

信用评分基本理论及其应用

石勇^{1,2,3}, 孟凡^{1,2,3}

1. 中国科学院大数据挖掘与知识管理重点实验室, 北京 100190;
2. 中国科学院虚拟经济与数据科学研究中心, 北京 100190;
3. 中国科学院大学经济与管理学院, 北京 100190

摘要

信用评分是市场交易的基础,以大数据分析挖掘为支撑的现代信用评分系统在当今全球经济发展中起着不可估量的作用。各行业与领域的信用评分系统将成为实现我国“十三五”规划中国国家大数据战略的首要目标。主要介绍了金融信用评分系统的发展、相关技术、应用案例,并对大数据信用评分的发展进行了展望。

关键词

信用评分系统;数据挖掘;大数据评分

中图分类号:F069

文献标识码:A

doi: 10.11959/j.issn.2096-0271.2017003

Credit scoring: basic theory and applications

SHI Yong^{1, 2, 3}, MENG Fan^{1, 2, 3}

1. Key Laboratory of Big Data Mining and Knowledge Management, Chinese Academy of Sciences, Beijing 100190, China
2. Research Center on Fictitious Economy and Data Science, Chinese Academy of Sciences, Beijing 100190, China
3. School of Economy and Management, University of Chinese Academy of Sciences, Beijing 100190, China

Abstract

Credit scoring is the foundation of market transactions all over the world. Big data analysis based credit scoring plays an immeasurable role in global economics. In the 13th Five-Year Plan of China, it has been proposed as strategic developmental goal to develop and accomplish credit scoring system in all the industries and economic departments. The development of credit scoring system, the advanced techniques in credit scoring and cases of big data based on credit scoring systems were outlined. Suggestions and prospects on big data scoring systems were concluded.

Key words

credit scoring system, data mining, big data credit

1 引言

对于现代市场交易环境而言,信用(credit)是一种建立在信任基础上的能力,不用立即付款就可以获取资金、物资、服务的能力。接受信任的一方(受信方)与授予信任的一方(授信方)达成协议:在约定的时间期限内,受信方为获得的资金、物资、服务向授信方付款。信用的存在与践行,为市场经济带来了秩序,促使市场经济健康发展。在一个社会信用体系健全的国家,公正、权威的信用产品和信用服务已全面普及,信用交易已成为其市场经济的主要交易手段,企业和个人的信用意识强烈,注重维护信用,有着明确的信用市场需求;在信用体系的保障下,市场参与者更加理性,行为更加规范,市场效率得到了大幅提高。

作为信用体系的重要组成部分,信用评分系统的发展对整体信用评分体系的进步起到关键作用,随着大数据技术的快速发展和普及,基于大数据的信用评分受到了广泛的关注,并产生了一系列基于大数据的信用评分应用。首先介绍信用评分系统的发展历史,明确信用评分系统的重要性;随后介绍在信用评分系统中常用的技术;接着介绍信用评分的案例,特别是近年来出现的基于大数据的信用评分案例;最后,对基于大数据技术的信用评分的前景进行展望。

2 信用评分系统的发展

信用评分系统的出现,最早可以追溯到1801年英国布鲁林克成立的征信局,经过200多年的发展,征信行业已经相当成熟,并在社会经济活动中起着不可或缺的作用。经典的信用评分系统(credit scoring

system)指根据借款方的借款与还债历史、信用状况等信息给予不同的分数,用来决定是否核准贷款或进行信用额度的统计和评估,这一模式现在仍然广泛应用在传统金融行业的信贷管理活动中。

西方发达国家的信用评估行业处于领先地位,从历史的角度,以第二次世界大战为界将西方发达国家的信用评估行业划分为两个发展阶段。在第二次世界大战以前,包括美国在内的西方发达国家的资信评级行业发展缓慢。1910年的摩利斯计划银行率先开始提供消费信贷。1916年,一个名为Russell Sage的慈善机构制定出“小型贷款统一法”,提出了第一个针对消费贷款者的信贷框架。但由于信用的价值和风险还未被深入地研究和识别,通信设施和设备未能提供有效的支持,信用销售在当时并未能普及开来。

