

金融大数据概述

【学习目标】

- 了解大数据的概念和特征。
- 了解金融与大数据的结合方式和特点。
- 理解金融大数据在行业中的应用。
- 了解金融大数据给行业带来的影响。

【素质目标】

通过本章内容的学习,学生能够具备不断学习和更新金融大数据领域知识的能力,定期关注和学习金融大数据领域的新技术和发展动态;不断拓宽视野、保持学习的动力;保持对新技术和新方法的敏感性,并具备自主学习的意识和能力。

第一节 大数据概述

一、大数据概念的产生

近年来,随着计算技术的进步及移动互联网、物联网、5G 移动通信网络技术的发展,信息技术已经明显呈现“人—机—物”三元融合的态势,新兴应用不断出现,引发了数据规模的爆炸式增长,大数据(big data)引起了国内外产业界、学术界和政府部门的高度关注,甚至被认为是继人力、资本之后一种新的非物质生产要素,蕴含巨大价值,成为国家不可或缺的战略资源。各类基于大数据的应用正日益对全球生产、流通、分配、消费活动及经济运行机制、社会生活方式和国家治理能力产生重要影响。

在全球范围内,运用大数据推动经济发展、完善社会治理、提升政府服务和监管能力正成为趋势,有关发达国家相继制定实施大数据战略性文件,大力推动大数据发展和应用。我国互联网、移动互联网用户规模庞大,拥有丰富的数据资源和应用市场优势,大数据的部分关键技术研发取得突破,涌现出一批互联网创新企业和创新应用,一些地方政府已启动大数据相关工作。坚持创新驱动发展,加快大数据部署,深化大数据应用,已成为稳增长、促改革、调结构、惠民生和推动政府治理能力现代化的内在需要和必要选择。

二、大数据的含义

大数据引起了广泛的关注,但直到今天,并没有形成一个公认的定义,比较被人们接受的说法如下。

(1) 维基百科将大数据定义为规模庞大、结构复杂、难以通过现有商业工具和技术在容许的时间内获取、管理和处理的数据集。

(2) 美国国家标准与技术研究院(NIST)认为,大数据由具有规模巨大、种类繁多、增长速度快和变化多样,且需要一个可扩展体系结构来有效存储、处理和分析的广泛的数据集组成。

(3) 麦肯锡全球研究院综合了“现有技术无法处理”和“数据特征”两个定义,认为大数据是指大小超过经典数据库软件工具收集、存储、管理和分析能力的数据集合。这一定义是站在经典数据库处理能力的基础上看待大数据的。

(4) 高德纳(Gartner)咨询公司认为大数据是使用高效的信息处理方式以具备更强的决策力、洞察发现力和流程优化能力的海量、高增长率和多样化的信息资产。从数据的类别上看,大数据是指无法使用传统流程或工具处理或分析的信息。它定义了迫使用户采用非传统处理方法处理的超出正常处理范围及大小的数据集,其价值在于提高数据使用者的最终决策力。

综上所述,大数据是数量巨大、结构复杂、类型各异的数据集合,也是可以通过数据共享、交叉复用的知识与智力资源,更是现代社会的一种核心资产。

三、大数据的特征

大数据的特征可以用“5V”来表示:大体量(volume)、多样性(variety)、时效性(velocity)、准确性(veracity)、价值性(value)。

(一) 大体量

一般数据库的大小在TB级别,而大数据的起始计量单位在PB($1\text{PB}=1\,024\text{TB}$)级别,有的甚至跃升至EB、ZB级别,其中采集、存储和计算的量都非常大。过去,如此量级的数据是难以想象和获取的,但是随着信息技术和网络技术的发展,这一切成为可能。微型计算机、移动终端设备技术的进步提供了大量的源数据,人们的商业或非商业活动得以被记录下来;传感器等各种仪器获取数据的能力大幅提高,越来越多的事物特征可以被感知,这些特征数据会以数据的形式被存储下来;存储介质的革命、单位芯片晶体管的几何级增加使超级数据的保存成为可能;超级计算机、云计算技术为数据的转换、分析提供了便捷。互联网的普及,使数据的分享和获取越来越容易,无论用户是有意还是无意地分享或浏览网页,都会产生大量数据。

(二) 多样性

大数据具有多样性。在目前所有数据中,98%以上是数字数据,不到2%是书本等非数字数据。即使是数字数据本身,也存在结构化数据和非结构化数据的区别。结构化数据是指存储在数据库中,可以通过二维表结构实现逻辑表达的数据。简而言之,结构化数据就是

