

P2P网络借贷个人信用风险评估

宋丽平(教授), 张利坤, 徐 玮

(哈尔滨理工大学经济学院, 哈尔滨 150040)

【摘要】 P2P网络借贷是由互联网和P2P借贷相结合的一种新型金融服务模式。个人信用风险是指P2P网络借贷平台的借款人无法偿还贷款的风险,这是P2P网络借贷平台面临的主要风险。借款人自身的客观条件、还款能力、历史表现都会对P2P网络借贷个人信用风险产生重要影响。本文针对P2P网络借贷平台的特点,确定个人信用风险评估指标,并以平台借款人个人信用等级作为预测输出目标,创建BP神经网络模型,使贷款人和网贷平台能够更好地了解借款人的信用状况。

【关键词】 P2P网络借贷; 个人信用风险; BP神经网络模型

一、引言

P2P网络借贷起源于P2P借贷,是指贷款人与借款人之间通过网络借贷平台而不是金融机构进行贷款交易,是互联网和P2P借贷相结合的一种新型金融服务模式。P2P网络借贷行业发展迅速,平台数量和交易额发展呈快速增长的趋势,国内P2P平台从2009年的9家增长到2014年约2 358家,贷款余额约1 386.72亿元。行业快速发展,风险也随之暴露,其中借款人的违约给P2P网络借贷平台和贷款人造成了严重的损失,个人信用风险已经成为P2P网络借贷平台面临的主要风险。目前大多数情况下贷款人无法了解借款人的信用状况,也无法在贷前对借款人的信用状况做出客观的评价。鉴于此,本文尝试构建一个评估P2P网络借贷个人信用风险的模型,使贷款人和网贷平台能够更好地了解借款人的信用状况,降低损失发生的可能性。

二、文献回顾

国内外已有不少有关P2P网络借贷信用风险的研究。Freedman(2008)指出,为了成功贷到自己所需的金额,借款人通常不会提供对自己不利的信息,甚至会利用虚假信息进行贷款,致使贷款人将贷款投向风险高的借款人,使贷款人遭受损失的可能性增加。王嵩青等(2014)认为,P2P网络借贷平台无法借助金融信息基础数据库来降低信息不对称引起的信用风险,同时平台的高利率加重了借款成本,可能使借款人无法按期还本付息,加剧了借款人的违约风险。在信用风险的影响因素方面,Emekter等(2015)指出信用等级在评估P2P网络贷款违约风险中发挥着重要作用,信用等级越低的借款人发生违约风险的概率越高。肖曼君等(2015)从借款信息、历史表现以及个人特征等方面对P2P网络借贷信用风险的影响因素进行

了研究,结果表明借款人的性别、年龄、认证等级等都会对其信用风险产生重要影响。

越来越多的学者在进行风险评估的研究时倾向于选择利用BP神经网络来构建相关模型。BP神经网络是一种基于误差反向传播算法(BP算法)的多层前向神经网络,可以实现任意线性或非线性的函数映射,包括输入层、隐含层、输出层,能够发现存在于未知模式的大量复杂数据中的规律,在处理非线性系统方面比传统的评估模型具有明显的优势。陈朝晖等(2011)选择BP神经网络构建模型进行商业银行信用风险预警分析,并达到了88%的准确率;许美玲等(2014)根据村镇银行信用风险的影响因素选取了12项风险指标,并将贷款质量作为输出值,构建了BP神经网络风险预警模型,模型的输出结果与样本数据非常接近,证明了该模型的合理性。

通过以上文献回顾可知,P2P网络借贷平台面临的个人信用风险是指借款人由于自身原因无法按期偿还贷款造成违约的可能性,主要受借款人自身条件、还款能力以及历史表现等非系统性因素影响,P2P网络借贷信用风险多为非系统性风险。鉴于BP神经网络可以处理任意类型的数据及具有较高的准确率,本文拟选择利用其构建P2P网络借贷个人信用风险评估模型。

三、P2P网络借贷个人信用风险的影响因素

根据P2P网络借贷个人信用风险的定义,其影响因素主要来自于借款者本身,包括借款者自身的客观条件、还款能力以及历史表现等。

1. 借款人自身的客观条件对信用风险的影响。随着年龄的不断变化,借款人的违约率也会产生变化。一般而言,30~45岁的借款人经济收入较稳定,其发生违约的风险概率较低;年龄较小的借款人储蓄意识较薄弱,难以养

成良好的消费习惯,其发生违约的风险概率较高;而年龄较大的借款人收入水平较低,突发消费较多,其违约风险也相对较高。相对于女性来说,男性通常具有更大的社会经济压力,其经济状况的变数较大,发生违约的可能性较大。借款人婚姻状况的改变会导致家庭收入结构的变化,相应的还款能力也会发生改变。已婚的借款人经济情况相对于未婚的借款人来说更加稳定,信用相对较高。一般而言随着文化程度的提高,个人素质、工作能力和专业知识水平也会更高,所以文化程度较高的借款人较文化程度低的人有更强的还款能力,发生道德风险的几率更低。

