



(12)发明专利申请

(10)申请公布号 CN 110851729 A

(43)申请公布日 2020.02.28

(21)申请号 201911145222.1

(22)申请日 2019.11.19

(71)申请人 深圳前海微众银行股份有限公司
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 钱锦锋 万磊 李毅 李旭锋
尹源春 柏勇 张双合 赵焕胜
颜超

(74)专利代理机构 深圳市世纪恒程知识产权代理事务所 44287
代理人 晏波

(51)Int.Cl.
G06F 16/9535(2019.01)

权利要求书2页 说明书12页 附图2页

(54)发明名称

资源信息推荐方法、装置、设备及计算机存储介质

(57)摘要

本发明涉及金融科技(Fintech)技术领域,公开了一种资源信息推荐方法、装置、设备和计算机存储介质,该方法包括:获取各资源信息的历史数据,基于特征选择在各历史数据中获取与用户账号具有相关性的多个历史特征值;计算用户账号的当前特征值与各历史特征值之间的相似度值,在各相似度值中获取预设数量的目标相似度值;基于各资源信息确定各目标相似度值对应的目标资源信息,获取各目标资源信息中的参数信息;基于预设推荐算法对各目标相似度值和各参数信息进行计算,以获取各资源信息对应的推荐指数,在各推荐指数中确定数值最高的目标推荐指数,对目标推荐指数对应的资源信息进行推荐。本发明提高了金融平台中推荐算法推荐的准确性。



1. 一种资源信息推荐方法,其特征在于,所述资源信息推荐方法包括如下步骤:

获取各资源信息的历史数据,并基于特征选择在各所述历史数据中获取与用户账号具有相关性的多个历史特征值;

计算用户账号的当前特征值与各所述历史特征值之间的相似度值,并在各所述相似度值中获取预设数量的目标相似度值;

基于各所述资源信息确定各所述目标相似度值对应的目标资源信息,并获取各所述目标资源信息中的参数信息;

基于预设推荐算法对各所述目标相似度值和各所述参数信息进行计算,以获取各所述资源信息对应的推荐指数,并在各所述推荐指数中确定数值最高的目标推荐指数,对所述目标推荐指数对应的资源信息进行推荐。

2. 如权利要求1所述的资源信息推荐方法,其特征在于,所述基于预设推荐算法对各所述目标相似度值和各所述参数信息进行计算,以获取各所述资源信息对应的推荐指数的步骤,包括:

依次遍历各所述目标相似度值,并确定当前遍历的当前目标相似度值对应的参数信息;

根据预设推荐算法对所述当前目标相似度值和所述当前目标相似度值对应的参数信息进行计算,以获取所述当前目标相似度值对应的资源信息的推荐指数,直至各所述目标相似度值遍历完成。

3. 如权利要求1所述的资源信息推荐方法,其特征在于,所述计算用户账号的当前特征值与各所述历史特征值之间的相似度值的步骤,包括:

获取所述当前特征值对应的当前标准化特征值和各所述历史特征值对应的历史标准化特征值;

根据所述当前标准化特征值和各所述历史标准化特征值确定所有相似度值。

4. 如权利要求3所述的资源信息推荐方法,其特征在于,所述获取所述当前特征值对应的当前标准化特征值和各所述历史特征值对应的历史标准化特征值的步骤,包括:

基于预设数据处理方式对所述当前特征值进行标准化,以获取所述当前特征值对应的当前标准化特征值;

基于所述预设数据处理方式对各所述历史特征值进行标准化,以获取各所述历史特征值对应的历史标准化特征值。

5. 如权利要求3所述的资源信息推荐方法,其特征在于,所述根据所述当前标准化特征值和各所述历史标准化特征值确定所有相似度值的步骤,包括:

依次遍历各所述历史标准化特征值,并计算当前遍历的历史标准化特征值和所述当前标准化特征值之间的欧式距离;

基于所述欧式距离确定当前遍历的历史标准化特征值和当前标准化特征值的相似度值,直至各所述历史标准化特征值遍历完成。

6. 如权利要求1所述的资源信息推荐方法,其特征在于,所述基于特征选择在各所述历史数据中获取与用户账号具有相关性的多个历史特征值的步骤,包括:

基于预设校验算法对各所述历史数据进行离散化,以获取各所述历史数据对应的初级特征值;