可以用相同或相似的表结构进行表达的数据;而非结构化数据则很难用二维逻辑表来表达。早期的非结构化数据主要指的是文本信息,如邮件、医疗档案、写作文档等;随着互联网和物联网的发展,网页搜索记录、社交媒体状态、图片、视频等也被纳入其中。大数据的多样性,同样体现在数据来源及数据应用过程中。例如,一个人的信用记录不仅仅源于银行,还可以源于保险公司、公安系统、就业单位等所有与生活和工作相关的地方;同样,运用个人信用记录的地方也并不单一,现实数据总是以多样化的形式存在。

(三) 时效性

大数据的时效性是指在数据量特别大的情况下,数据能够在一定的时间和范围内得到及时处理,这是大数据区别于传统数据最显著的特征。数据体量的增大对数据的处理度、时效性提出了更高的要求。例如,搜索引擎要求几分钟前的新闻能够被用户查询到,个性化推荐算法尽可能地要求实时完成推荐。而金融大数据技术正好能满足这一需求,这也是其区别于传统数据的显著特征。大数据的流动速度快,当处理的数据从 PB 增加至 TB 时,超大规模的数据快速变化,使用传统的软件工具将难以处理。只有对大数据做到实时创建、实时存储、实时处理和实时分析,才能及时有效地获得高价值的信息。

(四) 准确性

大数据的准确性是指保证处理的结果具有一定的准确性。结果的准确性涉及数据的可信度、偏差、噪声、异常等质量问题,原始数据的输入错误、缺失及数据预处理系统的失效等会导致数据的不准确,进而分析得出一些错误的结论。因此,保证正确的数据格式对大数据分析十分重要。

(五) 价值性

在大数据中,通过数据分析,可以在无序数据中建立关联进而获得大量高价值的、非显而易见的隐性知识,从而具有巨大价值。这一价值一方面体现在统计特征、事件检测、关联和假设检验等各个方面。但是,大数据的价值并不一定随数据集规模扩大而增加。若分析一个特定的问题,大数据中可能包含大量的无用数据,有价值的数据会淹没在大量的无用数据中,因而价值密度低。因此,在计算上,如何度量数据集的价值密度,如何针对应用问题快速定位有价值的数据,挖掘出有价值的数据,是大数据计算的核心问题之一。

四、大数据带来的变革

社会变革是科技创新的产物,蒸汽机的发明和电力的广泛应用将人们带进了工业社会,计算机和互联网的成熟则使人们进入了信息社会。而现在,基于 Web 2.0 的移动互连、物联网、云计算,则将开拓全新的大数据时代。大数据已经改变了我们传统的认识和对事物的方式,无论是医疗、商业、金融,还是教育、气象、军事等领域,大数据对于现代社会的改变都是全方位的。

(一) 技术等级

大数据的处理与分析正成为新兴信息技术应用的焦点,并持续推动信息产业高速增长。移动互联网、物联网、社交网络、电子商务等是新一代信息技术的应用形态,这些应用会不断地产生即时数据,成为大数据的重要来源。云计算技术则为这些海量、多样化的大数据提供

存储和运算平台,并通过分析优化,将结果反馈到应用中,使其创造出巨大的经济和社会价值。大数据价值的实现呼唤新技术、新产品、新服务、新业态的产生。这在硬件与集成设备领域表现为对芯片、存储性能提出更高的要求,并催生一体化数据存储处理服务、内存计算等市场;在软件与服务领域表现为引发了数据快速处理分析、数据挖掘技术和软件产品的发展。

（二）决策依据

在传统科学思维中,决策的制定往往是由目标驱动的。然而,大数据时代出现了另一种决策方式,即数据驱动型决策,数据成为决策制定的主要触发条件和重要依据。以天气预测为例,假如现在需要预测某天某地的天气如何,这个时候如果不掌握任何数据,就只能像抛硬币一样进行猜测。但如果知道前一天是晴天,那么预测结果的准确性就大一些。如果又知道大气云层、空气湿度、气温、风速等情况,就能更加准确地做出预测。在这个过程中,掌握的数据越多,做出的决策也就越准确。