第二次世界大战结束后,一方面,科学技术快速发展,生产力得到极大程度的提高,促进了消费者对产品、服务的需求;另一方面,计算机技术的快速发展,为信用交易提供了坚实的基础,信用交易成为在发达国家中非常重要的促销手段,随着信用交易额的猛增,征信公司的业务规模也迅速扩大。1970年美国颁布《公平信用报告法(Fair Credit Reporting Act)》和《平等信用机会法(Equal Credit Opportunity Act)》,随着这两个法案被广泛接受,信用管理行业步入了现代信用管理阶段,并取得了快速发展。目前,美国有3家主要的信用服务公司:Equifax、Experian和Trans Union,主要的信用风险模型生产公司则为Fair Issac和First Data Corporation。

由于起步较晚,我国社会信用体系发展相对于发达国家处于落后地位。中国共产党第十八次全国代表大会以来,我国社会信用体系的建设进入全面推进阶段。党的第十八次全国代表大会提出“加强政务诚

信、商务诚信、社会诚信和司法公信建设”，党的第十八届中央委员会第三次全体会议提出“建立健全社会征信体系，褒扬诚信，惩戒失信”，中共中央国务院在《关于加强和创新社会管理的意见》中提出“建立健全社会诚信制度”以及《中华人民共和国国民经济和社会发展第十二个五年规划纲要》（以下简称“十二五”规划纲要）提出“加快社会信用体系建设”的总体要求。2014年6月，国务院发布《社会信用体系建设规划纲要（2014—2020年）》，明确了我国社会信用体系建设的指导思想、基本原则、主要目标、重点任务和保障措施。

2.1 信用的分类

根据授信人的身份，通常信用可划分为公共信用（public credit）和私人信用（private credit）两大类。其中，公共信用就是政府信用，是指社会为了帮助政府成功实现其各项职能而授予政府的信用，其核心是政府的公债。私人信用包括企业信用（business credit）和个人信用（personal credit）。企业信用也叫商业信用，个人信用也叫消费者信用。

2.2 信用评级系统的结构

信用评级系统包括信用报告和信用评级两部分。信用报告是指受信人的历史信用记录，一旦有不良记录，在7年之内无法从信用记录中抹去。它在法律允许的范围内，可以有偿地提供给银行、公司等授信方。授信方可以根据此报告来决定是否向受款人提供贷款及贷款额度。信用评级是将有关个人信用的信息转换成信用评分数的过程。它包括授信方的两部分行为：准确地预测个人有关的信用风险；产生相对应的商业决策。

在建立信用评级系统时，考虑5个方面的前提：一是评分系统考虑各种信用信息，而不是单一的因素；二是每个因素的重要性取决于所有其他因素的相关性；三是评分的信息直接来源于个人信用历史报表；四是评分同时考虑正面和负面的信息；五是评分不考虑人种、宗教、性别、婚姻状况和国籍。

基于过去大量的信用历史数据，数据分析公司运用数据挖掘技术建立可靠的评分系统。评分系统将信用卡顾客分为不同的模式（如“好”与“坏”）。在评分系统应用时，给定一个顾客作为评判对象，评分系统将根据该顾客的信用历史计算起评分。其基本原则是制造“黑名单”，尽可能地寻找将要发生坏账记录的嫌疑者。

2.3 信用评级系统的重要性

信用评级系统不仅能为授信人（银行）带来巨大的价值，同时也对受信人（顾客）和社会有着重要的作用。具体而言，对于授信人，信用评级系统可以用来确定市场和开拓新顾客，使授信人更加懂得顾客的消费行为和购买力，有利于与顾客建立良好信誉关系，促进内部金融管理质量的提高。对于受信人（顾客），信用评级系统可以使受信人能够更快地申请到应有的贷款，评分将促使其改进偿债能力，进一步提高被选为“自动信用批准”的机率。对社会而言，信用评级系统可以降低借款方申请成本，出借方信用利息下降，整个社会的贷款利息（包括住房贷款率）将下降，经济效率将提高。

3 信用评级系统的主要技术

在已掌握的数据的基础上，设计出准确、高效、有较强拓展能力的模型一直是信

用评分领域的研究热点。最早的信用评分采用的是打分卡的方式；随着数据的不断积累和应用，到20世纪90年代，判别分析、逻辑回归等方法开始被应用于信用评分领域；此后，计算机技术和数据挖掘得到迅速发展，相关进展也在信用评分领域得到了广泛的应用，尤其以支持向量机、决策树、多目标数学规划为代表的数据挖掘算法取得了很好的评分效果。本节对信用评分系统采用的主要技术进行总结和梳理。

