

# 基于大数据分析的互联网金融风险预警研究

●杨虎 易丹辉 肖宏伟

**摘要:**在国家支持互联网创新的政策背景下,互联网金融迅速发展,亟需建立互联网金融风险预警系统进行风险防范。文章从大数据出发,针对互联网金融数据的特点,遵循以数据为中心的系统设计原则,按照数据的处理方式进行系统层级划分,构建了基于大数据分析的互联网金融风险预警系统。运用大数据管理工具和分析方法,可以提前发现互联网金融中潜在的金融风险,从而为互联网企业提供风险管理依据,有利于保障互联网金融企业健康可持续发展。

**关键词:**互联网金融;大数据;金融风险预警

## 一、互联网金融风险的概述

所谓互联网金融是指借助于互联网技术、移动通信技术实现资金融通、支付和信息中介等业务的新兴金融模式,既不同于商业银行间接融资,也不同于资本市场直接融资的融资模式。随着互联网技术的快速发展,互联网金融得到迅猛的发展,具体体现在用户数量和资金规模上。央行数据显示,电子支付业务增长较快,移动支付业务保持高位增长;2013年移动支付业务16.74亿笔,金额9.64万亿元,同比分别增长212.86%和317.56%。新华网的报道显示,互联网理财产品在近一年的时间内,发展已经超出人们的想象,例如互联网理财领域的“余额宝”,截止到2014年2月,其客户数已超过6000万人,资金规模已超过2500亿元。与传统金融相比,互联网金融具有成本低、效率高、覆盖广、发展快和管理弱等特点,并在国家支持互联网创新的背景下,得到迅速的发展。

“高效共享”、“平等自由”、“信任尊重”的互联网精神推动了互联网金融的快速发展,进而形成以点对点、网格化共享互联,信息交互,资源共享,优势互补的金融体系。虽然成熟可靠的互联网技术为互联网金融的正常运营提供了强有力的保证,但是互联网金融的风险监管体系与传统金融的监管体系相比,在合法性、规范性和安全性等方面尚存在很多问题,这些问题将会产生诸多不确定的金融风险,不仅会影响企业的可持续发展,甚至会影响国家和社会的繁荣稳定。要控制金融风险,需有完善的金融预警机制作为保障。所谓金融预警机制主要是指各种反映金融风险警情、警兆、警源及变动趋势的组织形式、指标体系和预警方法等所构成的有机整体,并且以经济金融统计资料为依据,以信息技术为基础,是金融风险防范的重要组成部分。如何判断和识别金融风险,是金融风险预警机制的核心问题。随着信息技术的快速发展,结合收集的历史数据,结合数学指标、统计模型、数据挖掘等模型、算法判断和识别金融风险,是当今金融风险预警机制中的研究热点。传统的金融风险预警方法主要有三大类:景气指数法、指标体系评分法和模型法。景气指数法,通过综合许多经济因素为一个或一组景气指数来发布经济动态走向;指标

体系评分法,通过筛选指标、编制指标体系、给与指标赋分来给出金融安全状态的较为完整的评价;模型法,通过将金融危机发生的相关因素纳入统计模型进行检验来预测金融危机发生的可能性。

对于互联网企业来说,爆炸式增长的客户数据是一个亟待开发的资源,数据中所蕴藏的无限信息金矿若以先进的分析技术加以利用,将之转换为极其有价值的洞察力,能够帮助金融企业执行实时风险管理,成为金融企业的强大保护盾,保证金融企业的正常运营。数据是下一个“Intel Inside”,未来属于将数据转换成产品的公司和人们。互联网金融风险的预警体系的建立,应根植于互联网中的大数据,结合传统的金融风险分析方法,利用统计、计算机、数据挖掘、人工智能等手段,从数据的海洋中甄别、判断互联网金融中潜在的风险;并且还能通过数据掌握客户动态,企业经营环节中可能出现的金融风险,从而提高企业经营管理效益。