基于所述预设校验算法对各所述初级特征值进行评分,并基于评分结果获取与用户账号具有相关性的多个历史特征值。

7.如权利要求1-6任一项所述的资源信息推荐方法,其特征在于,所述计算用户账号的当前特征值与各所述历史特征值之间的相似度值的步骤之前,包括:

获取用户账号的当前数据,并对所述当前数据进行离散数据编码,以获取所述当前账号的当前特征值。

8.一种资源信息推荐装置,其特征在于,所述资源信息推荐装置包括:

获取模块,用于获取各资源信息的历史数据,并基于特征选择在各所述历史数据中获取与用户账号具有相关性的多个历史特征值;

计算模块,用于计算用户账号的当前特征值与各所述历史特征值之间的相似度值,并在各所述相似度值中获取预设数量的目标相似度值;

确定模块,用于基于各所述资源信息确定各所述目标相似度值对应的目标资源信息,并获取各所述目标资源信息中的参数信息;

推荐模块,用于基于预设推荐算法对各所述目标相似度值和各所述参数信息进行计算,以获取各所述资源信息对应的推荐指数,并在各所述推荐指数中确定数值最高的目标推荐指数,对所述目标推荐指数对应的资源信息进行推荐。

9.一种资源信息推荐设备,其特征在于,所述资源信息推荐设备包括:存储器、处理器及存储在所述存储器上并可在所述处理器上运行的资源信息推荐程序,所述资源信息推荐程序被所述处理器执行时实现如权利要求1至7中任一项所述的资源信息推荐方法的步骤。

10.一种计算机存储介质,其特征在于,所述计算机存储介质上存储有资源信息推荐程序,所述资源信息推荐程序被处理器执行时实现如权利要求1至7中任一项所述的资源信息推荐方法的步骤。

资源信息推荐方法、装置、设备及计算机存储介质

技术领域

[0001] 本发明涉及金融科技 (Fintech) 技术领域, 尤其涉及资源信息推荐方法、装置、设备及计算机存储介质。

背景技术

[0002] 随着计算机技术的发展, 越来越多的技术应用在金融领域, 传统金融业正在逐步向金融科技 (Fintech) 转变, 大数据中的企业贷款技术也不例外, 但由于金融行业的安全性、实时性要求, 也对技术提出的更高的要求。例如在金融平台对各个企业推荐资源信息时, 往往需要对已有的历史数据进行分析, 并根据分析结果向各个企业进行推荐。但目前金融平台的推荐方式一般是基于用户的协同推荐算法或者是基于资源信息的协同推荐算法来进行推荐的, 但是由于这两种推荐方式中的推荐算法并没有考虑企业经营状况和资源信息的变化程度等, 往往导致推荐给企业的资源信息不准确。因此, 如何提高金融平台中推荐算法推荐的准确性成为了目前亟待解决的技术问题。

发明内容

[0003] 本发明的主要目的在于提出一种资源信息推荐方法、装置、设备及计算机存储介质, 旨在提高金融平台中推荐算法推荐的准确性。

[0004] 为实现上述目的, 本发明提供一种资源信息推荐方法, 所述资源信息推荐方法包括如下步骤:

[0005] 获取各资源信息的历史数据, 并基于特征选择在各所述历史数据中获取与用户账号具有相关性的多个历史特征值;

[0006] 计算用户账号的当前特征值与各所述历史特征值之间的相似度值, 并在各所述相似度值中获取预设数量的目标相似度值;

[0007] 基于各所述资源信息确定各所述目标相似度值对应的目标资源信息, 并获取各所述目标资源信息中的参数信息;

[0008] 基于预设推荐算法对各所述目标相似度值和各所述参数信息进行计算, 以获取各所述资源信息对应的推荐指数, 并在各所述推荐指数中确定数值最高的目标推荐指数, 对所述目标推荐指数对应的资源信息进行推荐。

[0009] 可选地, 所述基于预设推荐算法对各所述目标相似度值和各所述参数信息进行计算, 以获取各所述资源信息对应的推荐指数的步骤, 包括:

[0010] 依次遍历各所述目标相似度值, 并确定当前遍历的当前目标相似度值对应的参数信息;

[0011] 根据预设推荐算法对所述当前目标相似度值和所述当前目标相似度值对应的参数信息进行计算, 以获取所述当前目标相似度值对应的资源信息的推荐指数, 直至各所述目标相似度值遍历完成。

[0012] 可选地, 所述计算用户账号的当前特征值与各所述历史特征值之间的相似度值的

步骤,包括:

[0013] 获取所述当前特征值对应的当前标准化特征值和各所述历史特征值对应的历史标准化特征值;

[0014] 根据所述当前标准化特征值和各所述历史标准化特征值确定所有相似度值。

[0015] 可选地,所述获取所述当前特征值对应的当前标准化特征值和各所述历史特征值对应的历史标准化特征值的步骤,包括:

[0016] 基于预设数据处理方式对所述当前特征值进行标准化,以获取所述当前特征值对应的当前标准化特征值;

[0017] 基于所述预设数据处理方式对各所述历史特征值进行标准化,以获取各所述历史特征值对应的历史标准化特征值。

[0018] 可选地,所述根据所述当前标准化特征值和各所述历史标准化特征值确定所有相似度值的步骤,包括:

[0019] 依次遍历各所述历史标准化特征值,并计算当前遍历的历史标准化特征值和所述当前标准化特征值之间的欧式距离;

[0020] 基于所述欧式距离确定当前遍历的历史标准化特征值和当前标准化特征值的相似度值,直至各所述历史标准化特征值遍历完成。

[0021] 可选地,所述基于特征选择在各所述历史数据中获取与用户账号具有相关性的多个历史特征值的步骤,包括:

[0022] 基于预设校验算法对各所述历史数据进行离散化,以获取各所述历史数据对应的初级特征值;

[0023] 基于所述预设校验算法对各所述初级特征值进行评分,并基于评分结果获取与用户账号具有相关性的多个历史特征值。

[0024] 可选地,所述计算用户账号的当前特征值与各所述历史特征值之间的相似度值的步骤之前,包括:

[0025] 获取用户账号的当前数据,并对所述当前数据进行离散数据编码,以获取所述当前账号的当前特征值。

[0026] 此外,为实现上述目的,本发明还提供一种资源信息推荐装置,所述资源信息推荐装置包括:

[0027] 获取模块,用于获取各资源信息的历史数据,并基于特征选择在各所述历史数据中获取与用户账号具有相关性的多个历史特征值;

[0028] 计算模块,用于计算用户账号的当前特征值与各所述历史特征值之间的相似度值,并在各所述相似度值中获取预设数量的目标相似度值;

[0029] 确定模块,用于基于各所述资源信息确定各所述目标相似度值对应的目标资源信息,并获取各所述目标资源信息中的参数信息;

[0030] 推荐模块,用于基于预设推荐算法对各所述目标相似度值和各所述参数信息进行计算,以获取各所述资源信息对应的推荐指数,并在各所述推荐指数中确定数值最高的目标推荐指数,对所述目标推荐指数对应的资源信息进行推荐。

[0031] 此外,为实现上述目的,本发明还提供一种资源信息推荐设备,所述资源信息推荐设备包括:存储器、处理器及存储在所述存储器上并可在所述处理器上运行的资源信息推

荐程序,所述资源信息推荐程序被所述处理器执行时实现如上所述的资源信息推荐方法的步骤。

[0032] 此外,为实现上述目的,本发明还提供一种计算机存储介质,所述计算机存储介质上存储有资源信息推荐程序,所述资源信息推荐程序被处理器执行时实现如上所述的资源信息推荐方法的步骤。

[0033] 本发明通过获取各资源信息的历史数据,并基于特征选择在各所述历史数据中获取与用户账号具有相关性的多个历史特征值;计算用户账号的当前特征值与各所述历史特征值之间的相似度值,并在各所述相似度值中获取预设数量的目标相似度值;基于各所述资源信息确定各所述目标相似度值对应的目标资源信息,并获取各所述目标资源信息中的参数信息;基于预设推荐算法对各所述目标相似度值和各所述参数信息进行计算,以获取各所述资源信息对应的推荐指数,并在各所述推荐指数中确定数值最高的目标推荐指数,对所述目标推荐指数对应的资源信息进行推荐。通过获取各个资源信息的历史数据,并将用户账号的当前特征值和历史数据中的多个历史特征值进行比较,以确定相似度值,也就是在计算相似度值时,综合考虑了资源信息的历史数据,提高了向用户账号推荐资源信息的准确性,并且还需要在这些相似度值中确定目标相似度值,并根据参数信息确定目标推荐指数,将目标推荐指数对应的资源信息进行推荐,使得本次推荐的效果更好,更符合用户的需求,提高了金融平台中推荐算法推荐的准确性。