3.1 传统信用评分技术

传统信用评分技术主要包括信用评分卡、判别分析和逻辑回归等，这些技术便于实现和推广，提高了信用评分的科学性和可靠性，在早期的信用评分应用中起到了重要作用。直到现在，这些模型仍然被广泛应用在众多现代信用评分系统中。

信用评分卡是最传统的信用评分方法，其主要实现方式为，以业务理解为基础提出描述用户行为的指标，根据各个指标具体评分细则进行评分，依据预先设定好的指标权重进行加权汇总，得到最终的信用评分。信用评分卡的发展主要经历了3个阶段^[1]：第一阶段，以对客户细分和对客户行为的刻画为主要目的构建分析方法；第二阶段，以预测用户行为为目的构建模型；第三阶段，根据业务细分和应用场景的不同，信用评分卡进一步细分，并最终形成面向决策支持的信用评分卡。总体上，根据对数据的依赖程度不同，可以分为专家型数据评分卡、数据型信用评分卡和混合型信用评分卡。在应用时应考虑具体的数据条件进行选择 and 构建。

判别分析是早期就被引入信用评分的统计模型，从统计的角度，判别分析的原则是在判断“好人”“坏人”的过程中，使第Ⅰ类错误和第Ⅱ类错误带来的总损失最

小^[2]。在信用评分问题中，第Ⅰ类错误是指将不会违约的申请者判断为违约者，这类错误将降低贷款机构的业务规模；第Ⅱ类错误是指将会违约的申请者判断为非违约者，这类错误将带来坏账率的提高。判别模型有很多种具体的方法，包括贝叶斯判别、费舍尔判别等，在信用评分领域应用最为广泛的是贝叶斯判别。贝叶斯判别对特征的独立性有着较强的假设，但这些假设在实际数据中往往难以保证，因而贝叶斯判别模型在应用中的实际效果会受到一定的限制。

逻辑回归模型最早于1970年被应用于信用评分^[3]，该模型的形式与线性回归模型相似，区别在于在线性模型的输出上，增加了一步Logit变换，将实数域上的输出映射到了[0,1]上，由此便能进行分类和概率的预测。该模型形式简洁，便于求解，而且具有很强的可解释性，因而被广泛应用在信用评分模型中。

3.2 数据挖掘技术

决策树(decision tree)作为一种经典的数据挖掘方法，以一种自动生成判断规则的方式进行建模。目前较为流行的决策树算法主要包括使用信息增益进行划分准则的ID3、使用信息增益率进行节点划分的C4.5以及以Gini指数作为节点纯度度量的分类回归树(classification and regression tree, CART)。这些算法都能应用于信用评分系统，并能产生容易理解和直接应用的判别规则。但是由于容易产生过拟合、变量增多后规则复杂、稳健性较低等问题，决策树并没有在信用评分领域得到广泛的应用。

支持向量机(support vector machine, SVM)^[4]由Vapnik提出，它是一种分类算法，通过寻求结构化风险最小来提高学习机泛化能力，实现经验风险和置信范围的最小

化,从而达到在统计样本量较少的情况下,也能获得良好统计规律的目的。具体地讲,它是一种二分类模型,其基本模型定义为特征空间上间隔最大的线性分类器,即支持向量机的学习策略便是间隔最大化,最终可转化为一个凸二次规划问题的求解。支持向量机具有良好的学习能力,并成功应用于信用评分领域^[5-7]。与此同时,各种改进的SVM模型也在信用评分领域进行了尝试,并取得了成功,其中有代表性的成果为最小二乘支持向量机^[8,9]。

多目标线性规划(multiple-criteria linear programming, MCLP)模型^[10,11]也在信用评分领域取得了一定的竞争力^[12]。该模型给定一组类别和一组评价变量,用一些相关的边界变量区分类别。最简单的多目标描述为既求分类重叠的最小化,又求类别之间距离的最大化,其结果为最满意解。在此模型框架下开发出的多目标二次规划(multiple-criteria quadratic programming, MCQP)等模型也取得了较好结果^[13]。MCLP模型和MCQP模型在中国人民银行信用评分中取得了很好的效果,并被应用于最终的评分模型中,创造了巨大的商业价值。

基于数据挖掘方法的信用评分模型,通常能够取得较为准确的预测结果,但是由于多数模型(SVM、神经网络等)的黑箱的性质,模型的可解释性不强,在进行决策支持和客户管理的时候,往往难以得到更具体细致的分析结果。这些劣势在一定程度上影响了数据挖掘技术在具体业务场景下的应用。

