



# (12)发明专利申请

(10)申请公布号 CN 109146122 A

(43)申请公布日 2019.01.04

(21)申请号 201810676342.3

(22)申请日 2018.06.27

(71)申请人 深圳市买买提信息科技有限公司  
地址 518000 广东省深圳市前海深港合作区前湾一路1号A栋201室(入驻深圳市前海商务秘书有限公司)

(72)发明人 曾吟 任虎

(74)专利代理机构 广州三环专利商标代理有限公司 44202  
代理人 郝传鑫 熊永强

(51)Int.Cl.  
G06Q 10/04(2012.01)  
G06Q 40/02(2012.01)  
G06F 16/2458(2019.01)

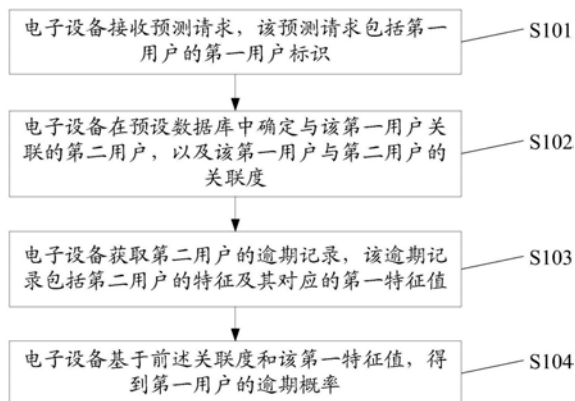
权利要求书2页 说明书9页 附图3页

## (54)发明名称

一种概率预测方法、装置、电子设备及计算机存储介质

## (57)摘要

本发明实施例公开了一种概率预测方法、装置、电子设备及计算机存储介质,其中,方法包括:通过在预设数据库中查找与第一用户关联的第二用户,确定第一用户与第二用户的关联度,并获取第二用户的逾期记录;进一步的,根据第二用户的逾期记录,得到第二用户的特征及其对应的第一特征值;然后根据第一用户与第二用户的关联度以及该第一特征值,得到第一用户的逾期概率。通过实施本发明实施例,可以根据第二用户的逾期记录来预测第一用户的逾期概率。



1. 一种概率预测方法,其特征在于,包括:  
接收预测请求,所述预测请求包括第一用户的第一用户标识;  
在预设数据库中确定与所述第一用户关联的第二用户,以及所述第一用户与所述第二用户的关联度;  
获取所述第二用户的逾期记录,所述逾期记录包括所述第二用户的特征及其对应的第一特征值;  
基于所述关联度和所述第一特征值,得到所述第一用户的逾期概率。
2. 根据权利要求1所述的方法,其特征在于,所述预测请求还包括所述第一用户的身份信息及其对应的第二特征值;  
所述基于所述关联度和所述第一特征值,得到所述第一用户的逾期概率,包括:  
基于所述关联度、所述第一特征值和所述第二特征值,得到所述第一用户的逾期概率。
3. 根据权利要求1或2所述的方法,其特征在于,所述在预设数据库中确定与所述第一用户关联的第二用户,包括:  
在所述预设数据库中获取包含所述第一用户标识的即时通讯信息和/或逾期信息;  
根据所述包含所述第一用户标识的即时通讯信息和/或逾期信息,确定与所述第一用户关联的第二用户。
4. 根据权利要求3所述的方法,其特征在于,所述在预设数据库中确定所述第一用户与所述第二用户的关联度,包括:  
根据所述包含所述第一用户标识的即时通讯信息和/或逾期信息,确定所述第一用户与所述第二用户的关联度。
5. 根据权利要求1所述的方法,其特征在于,所述第二用户的数量为多个;  
所述基于所述关联度和所述第一特征值,得到所述第一用户的逾期概率,包括:  
基于各个所述第二用户与所述第一用户的关联度,对各个所述第二用户的第一特征值进行加权求和,得到所述第一用户的逾期概率。
6. 一种概率预测装置,其特征在于,包括:  
接收单元,用于接收预测请求,所述预测请求包括第一用户的第一用户标识;  
处理单元,用于在预设数据库中确定与所述第一用户关联的第二用户,以及所述第一用户与所述第二用户的关联度;  
所述处理单元,还用于获取所述第二用户的逾期记录,所述逾期记录包括所述第二用户的特征及其对应的第一特征值;  
所述处理单元,还用于基于所述关联度和所述第一特征值,得到所述第一用户的逾期概率。
7. 根据权利要求6所述的装置,其特征在于,所述预测请求还包括所述第一用户的身份信息及其对应的第二特征值;  
所述处理单元,具体用于基于所述关联度、所述第一特征值和所述第二特征值,得到所述第一用户的逾期概率。
8. 根据权利要求6或7所述的装置,其特征在于,  
所述处理单元,具体用于在所述预设数据库中获取包含所述第一用户标识的即时通讯信息和/或逾期信息;根据所述包含所述第一用户标识的即时通讯信息和/或逾期信息,确

定与所述第一用户关联的第二用户。

9. 一种电子设备,其特征在于,包括处理器和存储器,所述存储器用于存储计算机程序,所述计算机程序包括程序指令,所述处理器被配置用于调用所述程序指令,执行如权利要求1~5任一项所述的方法。

10. 一种计算机存储介质,其特征在于,所述计算机存储介质存储有计算机程序,所述计算机程序包括程序指令,所述程序指令当被处理器执行时使所述处理器执行如权利要求1~5任一项所述的方法。