附图说明

[0034] 图1是本发明实施例方案涉及的硬件运行环境的设备结构示意图;

[0035] 图2为本发明资源信息推荐方法第一实施例的流程示意图;

[0036] 图3为本发明资源信息推荐装置的装置模块示意图;

[0037] 图4为本发明资源信息推荐方法中申请产品的流程示意图;

[0038] 图5为本发明资源信息推荐方法中的流程示意图。

[0039] 本发明目的的实现、功能特点及优点将结合实施例,参照附图做进一步说明。

具体实施方式

[0040] 应当理解,此处所描述的具体实施例仅仅用以解释本发明,并不用于限定本发明。

[0041] 如图1所示,图1是本发明实施例方案涉及的硬件运行环境的设备结构示意图。

[0042] 本发明实施例资源信息推荐设备可以是PC机或服务器设备,其上运行有Java虚拟机。

[0043] 如图1所示,该资源信息推荐设备可以包括:处理器1001,例如CPU,网络接口1004,用户接口1003,存储器1005,通信总线1002。其中,通信总线1002用于实现这些组件之间的连接通信。用户接口1003可以包括显示屏(Display)、输入单元比如键盘(Keyboard),可选用户接口1003还可以包括标准的有线接口、无线接口。网络接口1004可选的可以包括标准的有线接口、无线接口(如WI-FI接口)。存储器1005可以是高速RAM存储器,也可以是稳定的存储器(non-volatile memory),例如磁盘存储器。存储器1005可选的还可以是独立于前述处理器1001的存储装置。

[0044] 本领域技术人员可以理解,图1中示出的设备结构并不构成对设备的限定,可以包

括比图示更多或更少的部件,或者组合某些部件,或者不同的部件布置。

[0045] 如图1所示,作为一种计算机存储介质的存储器1005中可以包括操作系统、网络通信模块、用户接口模块以及资源信息推荐程序。

[0046] 在图1所示的设备中,网络接口1004主要用于连接后台服务器,与后台服务器进行数据通信;用户接口1003主要用于连接客户端(用户端),与客户端进行数据通信;而处理器1001可以用于调用存储器1005中存储的资源信息推荐程序,并执行下述资源信息推荐方法中的操作。

[0047] 基于上述硬件结构,提出本发明资源信息推荐方法实施例。

[0048] 参照图2,图2为本发明资源信息推荐方法第一实施例的流程示意图,所述方法包括:

[0049] 步骤S10,获取各资源信息的历史数据,并基于特征选择在各所述历史数据中获取与用户账号具有相关性的多个历史特征值;

[0050] 在本实施例中用户账号可以为企业在金融平台中登录的账号。目前,金融平台为企业推荐各个资源信息(如金融资源信息)的方式主要有:(1) topN:即按对金融平台上每个资源信息申请成功的企业数从多到少的排名,并按排名进行推荐;(2) 基于用户的协同推荐算法:主要原理是推荐和用户相似用户曾经成功申请的资源信息,主要步骤如下,找到与目标企业U相似的企业集合:即通过计算企业U和其它每个企业间的相似度找到最相似的企业集合,常用的相似度算法为Jaccard算法和余弦相似度算法:Jaccard算法:即两个企业申请资源信息的交集除以两个企业申请资源信息的并集。

[0051] 用户U和V相似度 $S_{uv} = \frac{|N(u) \cap N(v)|}{|N(u) \cup N(v)|}$ 余弦相似度算法:用户U和V相似度

$$S_{UV} = \frac{|N(u) \cap N(v)|}{\sqrt{|N(u) \cup N(v)|}}$$

[0052] 将企业U和其它企业相似度进行从大到小排序,选择前L个企业;针对这L个企业 $L = \{L_1, L_2, \dots, L_n\}$ 申请成功的资源信息集合 $P = \{P_1, P_2, \dots, P_n\}$, 资源信息 P_i 对用户的推荐指数 $C(P_i, U)$ 为:

$$[0053] \quad C(P_i, U) = \sum_{v \in L \cap N(p_i)} S_{uv}$$

[0054] $N(p_i)$ 为对申请成功过 P_i 的企业。对资源信息集合P根据每个 P_i 的推荐指数 $C(P_i, U)$ 值从大到小进行排序,选取前K个资源信息,得到企业U的信息推荐列表。