2. 借款人还款能力对信用风险的影响。借款人的还款能力越强,违约风险越低。借款人的工作时间越长,对自身行业或工作的熟悉程度越高,该指标有助于判断其收入稳定性和经营能力,从而判断其还款能力。借款人是否有房或者车,直接反映其经济实力,从侧面反映其还款能力。借款人月收入是借款人的主要还款来源,也是其实际的现金流,能够直接反映其还款能力。

3. 借款人的历史表现对信用风险的影响。借款人在此次借款之前成功借到资金的次数越多,表示其信用等级越高,发生信用风险的概率越低。借款人在借款的时候会选定还款方式,其中不能按期还款的次数称为逾期笔数,逾期笔数多代表着借款人的信用状况较差,发生违约的可能性较大。

四、P2P网络借贷平台个人信用风险评估模型的构建

1. 样本选取。本文从我国成立较早且经营模式是无抵押无担保的人人贷平台交易数据中选取119组借款人信息作为本次研究的样本,包含29个违约借款人的信息。样本选取时间截至2015年3月24日。将从P2P网络借贷平台选取的119条借款人信息进行指标梳理,并根据指标取值表对借款人信息进行了量化,方便进行模型数据输入。P2P网络借贷平台通常将借款人的信用从高到低划分为七个等级:AA、A、B、C、D、E、HR,本文对其取值依次为1、0.9、0.8、0.7、0.6、0.5、0.4,并将其作为模型的期望输出值。

2. 指标选取。P2P网络借贷平台的特点之一就是借款

人提供的个人信息较为简单,一般包括个人身份信息、基本财产状况、年龄、学历等基础情况,然后通过第三方认证平台对借款人提供的信息进行认证。根据P2P网络借贷平台的特点,结合P2P网络借贷平台个人信用风险的影响因素,可将借款人的基本资料归整为年龄、性别、婚姻状况、文化程度、工作年限、月收入范围、房产状况、是否购车、借款成功次数、逾期笔数,并将其选定为风险评估指标。

本文参考商业银行个人信用评级的取值标准,结合P2P网络借贷平台的自身特点,对每一个指标的具体取值如表1所示。

表1 借款人指标取值情况

指标	取值							
	年龄	20~25岁	26~30岁	31~35岁	36~40岁	41~45岁	46~50岁	51~55岁
	0.7	0.8	0.9	1	0.9	0.8	0.7	0.6
性别	男				女			
	0				1			
婚姻状况	未婚				已婚			
	0				1			
文化程度	高中或以下		大专		本科		硕士及以上	
	0.3		0.5		0.7		1	
工作年限	1年(含)以下		1~3年		3~5年		5年(含)以上	
	0.4		0.6		0.8		1	
月收入范围	1 001~2 000元	2 000~5 000元	5 000~10 000元	10 000~20 000元	20 000~50 000元	50 000元以上		
	0.5	0.6	0.7	0.8	0.9	1		
房产状况	无				有			
	0				1			
是否购车	无				有			
	0				1			
借款成功次数	10(含)以下		11~20		21~30		30以上	
	0.7		0.8		0.9		1	
逾期笔数	0	1~5	6~10	11~15	16~20	20以上		
	1	0.8	0.6	0.4	0.2	1		

3. BP神经网络模型的构建。在将BP神经网络应用于P2P网络借贷个人信用风险的评估时,模型的输入应为评估指标,本文拟选定10个指标对P2P网络借贷平台个人信用风险进行评估,因此输入层节点数设定为10;模型以借款人的信用等级作为输出值,输出层节点数确定为1,输出值为模型实际输出值。根据最常用的公式设置隐含层个数:

$$k = \sqrt{n + m} + a$$

其中:k为隐含层节点数;n为输入层节点数;m为输出层节点数;a为常数,a∈[1,10]。隐含层的节点数需要在不断训练的过程中根据结果确定最优数。其他参数设置如下:目标误差值设为0.000 01,最大训练迭代次数设为

1 000次,其他采用MATLAB7.0软件工具箱默认参数。

4. 模型训练过程和结果。在训练过程中,本文选取119条借款人数据中的前109条作为训练数据,其中含有26个有违约记录的借款人。调用MATLAB7.0软件BP神经网络工具箱训练样本,模型样本在训练过程中不断优化调整,最终确定隐含层节点数为10,因此将BP模型结构优化为(10,10,1)。