## 一种概率预测方法、装置、电子设备及计算机存储介质

### 技术领域

[0001] 本发明涉及机器学习领域,尤其涉及一种概率预测方法、装置、电子设备及计算机存储介质。

### 背景技术

[0002] 风险控制(简称风控)是指风险管理者采取各种措施和方法,消灭或减少风险事件发生的各种可能性,或者减少风险事件发生时造成的损失。

[0003] 现有的风控模型只能通过老用户以往的逾期情况,对老用户的逾期概率进行预测,而不能对新用户进行逾期预测。因此,如何预测新用户的逾期概率成为亟待解决的技术问题。

### 发明内容

[0004] 本发明实施例公开了一种概率预测方法、装置、电子设备及计算机存储介质,可以预测第一用户(即新用户)的逾期概率。

[0005] 第一方面,本发明实施例公开了一种概率预测方法,该方法可以包括:接收预测请求,所述预测请求包括第一用户的第一用户标识;在预设数据库中确定与所述第一用户关联的第二用户,以及所述第一用户与所述第二用户的关联度;获取所述第二用户的逾期记录,所述逾期记录包括所述第二用户的特征及其对应的第一特征值;基于所述关联度和所述第一特征值,得到所述第一用户的逾期概率。

[0006] 在一种实现方式中,所述预测请求还包括所述第一用户的身份信息及其对应的第二特征值;所述基于所述关联度和所述第一特征值,得到所述第一用户的逾期概率的具体实施方式可以为:基于所述关联度、所述第一特征值和所述第二特征值,得到所述第一用户的逾期概率。

[0007] 在一种实现方式中,所述在预设数据库中确定与所述第一用户关联的第二用户的具体实施方式可以为:在所述预设数据库中获取包含所述第一用户标识的即时通讯信息和/或逾期信息;根据所述包含所述第一用户标识的即时通讯信息和/或逾期信息,确定与所述第一用户关联的第二用户。

[0008] 在一种实现方式中,所述在预设数据库中确定所述第一用户与所述第二用户的关联度的具体实施方式可以为:根据所述包含所述第一用户标识的即时通讯信息和/或逾期信息,确定所述第一用户与所述第二用户的关联度。

[0009] 在一种实现方式中,所述第二用户的数量可以为多个;所述基于所述关联度和所述第一特征值,得到所述第一用户的逾期概率的具体实施方式可以为:基于各个所述第二用户与所述第一用户的关联度,对各个所述第二用户的第一特征值进行加权求和,得到所述第一用户的逾期概率。

[0010] 第二方面,本发明实施例公开了一种概率预测装置,该装置可以包括:接收单元,用于接收预测请求,所述预测请求包括第一用户的第一用户标识;处理单元,用于在预设数

数据库中确定与所述第一用户关联的第二用户,以及所述第一用户与所述第二用户的关联度;所述处理单元,还用于获取所述第二用户的逾期记录,所述逾期记录包括所述第二用户的特征及其对应的第一特征值;所述处理单元,还用于基于所述关联度和所述第一特征值,得到所述第一用户的逾期概率。

[0011] 在一种实现方式中,所述预测请求还包括所述第一用户的身份信息及其对应的第二特征值;所述处理单元具体可以用于基于所述关联度、所述第一特征值和所述第二特征值,得到所述第一用户的逾期概率。

[0012] 在一种实现方式中,所述处理单元具体可以用于在所述预设数据库中获取包含所述第一用户标识的即时通讯信息和/或逾期信息;根据所述包含所述第一用户标识的即时通讯信息和/或逾期信息,确定与所述第一用户关联的第二用户。

[0013] 在一种实现方式中,所述处理单元具体可以用于根据所述包含所述第一用户标识的即时通讯信息和/或逾期信息,确定所述第一用户与所述第二用户的关联度。

[0014] 在一种实现方式中,所述第二用户的数量可以为多个;所述处理单元具体可以用于基于各个所述第二用户与所述第一用户的关联度,对各个所述第二用户的第一特征值进行加权求和,得到所述第一用户的逾期概率。

[0015] 第三方面,本发明实施例公开了一种电子设备,该包括存储器和处理器,所述存储器用于存储计算机程序,所述计算机程序包括程序指令,所述处理器被配置用于调用所述程序指令,执行上述第一方面所述的方法。

[0016] 第四方面,本发明实施例公开了一种计算机存储介质,所述计算机存储介质存储有计算机程序,所述计算机程序包括程序指令,所述程序指令当被处理器执行时使所述处理器执行上述第一方面所述的方法。

[0017] 通过实施本发明实施例,可以根据与第一用户关联的第二用户(即老用户)的逾期记录,以及第一用户与第二用户的关联度,预测第一用户(即新用户)的逾期概率。

## 附图说明

[0018] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0019] 图1是本发明实施例提供的一种概率预测方法的流程示意图;

[0020] 图2是本发明实施例提供的另一种概率预测方法的流程示意图;

[0021] 图3是本发明实施例提供的一种知识图谱的示意图;

[0022] 图4是本发明实施例提供的一种概率预测装置的结构示意图;

