广泛应用的年代,确保预警体系的合规性与可控性至关重要。再者,构建多层次响应机制,将预警结果与金融政策工具动态对接,形成“预警—评估—响应—修正”的闭环管理,提高监管的针对性和有效性。最后,加强监管科技人才队伍建设,提升监管人员的数据素养与技术能力,为风险治理奠定坚实的人才基础。

6 结语

在当今复杂多变的金融环境下,系统性金融风险已成为威胁金融安全与经济稳定的重要因素。传统的风险识别与预警机制在多维风险交叉、多源数据爆炸的背景下,已难以适应监管新需求。大数据技术以其全方位的信息采集能力与强大的非线性分析能力,为构建科学、动态、高效的系统性金融风险预警指标体系提供了坚实支撑。本文围绕预警指标选取、数据处理方法、模型构建与政策响应等方面展开研究,提出了一整套以大数据为核心驱动力的风险监测框架。未来,随着人工智能、区块链、知识图谱等技术的发展,系统性金融风险管理将进一步走向智能化与精细化。监管部门应持续推动技术融合、机制创新与制度完善,构建更加全面、灵敏、协同的风险预警与处置体系,为实现金融高质量发展保驾护航。

参考文献

- [1] 高子航,黄新尧,袁新晨,等.地方政府债务与系统性金融风险的影响机制及政策选择[J].活力,2024,42(24):190-192.
- [2] 王韶华,杨铭,张伟,等.我国金融风险传染机制与防控策略研究——基于复杂网络下非连续治疗策略的SIRS模型[J].运筹与管理,2024,33(12):158-164.
- [3] 杨闯.生鲜农产品供应链金融风险评价研究[D].浙江农林大学,2024.
- [4] 赵先立,王睿儒,姜涓涓.地方政府债务、房地产风险和金融风险的动态交互影响研究[J].甘肃金融,2024,(12):46-56.
- [5] 吴弘,赵畅.习近平关于防范化解金融风险的几个重要论断[J].党的文献,2024,(06):21-30.