4 大数据信用评分案例

随着大数据时代的到来,大数据技术为信用评分带来了新的机遇。传统信用评分

系统利用的数据主要包括客户向金融机构提交的个人申请信息、金融机构内部积累的客户历史数据、中国人民银行征信中心等外部机构提供的数据。这些数据价值密度高、维度相对较低、质量相对较高。而基于大数据的信用评分会考察更加多元化的客户信息数据,主要包括各互联网金融机构掌控的生态体系内积累的客户信息以及通过外部各种渠道采集的客户信息。以互联网交易数据以及社交网络的数据为代表,互联网数据往往较为稀疏、价值密度较低、单变量区分能力较弱,但覆盖范围广泛。因而在基于大数据的信用评分系统中,需要采用先进的数据挖掘工具进行更加复杂的信用评分。

目前国内外已经出现众多成功的信用评分案例,在国外最具代表性的信用评分包括基于传统评分方法的FICO评分^①、基于手机数据源的First Access评分^②、基于丰富数据源的YODLEE^③、利用心理学分析VisualDNA^④、将非结构化数据融入模型的ZestFinance^⑤。国内互联网金融和大数据征信发展迅速,芝麻信用、腾讯信用、聚立信、51信用卡、Wecash闪银等众多公司已经基于所能获取的数据,开发出大数据信用评分模型,并成功应用于众多互联网金融的各个领域。

中国科学院虚拟经济与数据科学研究中心、中国科学院大数据挖掘与知识管理重点实验室作为国内信用评分的先驱,在近10年中与业界合作,在信用评分领域取得了多项成功的案例,具有代表性的项目包括:国内第一个全国个人信用评分决策系统——中国人民银行信用评分、1号店在线信用评分模型、考拉征信小微商户信用评分。

4.1 中国人民银行信用评分

2003—2006年,中国人民银行征信中心积累了全国各大金融机构的所有个人

① <http://www.fico.com/>
 ② <http://www.firstaccessmarket.com/>
 ③ <https://www.yodlee.com/>
 ④ <https://www.visualdna.com/>
 ⑤ <https://www.zestfinance.com/>

信贷账户的信息, 活跃自然人数超过1亿, 已经具备了研制和开发信用局风险评分的条件。中国科学院虚拟经济与数据科学研究中心与中国人民银行正式合作, 2006—2009年开展了《中国人民银行个人信用局评分项目》科研, 建立了中国评分(China score)。

该项目运用先进的数据挖掘和统计分析技术, 利用中国人民银行征信中心全国各大金融机构的消费者的住房贷款、汽车贷款、信用卡等历史信息, 通过对消费者的人口特征、信用历史记录、行为记录、交易记录等大量数据进行系统的分析, 挖掘出蕴含在数据中的行为模式, 找出历史信息与未来信用表现之间的关系, 建立预测性的模型, 预测出每个自然人在未来某个时期内发生“信贷违约”的概率, 并以一个分数来表示, 这个分数将对各信贷机构的各项信贷决策有重大的指导意义。在评分系统的建立过程中, 选取了MCLP、MCQP、逻辑回归模型、SVM、神经网络、线性回归、决策分析、判别分析等大量模型进行细致的分析和对比, 最终发现MCLP、MCQP、逻辑回归模型能够取得较好的结果, 其中MCLP和MCQP是笔者独创的基于最优化的数据挖掘方法。采用该模型得到的信用评分分布展现出和美国、澳大利亚等国家信用评分分布相似的模式。前中国人民银行副行长、国际货币基金组织副主席朱民博士对模型结果给予了高度评价, 认为“项目所开发的全国个人信用评分系统远远超过国际水平”^⑥。

项目完成后, 最终的模型被应用在中国人民银行信用评分系统中, 成为中国第一个针对全国个人的信用评分模型。从2008年起, 中国所有商业银行开始采用“中国评分”, 并将其用于贷款审批发放、信用卡申请审批等经济金融活动中, 创造了巨大的商业价值。比如, 仅在2008年, 该