## 二、互联网金融中的数据及特点

1. 互联网金融的数据。与互联网电子商务一样,互联网金融作为金融信息化的形式,离不开参与互联网金融活动的企业、客户,以及相关的金融服务或产品。与传统金融活动相比,互联网金融活动更容易收集、整理、存储用户信息、用户交易数据、服务或产品信息,甚至还能存储用户在交易过程中对互联网平台的使用情况、操作行为,以及沟通、留言等信息。从互联网金融数据的构成形式来看,主要包括:用户数据、交易数据、用户操作及行为数据、金融服务或产品供给情况,以及文本数据(包括:电子邮件,即时聊天,以及留言等)。

(1) 用户数据。互联网金融业务的开展,离不开用户的参与。为保障用户在交易过程中的金融安全,保证日常金融活动的顺利进行,金融企业针对用户信息的管理工作是十分严格的。通常情况下,用户的基本信息会被收录和存储到企业信息系统之中。作为互联网金融的服务对象,用户是不可缺少的组成部分,用户规模直接反映了企业的规模,还间接反映了企业的发展前景。

(2) 交易数据。互联网金融是传统金融向电子化

方向的发展,互联网金融的主要活动离不开用户交易。互联网金融企业为用户交易提供了互联网平台媒介及相关金融服务。为保证交易安全,提高企业的服务质量,便于回溯和取证,系统会记录用户通过互联网平台交易的过程。长期积累的交易数据不仅可以用来分析用户的交易偏好,也可用来侦测用户的异常交易行为,为防止交易风险提供依据。

(3)用户操作行为数据。互联网金融平台不仅是互联网交易的媒介,也承载着传递信息,宣传金融服务的作用。与传统金融不同,互联网金融平台无法通过面对面的交谈,感知客户的感受,发现客户的异常行为。因此,为了提升互联网金融平台的服务质量,了解客户的操作行为习惯,通常会记录客户的操作行为。

(4)文本数据。作为信息传递的平台,互联网中存在大量的评价、留言、沟通交流信息,这些信息体现了民众的舆论动向。金融运行的基础为信用与预期,这种特征使其更容易受社会信用与预期舆情的影响。金融舆情能够通过一定的作用机理对互联网金融运行产生现实的影响,如果不能及时关注和应对小的金融舆情,则有可能酿成大的金融危机事件。

(5)其他数据。除此之外,还有诸多外部因素会影响互联网金融的正常运行,例如国家宏观经济运行情况、物价水平、进出口、行业发展状况等都会对互联网金融产生影响。为保证互联网金融企业的正常运行,应该全面、细致的整理和收集相关的数据。

2. 互联网金融数据的特点。与大数据一样,互联网金融数据具有规模性(Volume)、多样性(Variety)和高速性(Velocity)等三大特点。三大特点交织在一起,形成了当今中国互联网金融的新局面。

(1)规模性。所谓规模性指的是,互联网金融数据的量达到了一定的程度,无法通过当前主流的分析工具来及时处理。互联网金融数据的规模体现在用户规模增大、交易规模增大两方面。一方面,由于互联网金融的门槛较低,效率较高,互联网金融的参与者更具有广泛性、规模性,更加平民化导致互联网金融用户规模较大。另一方面,结合互联网的特点,加之互联网企业的平台、用户、以及大数据优势,互联网金融的用户规模、交易规模很容易迅速提升。

(2)多样性。所谓多样性指的是,互联网金融数据的数据类型,除了有结构化的数据以外,还有半结构化和非结构化的数据,例如文本数据。此外,还体现为互联网金融活动的多样性,互联网金融提供了在线支付、还贷借贷、理财、保险等服务,丰富了互联网金融的形式。常见的互联网金融活动包括:B2B 电商金融、B2C 电商金融、网销基金(网络理财)、网销保险、银行电商、P2P 网贷、网络支付、众筹融资、虚拟货币等。

(3)高速性。所谓高速性指的是,互联网金融数据的到达与处理必须及时高效,不允许较长的延迟,不及时将会造成不必要的损失。同时,借助互联网平台的宣传,互联网用户的响应速度提高,加之用户规模较大,数据的增长速度呈指数增加。

可见,要从大数据中识别、发现互联网金融中潜在的风险,需要有处理大数据规模性、多样性、高速性的能力。要应对互联网金融中的大数据问题,需要建立完备科学的互联网金融风险预警体系。目前大数据主要的处理模式可以分为流处理和批处理两种,其中批处理是先存储后处理,而流处理则是直接处理。不论以哪种方式处理数据,互联网金融风险预警系统都要从数据出发,识别、发现、预警、监控、预测互联网金融中潜在的风险。