[0055] (3) 基于资源信息的协同推荐算法:主要原理是和用户曾经申请成功的资源信息越相似,推荐指数也越高,主要步骤同基于用户的协同推荐算法相似,具体如下:计算资源信息集合P中两两的相似度,即计算各资源信息之间的相似度 S_{ij} , 相似度算法采用Jaccard算法;计算资源信息P对于用户U的推荐指数 $C(P, U)$, 选取最高推荐指数的K个资源信息进行推荐,推荐指数 $C(P, U)$ 算法如下:

$$[0056] \quad C(P, U) = \sum_{j \in L \cap N(u)} S_{pj}$$

[0057] L为和资源信息j最为相似的1个资源信息。 $N(u)$ 为用户U曾经申请成功的资源信

息。

[0058] 但是由于大部分企业申请成功的资源信息数较少,特别是小微企业许多之前并没有申请成功过资源信息,因此计算相似度时绝大部分为0,基于用户和基于资源信息的协同推荐算法并不能获得很好的推荐效果。并且由于企业经营状况不断变化,资源信息的准入条件也在发生变化,之前的用户申请成功的资源信息及与其类型的资源信息,在经过一段时间后并不一定能申请成功,目前的算法并没有考虑到企业经营状况和产品的变化程度。而topN的算法推荐了申请成功企业最多的产品,但某个资源信息的申请成功的企业数量相比于企业的总量比例很小,并不适用于绝大部分企业,推荐效果不佳。

[0059] 因此,在本实施例中,通过根据平台企业基本信息及成功申请的历史数据,并通过特征选择筛选出独立且相关性较大的特征,即基于特征选择在各个历史数据中获取与用户账号具有相关性的多个历史特征值。

[0060] 并且由于金融平台会记录企业的基本信息和经营信息,以便银行进行初步筛选,市场化的金融平台会根据企业授权获得工商信息,政府化的金融平台可以通过当地征信平台或大数据局获得相关企业工商、税务数据。当企业申请资源信息时,金融平台会记录下企业申请资源信息时最新的工商、税务数据、申请结果等信息,将这些信息进行衍生特征加工可整理为如下表1。表中同一个企业会申请多个资源信息,同一个资源信息也会有多个企业申请,同一个企业也会申请同一个资源信息多次。

企业	资源信息	申请日期	结果	特征1-行业	特征2-上年营业收入	特征3-上年利润	特征4-上年营业成本	特征5-纳税人等级	..
[0061] Co mp1	Cp1	2016/7/3	失败	金融业	10000000	1000000	3000000	C	..
Co mp2	Cp2	2016/7/3	成功	采矿业	10000000	2000000	3000000	B	..
...

[0062] 表1

[0063] 其中资源信息一般会对企业所处行业进行判断,由于行业代码共分4层,1380个小类,不易区分,因此只取第1层的代码,既每个企业的行业类型只区分到第1层的大类。申请日期为当前日期的5年内。

[0064] 数据特征选择主要选择影响申请结果的主要特征,主要步骤如下:

[0065] 连续数据离散化:对表1中连续型变量通过卡方检验方法进行离散化,如对上年营业收入字段离散化后为:小于等于3000万为1,大于3000万并且小于等于1亿为2,大于1亿并且小于等于10亿为3,大于10亿为4。特征选择:采用过滤法中的卡方检验法对每个特征进行评分,选取最好的特征;通过特征选择,选取M个对企业成功申请资源信息相关性最高的特征。

[0066] 步骤S20,计算用户账号的当前特征值与各所述历史特征值之间的相似度值,并在各所述相似度值中获取预设数量的目标相似度值;

[0067] 当通过特征筛选,并且只选择申请成功的记录,可得如下表2:

企 业	资 源 信 息	申 请 日 期	选 择 特 征1- 行 业	选 择 特 征 2-上 年 营 业 收 入	选 择 特 征 3-上 年 利 润	选 择 特 征 4-上 年 营 业 成 本	选 择 特 征 5-纳 税 人 等 级	...
[0068] Cp 1	Cp 1	2016/ 7/3	金 融 业	10000000	100000 0	3000000	C	...
Cp 2	Cp 2	2016/ 7/3	采 矿 业	10000000	200000 0	3000000	B	...
...