评分模型使得国内7家主要商业银行大幅降低坏账风险, 共节约1 500亿元, 创造了巨大的商业价值, 并极大地推动了我国信用体系的建设和发展。

4.2 1号店在线信用评分模型

1号店作为一家规模、品类均占行业领先地位的B2C电子商务企业, 其商业链包括3部分: 供应商、商户和客户。作为国内领先的电子商务企业, 1号店积累了海量的信息数据, 这些数据不仅包括自身运营过程中产生的业务数据, 还保留了大量的网站日志数据。流量是电商存活的关键, 金融行业的竞争将紧紧围绕着客户, 尤其是优质客户这一稀缺资源。对用户的价值进行评分, 并以此作为客户关系管理的重要基础, 具有重要的商业价值。另外, 依赖所掌握的交易行为的数据, 对商户和客户的信用进行建模, 相关结果可以直接支持供应链金融以及客户的信用消费等方面。

2013—2014年, 中国科学院虚拟经济与数据科学研究中心与1号店合作, 共同开发针对商业的信用评分模型以及针对客户的信用评分和价值评分模型。根据数据情况和业务需求, 在具体建模过程中, 通过对1号店客户和供应商的海量历史数据(包括交易数据、网站行为数据等)以及在线数据进行分析和挖掘, 最终形成相关评分。项目的成果提高了1号店的资金流动性, 降低了风险, 并提高了整个商业链的运行效率。

4.3 考拉征信小微商户信用评分

考拉征信服务有限公司是一家第三方的信用评估及征信管理服务商, 其推出的考拉信用是面向政务、商务、社会、法务全方位的信用服务体系, 运用大数据及云计算技术客观呈现机构和个人的信用状况。

⑥ <http://www.chinanews.com/gn/news/2010/02-25/2137072.shtml>

2015年中国科学院大数据挖掘与知识管理重点实验室与考拉征信服务有限公司正式合作,建立联合实验室“考拉征信·中国科学院大学征信模型实验室”,开展大数据征信模型研究。在合作过程中,已经成功建立了基于大数据的小微商户信用评分模型。

中国有几千万的小微商户,其中部分小微商户由于业务发展需要,确实存在融资需求,但这些小微商户没有资产可以抵押,按照传统信贷思路很难获得贷款。为解决这一问题,征信模型实验室开发了小微商户信用评分模型,通过采集这些小微商户的基本属性、每日经营交易流水数据、工商信息以及外部的互联网公开信息,利用机器学习算法,从商户属性、信用记录、履约能力、成长能力、经营稳定、交易行为等维度评估小微商户信用,帮助他们凭借信用申请到贷款,缓解小微企业融资难问题。模型采用的数据整合了多个数据源,如金融、电信等行业的消费数据,政府的公开数据及互联网大数据等,其中包括:银联数据、拉卡拉数据、运营商数据、工商数据、公安身份信息验证、法院老赖名单、工商处罚名单、网络爬虫商户评价、考拉个人征信评分等,从多个场景还原、描述及反映了商户的信用行为和信用状况。

小微商户评分的应用场景主要包括:将考拉商户信用分嵌入“考拉征信应用”,小微商户可以通过考拉征信应用查询信用评分,方便用户了解自己的信用,并且线上就可以申请贷款让信用创造价值;与光大银行合作推出“信盈联名信用卡”,作为大数据征信平台下针对小微商户推出的专属金融服务,根据小微商户的信用分,创新性地推出以“考拉商户信用分”换取“个人信用额度”的金融模式,为小微商户提供信贷支持;推出考拉征信的“信用借”产品,基于信用评分的综合金融服务平台,提供多款借贷产品、信用卡等服务,为小微商户、个人用户和信用卡机构、贷款公司之间搭起一座信

用桥梁。目前,光大银行、拉卡拉、卡得万利等多家银行及小贷机构已将考拉商户信用分应用在贷款审核中,其中光大银行累积授信额度已过亿元。

5 结束语

经过多年的努力与发展,信用评分模型和系统的科研与应用已经取得了巨大的进展,为信贷行业以及社会经济发展创造了巨大的价值。随着大数据时代的到来,数据的大体量、来源的多样性和数据异构性为传统的信用评分技术带来了挑战。与此同时,描述用户特征的角度越来越多,为更精确地刻画用户行为提供了更丰富的数据基础。尽管国内外众多信用评分机构都已经在大数据信用评分领域取得了一定的进展,但是在实践中采用的模型和学习方法还未能充分利用包括深度学习在内的最先进的大数据挖掘技术。此外,各评分结构掌握和利用的数据仍然只是大数据中的一小部分,许多数据源,特别是分散在网络空间的社会数据没有得到充分应用。如何促进各评分结构间的数据共享和交换,将成为大数据信用评分取得突破性进展的重要条件,这也将成为信用评分系统进一步发展的方向。