（三）计算能力

只要拥有足够多的数据,我们就可以变得更聪明,这是对大数据时代的一个新认知。因此,在大数据时代,原本复杂的智能问题变成了简单的数据问题。只要对大数据进行简单分析,就可以达到基于复杂算法的智能计算的效果。为此,很多学者曾讨论过一个重要话题——大数据时代需要的是更多的数据还是更好的模型。

例如,机器翻译是传统自然语言技术领域的难点,虽曾提出过很多种算法,但应用效果并不理想。近年来,谷歌翻译不再仅靠复杂算法进行翻译,而是采用对它们之前收集的跨语言语料库进行简单分析的方式,提升了机器翻译的效果和效率。图灵奖的获得者吉姆·格雷(Jim Gray)曾提出科学研究的第四种范式——数据科学。在他看来,人类科学研究活动已经经历过三种范式的演变过程(早期的实验科学范式、以模型和归纳为特征的理论科学范式和以模拟仿真为特征的计算机科学范式),正在从计算机科学范式转向数据科学范式,即第四种范式。

（四）思维方式

大数据对社会的改变是显性的,推动这种改变的动力来自大数据对人们隐性的思维变革。从重要性来看,思维方式的改变甚至比社会的变革更加重要,因为思维的转变是社会变化的源泉。数据思维注重事物之间的相关关系;科学思维注重事物之间的因果关系。大数据的思维模式颠覆了千百年来人类的思维惯例,对人类的认知和与世界交流的方式提出了全新的挑战。

（五）统计方法

大数据时代对统计方法的要求也有着不同于以往的标准。

1. 统计数据采样更完整

在统计方法中,由于数据不容易获取,数据分析的主要手段是随机采样分析。随机采样的成功依赖于采样的绝对随机性,而实现绝对随机性非常困难,采样过程中出现任何偏差,都会使分析结果产生偏离。在大数据时代,数据的来源非常多,需要全面地考虑采样的范围,因此找到最优采样的标准非常困难。同时,随机采样的数据方法具有确定性,针对特定

的问题进行数据的随机采样后,一旦问题变化,采样的数据就不再可用。随机采样也受到数据变化的影响,一旦数据发生变化,就需要重新采样。随机采样的目的就是用最少的数据得到最多的信息,这取决于小数据的时代背景。小数据时代,数据的获取非常困难。大数据时代,获取全样本已经变得可行,互联网的普及使获取信息的成本非常低,计算机的发展也使处理大规模数据不再是难事。大数据不仅数据量大,而且数据全面。当有条件和方法获取到海量信息时,随机采样的方法和意义就变得不重要了。

2. 对数据精确性要求放宽

对小数据而言,由于收集的信息较少,对数据的基本要求是数据尽量精确、无错误。特别是在进行随机抽样时,少量错误将可能导致错误的无限放大,从而影响数据的准确性。同时,正是由于数据量小,才有可能保证数据的精确性。因此,数据的精确性是人们追求的目标。然而,对于大数据,保持数据的精确性几乎是不可能的。一方面,大数据通常源于不同领域产生的多个数据源,当由大数据产生所需信息时,通常会出现多源数据之间的不一致性。同时,也由于数据通过传感器、网络爬虫等形式获取,经常会产生数据丢失,而使数据不完整。虽然目前有方法和技术来进行数据清洗,试图保证数据的精确性,然而这不仅耗费巨大,而且保证所有数据都是精确的几乎是不可能的。因此,大数据无法实现精确性。另一方面,保持数据的精确性并不是必需的。经验表明,有时牺牲数据的精确性而获得更广泛来源的数据,反而可以通过数据集之间的关联提高数据分析结果的精确性。数据的丰富使得单个数据的价值密度变低,也使数据变得混杂不精确;然而,这是大数据时代必须接受的现实。

3. 数据相关性优于因果性

人们可以通过数据分析来预测某事是否会发生,其中基于因果关系分析和关联关系分析进行预测是常用的方法。然而,因果关系分析通常基于逻辑推理,耗费巨大;关联关系分析面临数据量不足的问题。在大数据时代,对于已经获取的大量数据,广泛采用的方法是使用关联关系来进行预测。经验表明,在大数据时代,由于因果关系的严格性,数据量的增加并不一定有利于得到因果关系,反而更容易得到关联关系。目前,基于关联关系分析的预测被广泛应用于各类推荐任务上。例如,在商业领域,企业只要能够从大数据分析中知道消费者喜欢哪类商品就足够了,而没有必要也没有时间去研究消费者偏好背后有什么特殊原因。