### 三、互联网金融风险预警系统

#### 1. 以数据为中心的体系设计原则。

(1)系统性原则。互联网金融风险预警体系是针对互联网金融风险的监测、预测、预警的系统,是一个大的体系,必须涵盖互联网金融活动的全过程。必须考虑到互联网金融活动中的每个参与者,包括金融服务、金融产品的提供者,中介机构、用户,以及政府、监管机构等;还须考虑各种交易行为,甚至民众舆论动向。同时,还需兼顾国家宏观经济运行情况、经济指标、行业发展情况等。

(2)时效性原则。由于互联网金融数据具有高速、变化的特点,说明实时处理分析的重要性,目的就是实时防范和减少金融风险,及时识别、判断金融风险,及时对风险进行预测和响应,在时间上要连续,在内容上要连贯和可比。

(3)可操作性原则。在数据的收集、管理时,要有利于风险的识别、判断、预测;在系统的构建时,要结合符合公司实际情况,简单、可靠、易行;在数据分析过程中,选取的指标、统计方法、相关判别准则要易于分析、有利于操作,不仅能快速的识别、判断、预测风险,做出预警,还能辨别风险的源头。

(4)科学性。设计过程中应尽量考虑采用可量化的指标,同时也要设置一定的定性指标,以进一步系统地反映定量指标所不能表征的金融风险。对于定性指标也要给出准确的判断标准,尽可能避免人为因素的误导,确保评价结果的科学性、合理性和准确性。

(5)弹性原则。系统的设计应兼容既有金融风险预警系统,保证企业正常运营的前提下,随着时间的推移,对系统进行不断改进和完善。保证系统中功能、模块应能独立运行,功能各异,相互补充,避免冗余。

#### 2. 以数据为中心的系统的层级。

(1)数据管理层。数据作为系统中的核心部分,是整个体系中的关键环节。企业在建立以数据为中心的互联网金融风险预警系统过程中,必须健全为企业服务的数据管理机制,建立与企业规模相匹配的数据中心。数据中心的职责包括:数据的收集、整理、加工、存储,提供方便、可靠的数据操纵接口,以便其他层级用户的使用。数据中心管理数据时,应保证数据的完整性、准确性以及安全性;并兼顾可靠性,保证数据中心正常运营,为风险的预警提供数据支持平台。

(2)数据整合层。要从互联网金融的大数据海洋中实现金融风险的预警,必须对金融风险有透彻的定义和认识。从金融风险的定义出发,确定分析需求,对数据进行重



新整合,提取与之对应的分析数据。数据整合是保证分析结果可靠性、准确性必不可少的环节。如果说数据是预警体系的基础,那么需求则是预警体系的灵魂。数据提取层的任务包括:风险的定义、分析需求的确定、数据的整合与提取。

(3)数据分析层。数据分析是互联网金融风险管理控制的实施手段。全面的数据分析系统,应包括现行的指标体系、统计模型,及人工智能方法;同时兼顾与企业相适应的相关指标体系、统计模型等方法。数据分析层的功能应包括:风险识别、判断,风险预警,风险监控,自动上报、信号系统,风险预测,风险评级等功能。

(4)数据解释层。来自数据分析层中的每一次预警、每一个报告,都须结合企业的经营管理状况,以及企业外部经济运行环境,行业背景来进行解读。目的是更系统的评估风险,评价风险的可靠性,风险的危害程度,产生的根源,可采取的控制手段,弥补数据分析层的不足,为企业决策管理者提供更完整的决策依据,从而减少企业为规避风险所产生的损益。数据解释层应健全风险响应机制,建立风险应急小组,为及时处理风险提供依据。

结合以“数据”为中心的体系设计原则,从系统性、时效性、可操作性、科学性和弹性来看,预警体系涵盖了以数据为中心的互联网金融风险分析的各个环节,即数据的收集、数据提取、数据分析和数据解释;各层级紧紧相扣,又相互独立,为企业风险控制管理提供有力支撑;通过数据中心的建设,有利于加快企业的信息化,提供企业管理水平,降低因企业管理缺陷导致的内部风险;统筹兼顾、持续改进,降低企业管理经验成本。