[0023] 图5是本发明实施例提供的一种电子设备的结构示意图。

## 具体实施方式

[0024] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他

实施例,都属于本发明保护的范围。

[0025] 本申请的技术方案的主要原理可以包括:通过在预设数据库中查找与第一用户关联的第二用户,确定第一用户与第二用户的关联度,并获取第二用户的逾期记录;进一步的,根据第二用户的逾期记录,得到第二用户的特征及其对应的第一特征值;然后根据第一用户与第二用户的关联度以及该第一特征值,得到第一用户的逾期概率。即本申请可以根据其他用户(即第二用户)的逾期记录来预测第一用户的逾期概率。

[0026] 请参阅图1,图1是本发明实施例提供的一种概率预测方法的流程示意图。具体的,如图1所示,本发明实施例的概率预测方法可以包括但不限于如下步骤:

[0027] S101、电子设备接收预测请求,该预测请求包括第一用户的第一用户标识。

[0028] 具体的,电子设备可以在接收到预测请求的情况下,提取该预测请求包括的第一用户的第一用户标识。其中,第一用户标识用于唯一标识第一用户,例如,第一用户标识可以是身份证号或者昵称。需要说明的是,第一用户为新用户,也就是说,电子设备中不存在该第一用户的逾期记录。例如,当张三在贷款平台A上提交贷款请求时,若张三在此之前并未在贷款平台A上有过贷款记录,所以无法在贷款平台A中查询到张三的逾期记录,此时,在贷款平台A中,张三为新用户。

[0029] 在一种实现方式中,电子设备可以是终端设备,也可以是服务器。其中,终端设备可以是智能手机、平板电脑、个人计算机(Personal Computer,PC)、智能电视、智能手表、车载设备、可穿戴设备、未来第五代移动通信技术(the 5th Generation,5G)网络中的终端设备等。在一种实现方式中,当电子设备为服务器时,该预测请求可以是终端设备发送的。当电子设备为终端设备时,该预测请求可以是根据用户的操作自动生成的,本发明实施例对此不作限定。

[0030] S102、电子设备在预设数据库中确定与该第一用户关联的第二用户,以及该第一用户与第二用户的关联度。

[0031] 具体的,电子设备可以在预设数据库中查找与第一用户标识关联的第二用户标识,将第二用户标识对应的用户确定为第二用户,并且,将第一用户标识与第二用户标识之间的关联度确定为第一用户与第二用户的关联度。其中,关联度可以是大于0的数字,关联度可以用于表示第一用户与第二用户之间的亲密程度,关联度越大,表示第一用户与第二用户之间的亲密程度越高。

[0032] 在一种实现方式中,电子设备可以预先设置第一用户与第二用户之间的关联关系及其关联度。在一种实现方式中,用户可以自定义第一用户与第二用户之间的关联关系及其关联度,本发明实施例对此不作限定。

[0033] 在一种实现方式中,电子设备可以将第一用户与第二用户的关联关系、以及第三用户与其他用户的关联关系存储在预设数据库中的同一个关联关系表中,其中,第三用户与第一用户(和第二用户)不同,其他用户是指与第三用户具有关联关系的用户。在另一种实现方式中,电子设备可以将第一用户与第二用户的关联关系、第三用户与其他用户的关联关系分别存储在不同的关联关系表中,本发明实施例对此不作限定。

[0034] 在一种实现方式中,与第一用户具有关联关系的第二用户的个数可以是1个或多个。当与第一用户具有关联关系的第二用户的个数为多个时,各个第二用户与该第一用户之间的关联度相互独立,即各个第二用户与该第一用户之间的关联度即可以相同,也可以

不同。例如，第一用户与第二用户的关联关系表可以如下所示：

[0035] 表1 第一用户与第二用户的关联关系表

[0036]

第一用户标识	第二用户标识	关联度
张三	李四	0.2
张三	王五	0.3

[0037] S103、电子设备获取第二用户的逾期记录，该逾期记录包括第二用户的特征及其对应的第一特征值。

[0038] 在一种实现方式中，逾期记录可以预先存储在电子设备中，或者，逾期记录可以存储在云端服务器中，电子设备通过访问云端服务器获取第二用户的逾期记录。在一种实现方式中，该逾期记录可以包括是否正常还款、是否逾期和是否被拒中的至少一个，及其相应的值。其中，被拒是指用户提交贷款信息时，审核结果为不通过，由此而被拒贷。例如，第二用户的逾期记录可以如下表所示：

[0039] 表2 第二用户的逾期记录

[0040]

第二用户标识	是否正常还款	是否逾期	是否被拒
李四	1	0	0

[0041] 需要说明的是，第二用户的逾期记录中，是否正常还款、是否逾期、是否被拒的值中只有一个值为1，也就是说，若第二用户属于正常还款、逾期和被拒中的任意一种情况时，其余两种情况的值为0。

[0042] 在一种实现方式中，第二用户的特征可以为正常还款、逾期和被拒中的一个。例如，若第二用户的逾期记录如表2所示，则第二用户的特征为正常还款，并且第二用户的特征对应的第一特征值为1。

[0043] S104、电子设备基于前述关联度和该第一特征值，得到第一用户的逾期概率。其中，第一用户的逾期概率的取值范围为[0,1]。

[0044] 具体的，当第二用户的特征为逾期时，电子设备可以将前述关联度和该第一特征值的乘积确定为第一用户的逾期概率。例如，当第一用户与第二用户的关联度为0.2，且第二用户的特征为逾期，第二用户的特征的第一特征值为1时，第一用户的逾期概率为： $P=0.2*1=0.2$ 。