五、大数据分析的意义

在全球信息化快速发展的大背景下,大数据已成为国家重要的基础性战略资源,正引领新一轮科技创新。对网络信息空间大数据的挖掘和应用将创造巨大的商业和社会价值,并催生科学研究模式的变革,对国家经济发展和安全具有战略性、全局性和长远性意义,是重塑国家竞争优势的新机遇。充分利用我国的数据规模优势,实现数据规模、质量和应用水平同步提升,发掘和释放数据资源的潜在价值,有利于更好地发挥数据资源的战略作用。

(一) 协助企业挖掘商业价值

大数据推动社会生产要素的网络化共享、集约化整合、协作化开发和高效化利用,改变了传统的生产方式 and 经济运行机制,可显著提升经济运行水平和效率。大数据持续激发商业模式创新,不断催生新业态,已成为互联网等新兴领域促进业务创新增值、提升企业核心价值的重要驱动力。大数据产业正在成为新的经济增长点,将对未来信息产业格局产生重

要影响。

通过大数据分析,能够获取巨大的商业价值。例如,共享单车、各类网约车等城市出行领域的共享经济应用,显著地改善了供需的共享、集约化整合与协作水平,促进了资源的有效利用。而大数据在传统工业和制造业领域的应用,则有助于制造企业打通产业链,延伸产品的价值链条,并支持产品更快地升级迭代和提供更好的个性化服务。

（二）提升政府治理能力

各国政府面临一系列问题及挑战,例如环境污染问题、疾病防御与预警、资源分配、交通拥堵和养老问题等。传统的政府部门管理方式和方法已经远远不足以适应如今瞬息万变的环境,解决这些不断变化的问题并应对挑战。因此,通过互联网、通信网、有线电视网及物联网完成数据抓取、数据融合、数据分析和数据决策的任务,采用云计算的技术消除信息孤岛,把存储器、服务器存储到有保障的云存储中心去收集和存储数据。为了在政务人员中培养使用大数据的习惯,政府部门应该率先应用大数据,形成示范作用。

大数据分析的核心就是预测。智慧政府平台决策是基于大数据的挖掘技术、统计分析技术、并行化技术、数据可视化技术及云计算、人工智能等技术的应用,辅助政府捕获实时的决策信息并精准预测和准确分析现实状况,建立虚拟决策模型,寻找规律并验证决策假设,使政府不仅对决策的合理性进行宏观把握,还能做到精准预测和客观分析。智慧政府平台架构有助于提升政府服务和监管效率、降低政府决策成本,并为政务智能的研究和应用提供新的思路。总之,智慧政府平台的应用,可大幅提高政务的透明度和及时响应度,满足民意需求和诉求。

大数据应用能够揭示传统技术方式难以展现的关联关系,推动政府数据开放共享,促进社会事业数据融合和资源整合,将极大地提升政府整体数据分析能力,为有效处理复杂社会问题提供新的手段。建立用数据说话、用数据决策、用数据管理、用数据创新的管理机制,实现基于数据的科学决策,将推动政府管理理念和社会治理模式进步,加快建设与社会主义市场经济体制和中国特色社会主义事业发展相适应的法治政府、创新廉洁政府和服务型政府,逐步实现政府治理能力现代化。

（三）拓宽科学研究途径

借助对大数据分析研究,能够发现医学、物理、经济和社会等领域的新现象,揭示自然与社会中的新规律,并预测未来趋势,使基于数据的探索成为科学发现继经验、理论、计算之后的第四种范式。

在大数据爆发式增长的时代,在理论层面,对数据利用的理论体系正逐步形成,数据科学的诞生成为必然。在实践层面,随着数据科学的理论架构逐步成熟,与之相应的对数据的加工提炼及挖掘技术也伴随而生。数据科学理论可以指导人们对大数据的利用,数据加工提炼及挖掘技术则是挖掘数据的有力工具。

（四）改变人们的生活方式

通过大数据分析,企业可以了解市场行情,获得更多收入;农民可以了解明年种什么菜才能赚更多钱;打工者可以知道哪里更需要工人,哪里待遇更高,哪里能租到房子。随着大数据技术的发展,人们的生活将会发生彻底的改变。目前的数据都是在即时通信过程中产生的,包括电话、短信、微信、邮件和浏览网页等,特别是社交自媒体每天产生大量的文本、音