[0069] 表2

[0070] 并且在企业在平台上基于用户账号进行登录后,也就是对于登录平台的企业,会根据企业当前最新的选择特征与各申请成功的企业特征进行相似度计算,计算步骤如下:对于用户账号的数据信息进行离散数据编码,即对于像纳税人等级的有序序列{A->B->C->D}转换为有序数列如{1、2、3、4};对于行业等无序序列,采用one-hot编码,即假设有3个行业{金融业、采矿业、文化娱乐},则金融业编码为{1,0,0},采矿业编码为{0,1,0},文化娱乐编码为{0,0,1}。既会增加特征向量,也就是获取到用户账号的当前特征值。并对当前特征值和各个历史特征值进行标准化特征,且方式均相同。

[0071] 例如,对表2中每个特征采用z-score(标准)方式进行标准化,标准化后的值为

$$x' = \frac{x - \bar{x}}{\sigma}$$

[0072] 计算企业当前标准化特征值与其它企业申请成功时的历史标准化特征值的相似度S,相似度采用欧式距离进行计算,距离越小,相似度越高。

$$[0073] \quad S_{(x, y)} = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

[0074] 其中n为标准化特征的个数;当计算出各个相似度后,选取与企业相似度最高的K(预设数量)个申请成功资源信息时的企业特征(即获取预设数量的目标相似度值),其对应的企业相似度为S(u,k),其对应的资源信息为L(L={L1,L2,L3……Ln})。

[0075] 步骤S30,基于各所述资源信息确定各所述目标相似度值对应的目标资源信息,并获取各所述目标资源信息中的参数信息;

[0076] 当获取到各个目标相似度值后,还需要获取各个资源信息的参数信息,如当资源信息为金融资源信息时,则参数信息可以包括金融资源信息的申请额度和申请时间。并在各个资源信息中确定与各个目标相似度值对应的目标资源信息,也就是目标资源信息的数量和目标相似度值的数量是相同的,并通过网络或者历史数据等方式获取各个目标资源信

息中的参数信息。其中,参数信息包括资源信息的额度限制,时间限制等。

[0077] 步骤S40,基于预设推荐算法对各所述目标相似度值和各所述参数信息进行计算,以获取各所述资源信息对应的推荐指数,并在各所述推荐指数中确定数值最高的目标推荐指数,对所述目标推荐指数对应的资源信息进行推荐。

[0078] 资源信息的推荐指数除了相似度指数外,由于企业申请到的额度越高,说明越符合资源信息的要求,因此和额度高的企业特征相似会有机会申请成功,同时由于资源信息根据经济周期也在变化,因此时间越早,参考性也越弱,因此对企业U和上一步骤产生的L中每个资源信息的推荐指数 $C(U, L_i)$ 计算公式如下:

$$[0079] \quad \sum_{N(L)} \frac{m * y}{S_{(u, k)} * \bar{M} * 10}$$

[0080] 其中m为K企业特征申请的信贷额度; \bar{M} 为L个资源信息最大贷款额度的平均值;y为K企业特征申请时的距离10年前的时间长度,以年为单位,取2位小数; $S(u, k)$ 为上一步获得的企业与K企业特征的相似度, L_i 为上一步获得的资源信息集合L中的资源信息; $N(L)$ 为 L_i 资源信息对应的K个企业特征个数;计算公式主要的含义是对资源信息集合L,可能多个企业特征 K_i 会对用同一个资源信息 L_i ,按公式计算资源信息 L_i 对应的多个企业特征 K_i 的推荐指数并汇总,就是用户对于资源信息 L_i 的推荐指数,最后选择推荐指数最高的D(预设数量)个资源信息进行推荐。

[0081] 另外,为辅助理解本申请中的资源信息推荐方法的流程,下面进行举例说明。

[0082] 例如,如图4所示,目前企业先在金融平台汇总进行申请,并通过金融平台进行数据转发(转发数据包括企业的申请信息)到金融机构,金融机构根据金融平台提供的数据对企业进行线上线下提供金融服务。单这样操作导致企业申请的成功率很低。而在本实施例中,则可以提高企业申请的成功率,如图5所示,先进行特征选择,即根据平台企业基本信息及成功申请的历史数据,通过特征选择筛选出独立且相关性较大的特征。再进行相似度计算,即使用企业当前特征信息计算企业与各资源信息成功申请时的特征的相似度,并筛选出相似度最高的K个企业申请时特征及对应的L个资源信息;再进行推荐指数计算,即根据成功申请的额度及申请时间计算这L个资源信息的推荐指数,选取推荐指数最高的信贷进行推荐;并进行效果定期评估,即跟踪推荐效果,并对特征进行定期更新,再每隔一段时间进行重新特征选取。