[0045] 在一种实现方式中，当第二用户的特征为正常还款时，电子设备可以将第一用户的逾期概率确定为0。当第二用户的特征为被拒时，电子设备可以将第一用户的逾期概率确定为1，进一步的，电子设备可以将第一用户加入黑名单。当电子设备接收到关于黑名单中的用户的贷款请求时，可以不用对该用户进行逾期预测，直接拒绝该用户的贷款请求。

[0046] 在一种实现方式中，当与第一用户具有关联关系的第二用户的个数为多个时，电子设备可以基于各个第二用户与第一用户的关联度，对各个第二用户的第一特征值进行加权求和，得到第一用户的逾期概率。例如，当第一用户为用户1，与用户1具有关联关系的第二用户的个数为2个，分别为用户2和用户3，并且用户1与用户2的关联度为0.2，用户1与用户3的关联度为0.3，用户2和用户3的特征均为逾期，用户2和用户3的特征的第一特征值均为1时，第一用户的逾期概率为： $P=(0.2*1+0.3*1)/2=0.25$ 。

[0047] 需要说明的是,上述例子中的数值仅用于方便举例,并不构成对本发明实施例的限定。

[0048] 通过实施本发明实施例,可以根据与第一用户关联的第二用户(即老用户)的逾期记录,以及第一用户与第二用户的关联度,预测第一用户(即新用户)的逾期概率。

[0049] 请参阅图2,图2是本发明实施例提供的另一种概率预测方法的流程示意图。具体的,如图2所示,本发明实施例的另一种概率预测方法可以包括但不限于以下步骤:

[0050] S201、电子设备接收预测请求,该预测请求包括第一用户的第一用户标识,以及第一用户的身份信息及其对应的第二特征值。

[0051] 具体的,电子设备可以在接收到预测请求的情况下,提取该预测请求包括的第一用户的第一用户标识,并提取第一用户的身份信息及其对应的第二特征值。其中,第一用户的身份信息可以包括第一用户的性别、年龄段、学历,是否黑名单和第一用户的贷款申请信息中的一个或多个。第一用户的身份信息对应的第二特征值的个数与身份信息包括的类别数相同,例如,当第一用户的身份信息包括第一用户的性别、年龄段和学历这3个类别时,第一用户的身份信息对应的第二特征值的数量为3个,且分别为第一用户的性别对应的第二特征值、第一用户的年龄段对应的第二特征值、第一用户的学历对应的第二特征值。

[0052] 在一种实现方式中,当第一用户的身份信息包括多个类别时,不同类别的身份信息对应的第二特征值相互独立,即不同类别的身份信息对应的第二特征值可以相同,也可以不同,本发明实施例对此不作限定。

[0053] 在一种实现方式中,电子设备提取到第一用户的身份信息之后,还可以对第一用户的身份信息进行特征处理,得到第一用户的身份信息对应的第二特征值。例如,当第一用户的身份信息包括第一用户的性别、年龄段和学历这3个类别,且第一用户的性别为男,第一用户的年龄段为中年,第一用户的学历为硕士时,电子设备对第一用户的身份信息进行特征处理后,得到的第一用户的性别对应的第二特征值可以为0.1,第一用户的年龄段对应的第二特征值可以为0.2,第一用户的学历对应的第二特征值可以为0.1。又如,当第一用户的性别为女,第一用户的年龄段为青年,第一用户的学历为本科时,电子设备对第一用户的身份信息进行特征处理后,得到的第一用户的性别对应的第二特征值可以为0.2,第一用户的年龄段对应的第二特征值可以为0.3,第一用户的学历对应的第二特征值可以为0.2。

[0054] 在一种实现方式中,第一用户的身份信息可以是第一用户在提交贷款请求时,由第一用户输入的。

[0055] S202、电子设备在预设数据库中获取包含第一用户标识的即时通讯信息和/或逾期信息。

[0056] 其中,即时通讯信息可以是通话记录信息、微信或者QQ等即时通讯软件中的聊天记录信息中的至少一种,逾期信息可以是贷款合同信息或者前述逾期记录中的至少一种。相应的,第一用户标识可以是第一用户的昵称、身份证号或者第一用户的电话号码。

[0057] 在一种实现方式中,即时通讯信息和/或逾期信息可以记录与第一用户具有关联关系的第二用户的用户标识,或者,即时通讯信息和/或逾期信息可以记录与第一用户标识具有关联关系的第二用户标识,其中,第二用户标识与第二用户相对应。

[0058] S203、电子设备根据所述包含第一用户标识的即时通讯信息和/或逾期信息,确定与该第一用户关联的第二用户。



[0059] 在一种实现方式中,电子设备可以根据所述包含第一用户标识的逾期信息,确定与该第一用户关联的第二用户。在一种实现方式中,电子设备可以根据所述包含第一用户标识的即时通讯信息和逾期信息,确定与该第一用户关联的第二用户。在一种实现方式中,电子设备可以先根据所述包含第一用户标识的即时通讯信息,再根据所述包含第一用户标识的逾期信息,确定与该第一用户关联的第二用户。在一种实现方式中,电子设备可以先根据所述包含第一用户标识的逾期信息,再根据所述包含第一用户标识的即时通讯信息,确定与该第一用户关联的第二用户。