频及视频,也是大数据的主要来源。随着大数据技术与云计算、物联网的进一步融合,未来物联网中的数据将更多地来源于大量传感器。例如,所有的物体上都带有一个标签式的小型传感器,每隔一定时间对外发射信号。当人们去商场购物时,商场里的多个探测器会对所有商品进行扫描,人们只需刷卡即可完成购买。下班回家前,人们可以通过手机遥控的方式提前打开空调、做饭、放洗澡水。诸如此类,如果每一个物品都联网,时间、能源等将得到更有效的利用,人们就被解放出来去从事更有创造力的活动。

六、促进大数据发展的政策支持

由于大数据对经济、社会和科研的巨大价值,世界主要发达国家给予广泛关注,投入大量的人力和财力,各国也相继制定了促进大数据产业发展的政策法规。我国也充分认识到大数据时代带来的重大机遇,部署了一系列与大数据研究密切相关的科研计划。2016年3月17日,《中华人民共和国国民经济和社会发展第十三个五年规划纲要》发布,其中第二十七章实施国家大数据战略提出:把大数据作为基础性战略资源,全面实施促进大数据发展行动,加快推动数据资源共享开放和开发应用,助力产业转型升级和社会治理创新。2021年11月,工业和信息化部正式发布《“十四五”大数据产业发展规划》,要求到2025年,大数据产业测算规模突破3万亿元,年均复合增长率保持在25%左右,创新力强、附加值高、自主可控的现代化大数据产业体系基本形成。2022年12月,《中共中央 国务院关于构建数据基础制度更好发挥数据要素作用的意见》出台,旨在加快构建数据基础制度,充分发挥我国海量数据规模和丰富应用场景优势,激活数据要素潜能,做强做优做大数字经济,增强经济发展新动能,构筑国家竞争新优势。2024年1月,国家数据局等17部门联合印发《“数据要素×”三年行动计划(2024—2026年)》,旨在充分发挥数据要素乘数效应,赋能经济社会发展。

在产业方面,一批具有国际影响力的中国互联网企业已经积累了大量实际运行数据,具备了较强的研发能力。国内一些高校及科研院所也在开展与大数据相关的理论和技术研究。因此,大数据已成为关系国家经济发展、社会安全和科技进步的重要战略资源,是国际竞争的焦点和制高点。开展大数据计算的基础研究,推动大数据的技术和应用,提升我国在相关领域的自主创新能力和核心竞争力,对推动经济转型,加强社会治理,提高我国科技竞争力具有至关重要的意义。

第二节 大数据与金融的结合

互联网的应用普及使得电子银行、在线支付等通过网络提供服务的金融产品快速扩散,金融服务的质量取决于对客户行为的理解挖掘,金融业务与大数据技术逐步融合,进而催生了金融大数据。较于其他业态,金融业是信息密集型业态,现代金融企业拥有庞大的数据库,数据密度大、技术储备相对充裕,使得金融大数据创造价值的机会很大,应用也广泛。

一、金融大数据的概念

金融大数据是指集合海量非结构化数据,通过对其进行实时分析,可以为互联网金融机

构提供客户的全方位信息,通过分析和挖掘客户的交易和消费信息,掌握客户的消费习惯,并准确预测客户行为,使金融机构和金融服务平台在营销和风险控制方面有的放矢。金融大数据包含结构化和非结构化的大量数据。金融大数据的管理目标是在数据海洋中分析和挖掘出有价值的规律。金融大数据是在线的和实时的,金融大数据在运用时不追求精确,而是在混杂的现实条件下追求真实时效,在这期间一旦发现某些有用的规律,马上加以利用,从而使服务更加灵活,对市场机会更加敏感。金融行业的大数据大致分为以下几种。①传统的结构化数据,例如各种数据库和文件信息等。②以社交媒体为代表的过程数据,涵盖了用户偏好、习惯、特点、发表的评论及朋友圈之间的关系等。③日益增长的机器设备及传感器所产生的数据,例如柜面监控视频、呼叫中心、手机、自动取款机(ATM)等记录的位置信息等。

根据传统金融行业的分类,可以将金融大数据细分为大数据银行、大数据保险和大数据证券。例如,差异化车险定价是典型的大数据保险形式之一,是指保险行业利用驾驶信息来确定车险价格,具有良好驾驶习惯的车主,其车险价格就较低;反之,车险价格就较高。信用卡自动授信是典型的银行大数据应用,银行根据用卡客户数据确定是否授信并计算信用额度。机器人投资是证券公司的大数据创新模式之一,证券公司根据股价的影响因素建立模型,自动选择股票或寻找交易时机,在适当的风控模型下建立机器人投资云交易模式。