[0083] 在本实施例中,通过获取各资源信息的历史数据,并基于特征选择在各所述历史数据中获取与用户账号具有相关性的多个历史特征值;计算用户账号的当前特征值与各所述历史特征值之间的相似度值,并在各所述相似度值中获取预设数量的目标相似度值;基于各所述资源信息确定各所述目标相似度值对应的目标资源信息,并获取各所述目标资源信息中的参数信息;基于预设推荐算法对各所述目标相似度值和各所述参数信息进行计算,以获取各所述资源信息对应的推荐指数,并在各所述推荐指数中确定数值最高的目标推荐指数,对所述目标推荐指数对应的资源信息进行推荐。通过获取各个资源信息的历史数据,并将用户账号的当前特征值和历史数据中的多个历史特征值进行比较,以确定相似度值,也就是在计算相似度值时,综合考虑了资源信息的历史数据,提高了向用户账号推荐资源信息的准确性,并且还需要在这些相似度值中确定目标相似度值,并根据参数信息确定目标推荐指数,将目标推荐指数对应的资源信息进行推荐,使得本次推荐的效果更好,更

符合用户的需求,提高了金融平台中推荐算法推荐的准确性。

[0084] 进一步地,基于本发明资源信息推荐方法第一实施例,提出本发明资源信息推荐方法第二实施例。本实施例是本发明第一实施例的步骤S40,基于预设推荐算法对各所述目标相似度值和各所述参数信息进行计算,以获取各所述资源信息对应的推荐指数的步骤的细化,包括:

[0085] 步骤a,依次遍历各所述目标相似度值,并确定当前遍历的当前目标相似度值对应的参数信息;

[0086] 在获取到各个目标相似度值和各个参数信息后,依次遍历各个目标相似度值,并在各个参数信息中确定当前遍历的当前目标相似度值对应的参数信息。

[0087] 步骤b,根据预设推荐算法对所述当前目标相似度值和所述当前目标相似度值对应的参数信息进行计算,以获取所述当前目标相似度值对应的资源信息的推荐指数,直至各所述目标相似度值遍历完成。

[0088] 通过预设推荐指数计算公式来对当前目标相似度值、当前目标相似度值对应的参数信息(如申请额度和申请时间)进行计算,以获取到当前目标相似度值对应的资源信息的推荐指数,并对所有目标相似度采用同样的方式进行计算,得到各个目标相似度对应的资源信息的推荐指数。其中预设推荐指数计算公式可以是:

$$[0089] \quad \sum_{N(L)} \frac{m * y}{S_{(u, k)} * \bar{M} * 10}$$

[0090] 其中m为K企业特征申请的信贷额度; \bar{M} 为L个资源信息最大贷款额度的平均值;y为K企业特征申请时的距离10年前的时间长度,以年为单位,取2位小数; $S(u, k)$ 为上一步获得的企业与K企业特征的相似度, L_i 为上一步获得的资源信息集合L中的资源信息; $N(L)$ 为 L_i 资源信息对应的K个企业特征个数;计算公式主要的含义是对资源信息集合L,可能多个企业特征 K_i 会对用同一个资源信息 L_i ,按公式计算资源信息 L_i 对应的多个企业特征 K_i 的推荐指数并汇总,就是用户对于资源信息 L_i 的推荐指数,最后选择推荐指数最高的D(预设数量)个资源信息进行推荐。也就是对每个资源信息(即资源信息)均采用相同的方式进行计算得到各个推荐指数。

[0091] 在本实施例中,通过根据预设推荐指数计算公式、目标相似度值、参数信息计算推荐指数,从而保障了获取到的推荐指数的准确性和有效性。

[0092] 进一步地,计算用户账号的当前特征值与各所述历史特征值之间的相似度值的步骤,包括:

[0093] 步骤d,获取所述当前特征值对应的当前标准化特征值和各所述历史特征值对应的历史标准化特征值;

[0094] 当获取到当前特征值后,可以采用如z-score方式对当前当前特征值进行标准化,以得到当前特征值对应的当前标准化特征值,并采用同样的方式如z-score方式对各个历史特征值进行标准化,得到各个历史特征值对应的历史标准化特征值。