[0060] 在一种实现方式中,电子设备可以根据所述包含第一用户标识的即时通讯信息和逾期信息,构建包含第一用户标识的知识图谱,该知识图谱记录了与第一用户具有关联关系的第二用户的用户标识。在一种实现方式中,电子设备可以根据知识图谱,确定与该第一用户关联的第二用户。具体的,电子设备可以在预设数据库中获取该知识图谱,该知识图谱的顶点包括用户标识或电话号码;并在该知识图谱中,确定与第一用户标识对应的顶点关联的至少一个目标顶点,该目标顶点通过满足预设数量的边与该第一用户标识对应的顶点相连;将各个目标顶点包含的用户标识对应的用户确定为与该第一用户关联的第二用户。以图3所示的知识图谱的示意图为例,其中,甲、乙、丙、丁、己为用户标识,A、B、C、D、E为电话号码,当第一用户标识为甲时,第一用户标识对应的顶点关联的目标顶点可以为乙、丁、己和丙(即第二用户可以为乙、丁、己和丙),其中,乙和丁均通过2条边与甲相连,己和丙均通过3条边与甲相连。

[0061] 需要说明的是,图3所示知识图谱仅用于举例,并不构成对本发明实施例的限定。在其他可行的实现方式中,知识图谱还可以包括其他用户标识和电话号码,可以理解的是,知识图谱中的两个用户标识之间还可以通过4、5、8或者其他数量的边相连,本发明实施例对此不作限定。

[0062] S204、电子设备根据所述包含第一用户标识的即时通讯信息和/或逾期信息,确定该第一用户与第二用户的关联度。

[0063] 在一种实现方式中,若电子设备根据所述包含第一用户标识的逾期信息,确定与该第一用户关联的第二用户,则电子设备可以根据所述包含第一用户标识的逾期信息,确定该第一用户与第二用户的关联度。在一种实现方式中,若电子设备根据所述包含第一用户标识的即时通讯信息和逾期信息,确定与该第一用户关联的第二用户,则电子设备可以根据所述包含第一用户标识的即时通讯信息和逾期信息,确定该第一用户与第二用户的关联度。

[0064] 在一种实现方式中,若电子设备根据知识图谱,确定与该第一用户关联的第二用户,则电子设备可以根据知识图谱,确定该第一用户与第二用户的关联度。在一种实现方式中,若第一用户对应的顶点在知识图谱中通过预设数量的边与第二用户对应的顶点相连,则该预设数量越小,第一用户和第二用户的关联度越大,同理,该预设数量越大,第一用户和第二用户的关联度越小。以图3所示的知识图谱的示意图为例,当第一用户标识为甲时,第一用户标识对应的顶点关联的目标顶点可以为乙、丁、己和丙(即第二用户可以为乙、丁、己和丙),其中,乙和丁均通过2条边与甲相连,己和丙均通过3条边与甲相连,此时,甲和乙(或者丁)的关联度可以为0.6,甲和己(或者丙)的关联度可以为0.4。

[0065] 在一种实现方式中,电子设备可以根据第一用户与第二用户的第一关联度和第二

关联度,得到第一用户与第二用户的关联度。具体的,电子设备可以将第一关联度和第二关联度的乘积,确定为第一用户与第二用户的关联度。其中,第一关联度可以根据前述预设数量得到的。在一种实现方式中,知识图谱中的各条边具有权重,电子设备可以根据知识图谱中与第一用户相关联的边的权重,确定第二关联度。在一种实现方式中,电子设备根据知识图谱,确定与该第一用户关联的第二用户之后,还可以执行如下步骤:在知识图谱中,确定与第一用户标识对应的至少一个目标边集合,该目标边集合包括前述预设数量的边,第一用户标识对应的顶点和前述目标顶点通过该目标边集合中的所有边相连;针对各个目标边集合,根据该目标边集合中的所有边的权重,确定第二关联度。

[0066] 以图3所示的知识图谱的示意图为例,其中,知识图谱中的边旁边的数字为该条边的权重,当第一用户标识为甲时,第一用户标识对应的顶点关联的目标顶点为乙、丁、己和丙(即第二用户可以为乙、丁、己和丙),其中,乙和丁均通过2条边与甲相连,己和丙均通过3条边与甲相连。此时,甲和乙(或者丁)的第一关联度可以为0.6,甲和己(或者丙)的第一关联度可以为0.4,甲和乙的第二关联度可以为 $0.4 \times 0.4 = 0.16$ ,甲和丁的第二关联度可以为 $0.2 \times 1 = 0.2$ ,甲和己的第二关联度可以为 $0.3 \times 0.7 \times 1 = 0.21$ ,甲和丙的第二关联度可以为 $0.2 \times 0.3 \times 1 = 0.06$ ,因此,甲和乙的关联度可以为 $0.6 \times 0.16 = 0.096$ ,甲和丁的关联度可以为 $0.6 \times 0.2 = 0.12$ ,甲和己的关联度可以为 $0.4 \times 0.21 = 0.084$ ,甲和丙的关联度可以为 $0.4 \times 0.06 = 0.024$ 。