二、金融大数据的特点

(一) 数字网络扩大化

在金融大数据时代,大量的金融产品和服务通过网络呈现。例如,支付结算、网络借贷、P2P、众筹融资、资产管理、现金管理、产品销售、金融咨询等都将主要通过网络实现。网络也包括固定网络和移动网络,其中移动网络将逐步成为金融大数据服务的主要途径。从长远来看,数据化和网络化全面深入的发展将极大地改变金融行业,大数据的应用将改变传统金融机构的资金中介的职能,使其表现出虚拟化和电子化的交易特征,整个金融行业未来的发展方向将是虚拟化的,全面颠覆当前的金融服务形态。传统的资金流将逐渐体现为数据信号的交换,电子货币等数字化金融产品在经济生活中将成为主流。传统的人工服务将逐渐被移动互联网、全息仿真技术等科技手段所替代,银行通过完全虚拟的渠道更广泛地向客户提供金融服务。银行业务流程中各种单据和凭证等将逐渐由传统的纸质形式转变为数字文件的形式来处理,将极大地提高工作效率和便利性。可以合理地预见,在大数据时代,传统金融机构将在涉及管理理念和运营方式的多个方面面临挑战。未来金融机构的整体运作将是一个数据的洪流,数字金融得以全面实现。

(二) 信息渠道多样化

传统金融机构拥有两项基础功能:一是资金中介,即传统金融机构通过专有技术达到规模经济,降低资金融通的交易成本;二是信息中介,即传统金融机构采用专门的信息处理技术,解决资金借贷双方之间由于信息不对称而引发的逆向选择和道德风险问题。过去,我们主要依靠银行、券商、保险等传统金融中介机构收集信息。在大数据时代,新兴技术如社交网络、物联网、搜索引擎、移动互联网、大数据、云计算等改变了信息产生、传播、处理并运用的方式,尤其是依托以互联网与移动支付技术为基础的互联网金融降低了信息不对称和物理区域所带来的障碍,通过信息流、数据流引导各类资源的全面有效分配,甚至资金供需双

方可以直接通过网络获取信息并参与交易,促使传统的生产关系发生变革,形成了各机构主体之间关系的相对平等。这对传统金融机构的业务发展提出了巨大挑战,金融机构将改变过去自然的、被动的社会经济信息收集中心的角色,以开放的方式与客户平等交流,主动收集客户信息。大数据时代,信息的来源渠道得到极大拓展,通过互联网工具,每一位互联网用户都能成为信息源,尤其是网上购物、网上支付、网上金融产品交易的行为,更是为金融活动提供了源源不断的数据流。掌握用户大数据的企业具备了成为新的金融中介的基础条件。从信息传递和处理方面来看,大数据企业也与传统金融中介形式存在着显著的差别。

(三) 管理科学化

在未来的经济活动中,大数据将与物质资本、人力资本一起,成为生产过程中的一个重要生产要素。它可以转变成现实的生产力,并创造出巨大的经济价值。随着大数据的广泛应用,开放的、数字化的金融机构可以实现更高的生产力。信息技术的发展带来金融产品交易的虚拟化,使金融供应链对外延伸,降低了社会融资成本和财务费用,提高了整个市场的生产效率。大数据的积累使得金融机构可以通过全面分析自身的内部数据和外部的社会化数据,进而获得更为完整的客户信息,避免因客户信息不对称而导致错误认知,使得金融产品和服务销售更具精准性。此外,银行还可以通过现有客户和他们的社交网络或商业网络找到更有价值的潜在客户,并发起精准营销。随着科技的进步和数据类型的扩展,出现了越来越多的半结构化数据,利用大数据技术能够整合结构化和半结构化的交易数据、非结构化交互数据,开展全面的模式识别、分析,帮助银行实现事前风险预警、事中风险控制,建立动态的、可靠的信用系统,并识别各种交易风险,有效防范和控制金融风险,深度挖掘高价值的目标客户。金融机构可以利用大数据分析技术对海量结构化、非结构化数据进行深度分析和挖掘,更好地了解客户的消费习惯、行为特征、客户群体及个体网络行为模式,并基于这些有价值的信息,为客户制定个性化、智能化的服务模式,设计开发出更贴近用户需求的新产品。