参考文献:

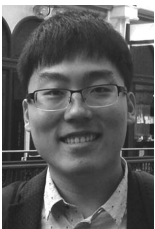
- [1] 李延东,郑小娟. 信用评分卡体系的发展及应用[J]. 甘肃金融, 2016(3): 43-45.
LI Y D, ZHENG X J. Credit scoring card system: advances and applications [J]. Gansu Finance, 2016(3): 43-45.
- [2] 朱晓明, 刘治国. 信用评分模型综述[J]. 统计与决策, 2007(2): 103-105.
ZHU X M, LIU Z G. Reviews on credit scoring models[J]. Statistics and Decision, 2007(2): 103-105.

- [3] ORGLER Y E. A credit scoring model for commercial loans [J]. *Journal of Money, Credit and Banking*, 1970, 2(4): 435-445.
- [4] VAPNIK V N. *Statistical learning theory* [M]. New York: Wiley, 1998.
- [5] HUANG C L, CHEN M C, WANG C J. Credit scoring with a data mining approach based on support vector machines[J]. *Expert Systems with Applications*, 2007, 33(4): 847-856.
- [6] MARTENS D, BAESENS B, GESTEL T V, et al. Comprehensible credit scoring models using rule extraction from support vector machines [J]. *European Journal of Operational Research*, 2007, 183(3): 1466-1476.
- [7] BELLOTTI T, CROOK J. Support vector machines for credit scoring and discovery of significant features[J]. *Expert Systems with Applications*, 2009, 36(2): 3302-3308.
- [8] ZHOU L, LAI K K, YU L. Least squares support vector machines ensemble models for credit scoring [J]. *Expert Systems with Applications*, 2010, 37(1): 127-133.
- [9] LAI K K, YU L, ZHOU L, et al. Credit risk evaluation with least square support vector machine[C]//*International Conference on Rough Sets and Knowledge Technology*, July 24-26, Chongqing, China. Berlin Heidelberg: Springer, 2006: 490-495.
- [10] KOU G, LIU X, PENG Y, et al. Multiple criteria linear programming approach to data mining: models, algorithm designs and software development[J]. *Optimization Methods and Software*, 2003, 18(4): 453-473.
- [11] SHI Y, TIAN Y, KOU G, et al. Multiple criteria linear programming[M]//*Optimization based data mining: theory and applications*. London: Springer London, 2011: 119-132.
- [12] SHI Y, PENG Y, XU W, et al. Data mining via multiple criteria linear programming: applications in credit card portfolio management [J]. *International Journal of Information Technology & Decision Making*, 2002, 1(1): 131-151.
- [13] SHI Y, PENG Y, KOU G, et al. Classifying credit card accounts for business intelligence and decision making: a multiple-criteria quadratic programming approach [J]. *International Journal of Information Technology & Decision Making*, 2005, 4(4): 581-599.

作者简介



石勇 (1956-), 男, 中国科学院大学经济管理学院教授、博士生导师, 发展中国家科学院院士, 国务院参事, 中国科学院虚拟经济与数据科学研究中心主任, 中国科学院大数据挖掘与知识管理重点实验室主任。被列为 Elsevier 发布 2014/2015 中国高被引学者榜单之一、入选汤普生路透 (Thomson Reuters) 2016 全球高被引作者 (计算机科学领域)。现在担任《国际信息技术和决策杂志 (IJITDM)》《国际数据科学年鉴 (Annals of Data Science)》创立主编。先后荣获以集合论创始人命名的康拓学术奖、中国复旦管理学杰出贡献奖、教育部自然科学奖一等奖 (2012 年、2016 年)。中国科学院百人计划学者、国家杰出青年科学基金获得者。



孟凡 (1989-), 男, 中国科学院大学经济与管理学院、中国科学院大数据挖掘与知识管理重点实验室博士生, 主要研究方向为大数据信用评分、量化投资、大数据图像挖掘、弱标签学习问题等。

收稿日期: 2016-12-22

基金项目: 国家自然科学基金资助项目 (No.91546201, No.71331005, No. 71110107026)

Foundation Items: The National Natural Science Foundation of China (No.91546201, No.71331005, No. 71110107026)