[0067] S205、电子设备基于该关联度、前述第一特征值和前述第二特征值,得到该第一用户的逾期概率。

[0068] 具体的,当与第一用户具有关联关系的第二用户的个数为多个时,电子设备可以基于各个第二用户与第一用户的关联度,对各个第二用户的第一特征值进行加权求和,得到第一数值,并对第一数值和前述第二特征值进行求和,将得到的结果确定为第一用户的逾期概率。例如,当第一用户为用户1,与用户1具有关联关系的第二用户的个数为2个,分别为用户2和用户3,并且用户1与用户2的关联度为0.2,用户1与用户3的关联度为0.3,用户2和用户3的特征均为逾期,用户2和用户3的特征的第一特征值均为1时,第一数值为 $(0.2 \times 1 + 0.3 \times 1) / 2 = 0.25$ ;当第一用户的身份信息包括第一用户的性别、年龄段和学历这3个类别,且第一用户的性别为男,第一用户的年龄段为中年,第一用户的学历为硕士时,电子设备对第一用户的身份信息进行特征处理后,得到的第一用户的性别对应的第二特征值为0.1,第一用户的年龄段对应的第二特征值为0.2,第一用户的学历对应的第二特征值为0.1时,第一用户的逾期概率为: $P = 0.25 + (0.1 + 0.2 + 0.1) = 0.65$ 。

[0069] 在一种实现方式中,电子设备还可以获取第一系数和第二系数,其中,第一系数对应第一用户的身份信息,第二系数对应与第一用户关联的所有第二用户。在一种实现方式中,第一系数和第二系数可以是电子设备默认设置的,也可以是用户自定义的,本发明实施例对此不作限定。

[0070] 在一种实现方式中,电子设备获取第一系数和第二系数之后,还可以基于该关联度、第一系数、第二系数、前述第一特征值和前述第二特征值,得到该第一用户的逾期概率。例如,当对与第一用户关联的各个第二用户的第一特征值进行加权求和,得到的第一数值为 $(0.2 \times 1 + 0.3 \times 1) / 2 = 0.25$ ,且根据第一用户的身份信息得到的第一用户的性别对应的第二特征值为0.1,第一用户的年龄段对应的第二特征值为0.2,第一用户的学历对应的第二

特征值为0.1,第一系数为0.6,第二系数为0.4时,第一用户的逾期概率为: $P=0.6*0.25+0.4*(0.1+0.2+0.1)=0.31$ 。

[0071] 通过实施本发明实施例,不仅根据与第一用户关联的第二用户(即老用户)的逾期记录以及第一用户与第二用户的关联度,还根据第一用户的身份信息,对第一用户(即新用户)的逾期概率进行预测,相较于仅根据第二用户的逾期记录以及第一用户与第二用户的关联度,对第一用户的逾期概率进行预测的方式,本发明实施例所采用的方法由于考虑了第一用户本人的身份信息,对第一用户的逾期概率进行预测,可以进一步提高预测准确度。

[0072] 请参见图4,图4是本发明实施例提供的一种概率预测装置的结构示意图,具体的,如图4所示,该概率预测装置40,可以包括:

[0073] 接收单元401,用于接收预测请求,该预测请求包括第一用户的第一用户标识。

[0074] 处理单元402,用于在预设数据库中确定与第一用户关联的第二用户,以及该第一用户与第二用户的关联度。

[0075] 该处理单元402,还用于获取第二用户的逾期记录,该逾期记录包括第二用户的特征及其对应的第一特征值。

[0076] 该处理单元402,还用于基于前述关联度和该第一特征值,得到第一用户的逾期概率。

[0077] 在一种实现方式中,前述预测请求还包括第一用户的身份信息及其对应的第二特征值;处理单元402,具体可以用于基于前述关联度、第一特征值和第二特征值,得到第一用户的逾期概率。

[0078] 在一种实现方式中,处理单元402,具体可以用于在预设数据库中获取包含第一用户标识的即时通讯信息和/或逾期信息;根据所述包含第一用户标识的即时通讯信息和/或逾期信息,确定与第一用户关联的第二用户。

[0079] 在一种实现方式中,处理单元402,具体可以用于根据所述包含第一用户标识的即时通讯信息和/或逾期信息,确定第一用户与第二用户的关联度。

[0080] 在一种实现方式中,第二用户的数量可以为多个;处理单元402,具体可以用于基于各个第二用户与该第一用户的关联度,对各个第二用户的第一特征值进行加权求和,得到该第一用户的逾期概率。

[0081] 本发明实施例和图1、图2所示方法实施例基于同一构思,其带来的技术效果也相同,具体原理请参照图1、图2所示实施例的描述,在此不赘述。

[0082] 请参阅图5,图5是本发明实施例提供的一种电子设备的结构示意图。该电子设备50可以包括接收器501、存储器502和处理器503,接收器501、存储器502和处理器503通过一条或多条通信总线连接。

[0083] 接收器501可以用于接收数据,例如,接收器501可以用于接收预测请求。

[0084] 存储器502可以包括只读存储器和随机存取存储器,并向处理器503提供指令和数据。存储器502的一部分还可以包括非易失性随机存取存储器。

[0085] 处理器503可以是中央处理单元(Central Processing Unit,CPU),该处理器503还可以是其他通用处理器、数字信号处理器(Digital Signal Processor,DSP)、专用集成电路(Application Specific Integrated Circuit,ASIC)、现成可编程门阵列(Field-Programmable Gate Array,FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、

分立硬件组件等。通用处理器可以是微处理器,可选的,该处理器503也可以是任何常规的处理器等。其中:

[0086] 存储器502,用于存储程序指令。

[0087] 处理器503,用于调用存储器502中存储的程序指令,以用于:

[0088] 接收预测请求,该预测请求包括第一用户的第一用户标识;

[0089] 在预设数据库中确定与该第一用户关联的第二用户,以及该第一用户与第二用户的关联度;

[0090] 获取第二用户的逾期记录,该逾期记录包括第二用户的特征及其对应的第一特征值;