(四) 决策精准度高

大数据的客观性和价值性将彻底重塑传统的银行决策机制,大数据时代为金融机构经营管理提供了全面、及时的决策支持信息。金融机构可以从每一个经营环节中挖掘出数据的价值,通过大数据分析更好地了解客户的行为特征、客户群体网络行为模式,优化运营流程,并进行业务创新。传统金融机构的决策模式依赖于样本数据分析和高层管理经验,而大数据时代的全面数据分析使得分析结果更具客观性和决策支持性,金融机构的决策过程将以数据为核心进行决策判断。对于金融机构的管理者来说,这是一场管理革命,能够极大地改变思维习惯。我们知道,全面数据分析是大数据的一个显著特征,在金融大数据体系下,金融机构数据获取、分析和运用的渠道机制都和传统金融运行方式截然不同。金融机构通过大数据分析技术,在对海量结构化数据和非结构化数据进行分析、判断和提取后,能够及时准确地发现业务和管理领域可能存在的机会与风险,为业务发展和风险防范提供重要决策依据。

(五) 其他

除上述特点之外,金融大数据还有其他一些特征。首先,对于单个金融企业来说,最适合扩大经营规模,由于效率提升导致其经营成本下降。金融企业的成本曲线形态也会发生

变化,长期平均成本曲线的底部会更快来临,也会更平坦、更宽。其次,基于大数据技术,金融从业人员的个体服务对象会更多,即单个金融企业从业人员会有减少的趋势,或至少其市场人员有减少的趋势。金融大数据的高效率性及扩展的服务边界,使金融服务的对象和范围也极大扩展,金融服务也更接地气。例如,在金融大数据时代,普通人都可以享受到极小金额的理财服务、存款服务、支付结算服务等,甚至极小金额的融资服务也会普遍发展起来,金融深化可以实现。

三、金融大数据的发展

在大数据时代,金融业将向着更加精细化、高效化的方向发展。

首先,金融市场各参与方有了更坚实的合作基础,证券信息自由流动,信息非对称程度大幅降低。这将使资源配置突破时间、空间和行业的限制,成本大为降低,效率进一步提高,甚至在一些社交网络上就可以形成交易市场。

其次,信用评价和征信体系更加有效,大数据所具有的预测能力将使风险管理和决策的模式由静态变为实时动态,个人的网络行为及动机也将被纳入风险定价和金融决策。

再次,非对称信息的减少及参与个体的信用能够被有效纳入定价模型,使金融市场的定价能力、范围和效率极大增强,价格信号的作用更及时有效。

最后,服务的综合化、一体化程度将加深,使金融服务的边界扩大;服务的精细化程度将无限延伸,个性化程度将极大增加。

因此,大数据时代的金融以无所不在的信息为中介,自然地融入我们的经济、生活 and 工作中。随着移动互联网和信息技术的迅速发展,金融行业的数据收集能力得到很大提高,大量连续、动态变化的数据存储成为可能。与其他行业相比,大数据决策模式对金融业更具针对性,而且金融业具备实施大数据的基本条件,所以大数据对金融业来说更具应用价值。

四、金融大数据对行业的影响

(一) 提升效率降低成本

依托金融大数据的管理模式能够降低社会金融融资成本和财务费用,提高市场效率;商业模式中大数据的积累使金融机构销售更具有精准性,金融机构能够发现更多具有价值的潜在客户,并对其展开精准营销。此外,硬件的开发和数据平台的建设可以针对不同系统中分散的、底层的交易数据进行模式识别和模式分析,有助于事前预防、事中控制;同时,大数据有利于促进金融机构创新,金融机构通过对用户的分类和信用能力分析,可以高效快捷地建立并管理由不同品种、不同数量组合而成的金融产品,还可以利用计算机语言,制定复杂的交易策略,处理海量的市场信息,捕捉短暂的市场波动,把握交易机会,提升盈利水平。

金融大数据技术还能够降低金融机构的管理和运行成本。通过金融大数据应用和分析,可以帮助金融机构准确地定位内部管理缺陷,从而制定有针对性的改进措施,实行符合自身特点的管理模式,进而降低管理运营成本。除此之外,金融大数据可以提供全新的沟通渠道和营销手段,可以更好地了解客户的消费习惯和行为特征,及时、准确地把握市场营销效果。