运行MATLAB7.0,通过29次迭代,误差达到预期目标,由此可见,BP神经网络模型训练时间较短,模型较优。训练结果如图1所示。

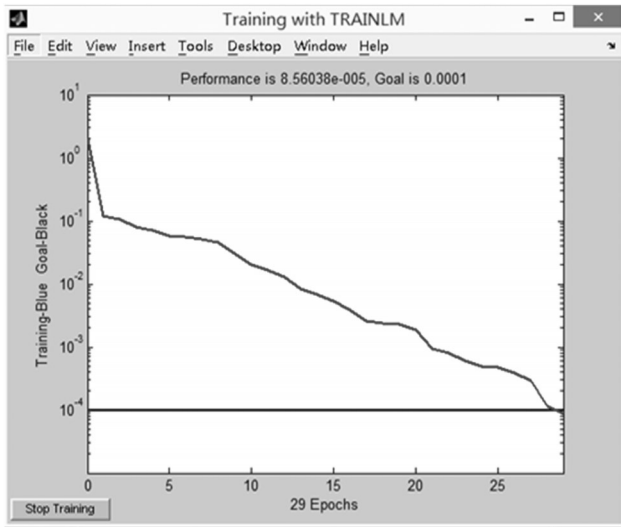


图1 训练结果

BP神经网络模型在训练过程中的误差大概服从正态分布,达到了预期目标,情况良好;109条样本期望输出值和实际输出值的分布情况如图2所示,其中黑点表示网络的输出结果,“△”表示期望输出值,拟合系数较高,拟合情况良好。

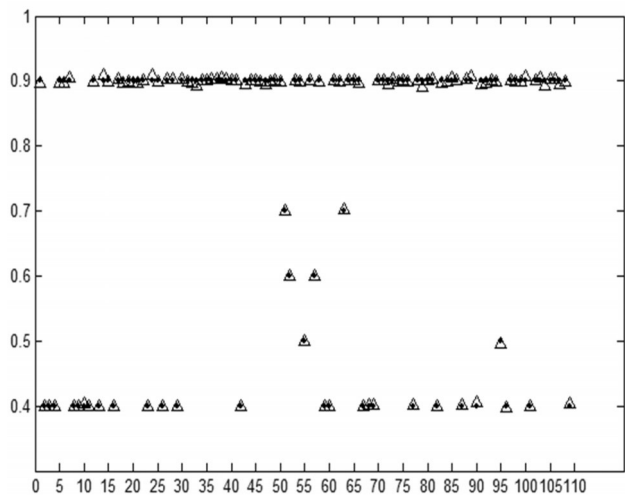


图2 函数逼近结果

5. 模型仿真过程和结果。训练及拟合完成后,对经过训练生成的神经网络模型的预测结果进行验证,将第110~119条数据作为验证数据,其中有违约记录的借款人样本为3个,无违约记录的借款人样本为7个。输入第110~119条验证数据后,期望输出值和模型输出结果对比如表2所示。

表2 期望输出值与模型输出结果对比

编号	110	111	112	113	114
期望输出值	0.4	0.4	0.9	0.4	0.9
模型输出值	0.401 71	0.398 66	0.897 11	0.403 25	0.907 43
信用等级	HR	HR	A	HR	A
编号	115	116	117	118	119
期望输出值	0.9	0.9	0.9	0.9	0.9
模型输出值	0.901 34	0.898 04	0.907 93	0.908 05	0.865 81
信用等级	A	A	A	A	A

由表2可以看出,模型根据109条借款人信息数据进行训练和学习后,对仿真数据预测的输出结果和期望输出结果基本一致。根据BP神经算法建立了P2P网络借贷平台个人信用风险评估的神经网络模型,通过对输入的借款人数据的训练和学习,不断调整连接各层的权值和阈值,最终确定输入和输出之间的内在联系,进而使构建的模型具有了对P2P网络借贷个人信用风险的评估预测能力,且准确率较高。

五、结论

本文在分析P2P网络借贷个人信用风险影响因素的基础上,根据P2P网络借贷的特点,将借款人年龄、性别、婚姻状况、文化程度、工作年限、月收入范围、房产状况和是否购车、借款成功次数、逾期笔数等选为评估指标并取值,构建了基于BP算法的神经网络模型,对其信用等级进行预测。本文选取样本数据以及验证数据的数量不够大,可能会导致实验结果有向好的偏差,但训练仿真的结果表明,该模型适用于对P2P网络借贷个人信用状况的预测,可以作为贷款人在贷前对借款人信用状况做出客观评价的简便方法。

主要参考文献

- 王嵩青,田芸,沈霞.征信视角下P2P网贷模式的信用风险探析[J].征信,2014(12).
- Emekter R., Tu Y., Jirasakuldech B., Lu M.. Evaluating credit risk and loan performance in online Peer-to-Peer (P2P) lending[J]. Applied Economics, 2015(1).
- 肖曼君,欧缘媛,李颖.我国P2P网络借贷信用风险影响因素研究——基于排序选择模型的实证分析[J].财经理论与实践,2015(1).
- 江训艳.基于BP神经网络的商业银行信用风险预警研究[J].财经问题研究,2014(S1).