[0091] 基于前述关联度和该第一特征值,得到第一用户的逾期概率。

[0092] 在一种实现方式中,预测请求还包括该第一用户的身份信息及其对应的第二特征值;处理器503用于基于前述关联度和该第一特征值,得到第一用户的逾期概率时,具体可以用于基于前述关联度、该第一特征值和第二特征值,得到第一用户的逾期概率。

[0093] 在一种实现方式中,处理器503用于在预设数据库中确定与第一用户关联的第二用户时,具体可以用于在预设数据库中获取包含第一用户标识的即时通讯信息和/或逾期信息;根据所述包含第一用户标识的即时通讯信息和/或逾期信息,确定与第一用户关联的第二用户。

[0094] 在一种实现方式中,处理器503用于在预设数据库中确定所述第一用户与所述第二用户的关联度时,具体可以用于根据所述包含第一用户标识的即时通讯信息和/或逾期信息,确定第一用户与第二用户的关联度。

[0095] 在一种实现方式中,第二用户的数量可以为多个;处理器503用于基于关联度和第一特征值,得到第一用户的逾期概率时,具体可以用于基于各个第二用户与第一用户的关联度,对各个第二用户的第一特征值进行加权求和,得到第一用户的逾期概率。

[0096] 需要说明的是,图5对应的实施例中未提及的内容以及各个步骤的具体实现方式可参见图1-图4所示实施例以及前述内容,这里不再赘述。

[0097] 本发明实施例还提供一种计算机可读存储介质,计算机可读存储介质存储有计算机程序,计算机程序包括程序指令,程序指令被处理器执行时,使处理器执行如图1-图2所示方法实施例中所执行的步骤。

[0098] 以上所揭露的仅为本发明的部分实施例而已,当然不能以此来限定本发明之权利范围,本领域普通技术人员可以理解实现上述实施例的全部或部分流程,并依本发明权利要求所作的等同变化,仍属于发明所涵盖的范围。