#### 四、结论与机制实施建议

建立互联网金融风险预警体系的目的是,预防或降低企业在经营过程中,由于决策失误,客观情况变化或其他原因使资金、财产、信誉遭受损失。本文介绍了从互联网金融的发展状况入手,介绍了互联网金融的数据及特点,说明了互联网金融风险预警系统的设计原则和系统层级。建立以数据为中心的金融风险预警系统,不仅能够帮助企业降低和减少金融风险带来的损失,也能帮助企业提高、完善企业经营管理水平。基于大数据的金融风险预警系统作为保障互联网金融正常运行的工具,在传统金融互联网化的时代背景下,将会得到快速的发展。在系统的实施过程中,我们提出如下建议:

(1)建立科学、体系的考核评价机制。数据作为风险预警机制的核心,一旦离开操作数据的“人”,将毫无用处。因此在系统建设的过程中,应建立科学、体系的考核评价机制,提高参与者的主观能动性,保证系统顺利实施。考核机制应从数据的角度出发,以建立全面、可靠、弹性、实时、安全的数据体系为目标,对参与者在体系建设中的效能进行评估,量化参与者任务完成情况考核,奖励为体系建设做出贡献的参与者。

(2)要注意事物发展的阶段性,由易到难,逐步金融风险预警系统。互联网金融风险预警系统不是简单地借用传

统金融风险系统,或者新系统的重新开发,而是在传统金融风险系统的基础上,结合互联网的特点,建立与大数据为中心的风险预警系统,本质是传统金融行业向互联网金融行业的转变。实施的过程中,企业要做系统的评估,从简到繁,从易到难,保证企业正常运行的情况下,稳步建设以数据为中心的金融风险体系。同时,要结合企业的经营管理水平,充分利用互联网技术,在风险管理的过程中不断实践,有条件的创新,建立符合企业自身发展要求的金融风险预警系统。

(3)制定科学规范的金融风险预警系统实施、操作程序。为保证系统实施、操作的规范性,应制定科学、规范的程序。在预警系统实施的过程中,应以数据为中心,制定明确的系统实施计划,包括确定系统实施的进度、参与者、目标以及突发事件的处理等。同时要制定系统使用的行为规范、操作流程,明确参与者的权责、业务范围、数据权限等;制定风险分析、上报、反馈和监测机制,保证及时发现风险,且得到及时响应。

此外,互联网金融的正常运行离不开国家法律、法规、政策的支持,以及投资者、参与者的监督。监管部门应尽快完善法律、法规及相关政策,创造公平的竞争环境,保护互联网金融参与者的财产安全;加快相关政策的出台,明确互联网金融的业务范围,建立有效的准入和退出机制,奖励金融创新,加大金融投机行为的处罚力度;提高政府监管水平,保障互联网金融市场有序健康的发展;加强舆论监督和舆论导向,弥补政府监管不足。企业应公开披露相关数据,充分发挥互联网金融的投资者、参与者的外部监管作用,避免金融风险的发生。

#### 参考文献:

1. 陈萍.金融舆情监测分析运用及政府策略.中国国情国力 2012 (8) 26-28.
  2. 董小君.金融风险:预警机制研究.北京:经济管理出版社 2004.
  3. 胡辉.我国金融风险预警机制研究.江苏大学学位论文 2008.
  4. 孟小峰,慈祥.大数据管理:概念、技术与挑战.计算机研究与发展 2013 50(1):146-169.
  5. 李国杰.大数据研究的科学价值.中国计算机学会通讯 2012 8(9) 8-15.
  6. 刘静如.大数据:金融企业的盾与金——浅谈风险管理与消费智能.中国金融电脑 2012 (9) 83-83.
- 基金项目:中国人民大学科学研究基金(中央高校基本科研业务费专项资金资助)项目“基于高维联合模型的重复测量与生存时间资料联合评价的扩展研究”(项目号:13XNH190)。
- 作者简介:易丹辉,中国人民大学统计学院教授、博士生导师;杨虎,中国人民大学统计学院博士生;肖宏伟,中国人民大学经济学博士,国家信息中心经济预测部助理研究员。

收稿日期 2014-02-18。