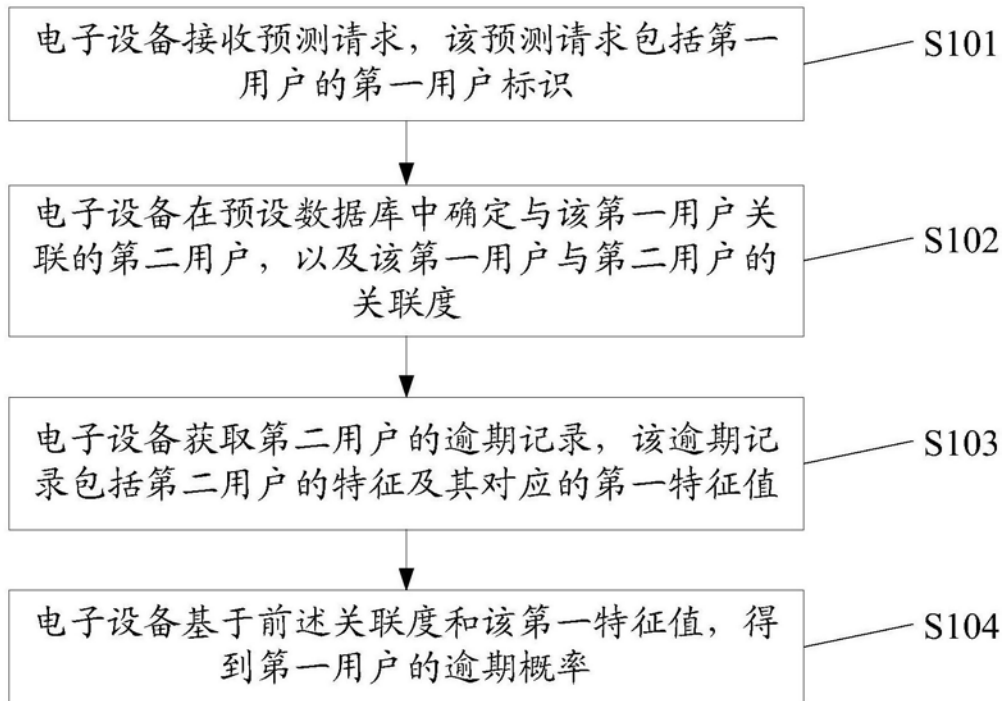


图1

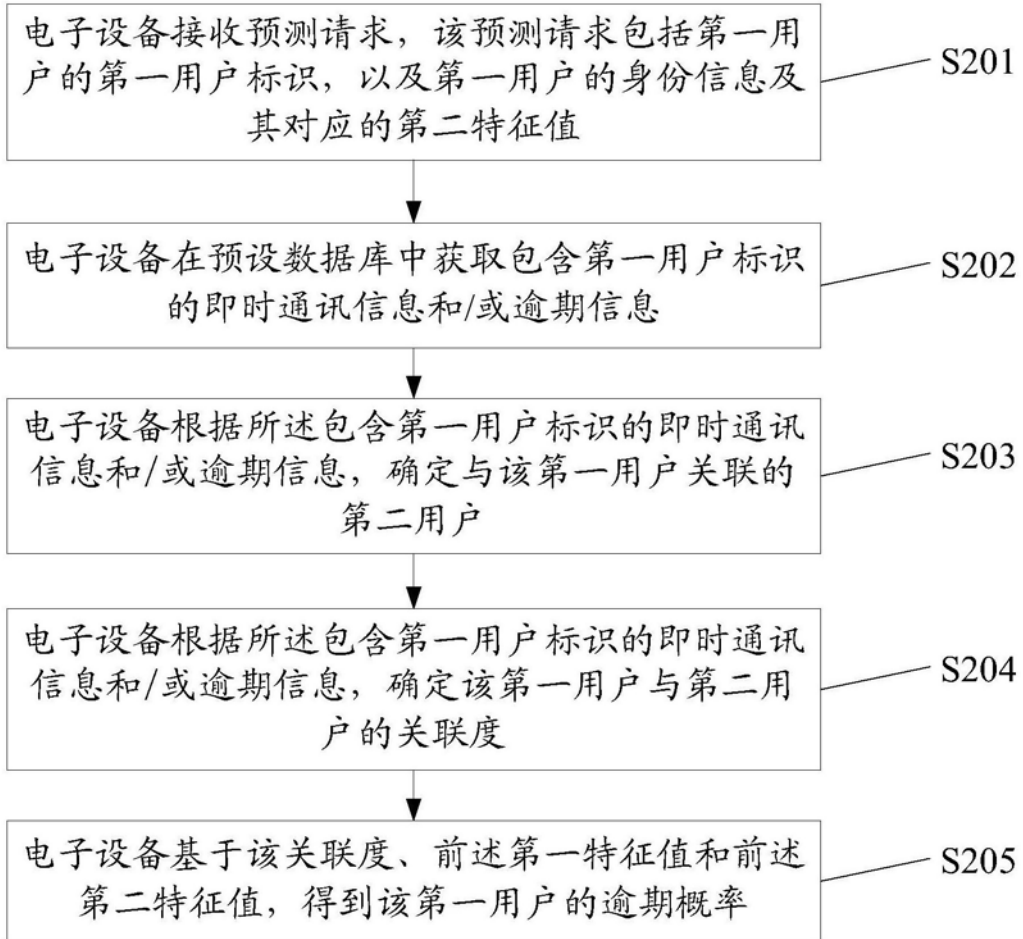


图2

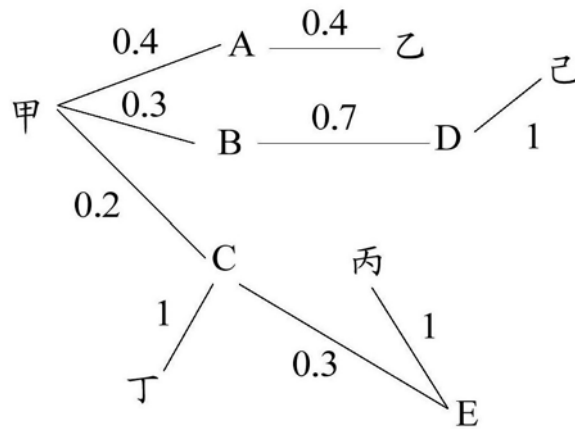


图3

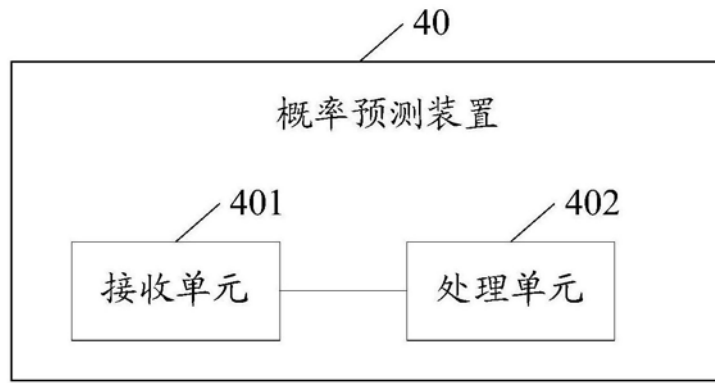


图4

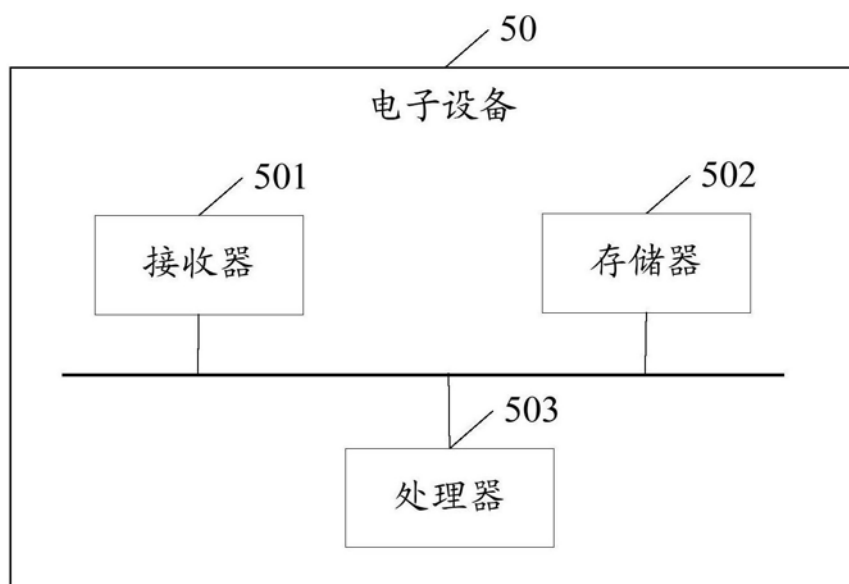


图5