[0095] 步骤e,根据所述当前标准化特征值和各所述历史标准化特征值确定所有相似度值。

[0096] 当获取到当前标准化特征值和各个历史标准化特征值后,可以根据欧式距离来计算当前标准化特征值和各个历史标准化特征值之间的相似度值。

[0097] 在本实施例中,通过获取当前标准化特征值和各个历史标准化特征值,再获取到各个相似度值,从而保障了获取到的相似度值的准确性。

[0098] 具体地,获取所述当前特征值对应的当前标准化特征值和各所述历史特征值对应的历史标准化特征值的步骤,包括:

[0099] 步骤f,基于预设数据处理方式对所述当前特征值进行标准化,以获取所述当前特征值对应的当前标准化特征值;

[0100] 当获取到当前特征值后,采用预设数据处理方式对当前特征值进行标准化,以获取当前特征值对应的当前标准化特征值,如采用z-score方式进行。

[0101] 步骤h,基于所述预设数据处理方式对各所述历史特征值进行标准化,以获取各所述历史特征值对应的历史标准化特征值。

[0102] 需要说明的是,在对当前特征值进行标准化的方式和对各个历史特征值进行标准化的方式相同,也就是同样采用预设数据处理方式对各个历史特征值进行标准化,得到各个历史特征值对应的历史标准化特征值。

[0103] 在本实施例中,通过根据预设数据处理方式对当前特征值进行标准化,并对各个历史特征值进行标准化,从而保障了获取到的当前标准化特征值和历史标准化特征值的准确性。

[0104] 具体地,根据所述当前标准化特征值和各所述历史标准化特征值确定所有相似度值的步骤,包括:

[0105] 步骤k,依次遍历各所述历史标准化特征值,并计算当前遍历的历史标准化特征值和所述当前标准化特征值之间的欧式距离;

[0106] 当获取到各个历史标准化特征值后,可以依次遍历各个历史标准化特征值,并计算当前遍历的历史标准化特征值和当前标准化特征值之间的欧式距离。

[0107] 步骤m,基于所述欧式距离确定当前遍历的历史标准化特征值和当前标准化特征值的相似度值,直至各所述历史标准化特征值遍历完成。

[0108] 根据计算得到的欧式距离来确定当前遍历的历史标准化特征值和当前标准化特征值之间的相似度值,可以是欧式距离越小,两者之间的相似度值就越高。直至获取到所有的历史标准化特征值和当前标准化特征值的相似度值。并且对所有历史标准化特征值和打过去标准化特征值之间的相似度值计算方式相同。

[0109] 在本实施例中,通过计算当前标准化特征值和历史标准化特征值之间的欧式距离,并根据欧式距离计算相似度值,从而保障了计算得到的相似度值的准确性。

[0110] 进一步地,基于本发明资源信息推荐方法第一至第二任意一个实施例的基础上,提出本发明资源信息推荐方法第三实施例。本实施例是本发明第一实施例的步骤S10,基于特征选择在各所述历史数据中获取与用户账号具有相关性的多个历史特征值的步骤的细化,包括:

[0111] 步骤n,基于预设校验算法对各所述历史数据进行离散化,以获取各所述历史数据对应的初级特征值;

[0112] 当获取到各个历史数据后,可以通过预设校验算法(如卡方校验方法)对这些历史数据进行离散化,如对上年营业收入字段离散化后为:小于等于3000万为1,大于3000万并且小于等于1亿为2,大于1亿并且小于等于10亿为4,当对这些历史数据进行离散化后得到

的各个结果就是各个历史数据对应的初级特征值。

[0113] 其中,卡方检验是一种假设性检验的方法,它能够检验两个分类变量之间是否是独立无关的。它通过观察实际值和理论值的偏差来确定原假设是否成立。

[0114] 步骤x,基于所述预设校验算法对各所述初级特征值进行评分,并基于评分结果获取与用户账号具有相关性的多个历史特征值。

[0115] 当获取到各个初级特征值后,还可以采用过滤法中的卡方校验法(即预设校验算法)对每个初级特征值进行评分,并根据评分结果选取一定数量的评分较高的初级特征值作为与用户具有相关性的多个历史特征值。

[0116] 在本实施例中,通过根据预设校验算法对各个历史数据进行离散化得到各个初级特征值,并对这些初级特征值进行评分,基于评分结果获取历史特征值,从而保障了获取到的历史特征值的准确性。

[0117] 进一步地,计算用户账号的当前特征值与各所述历史特征值之间的相似度值的步骤之前,包括:

[0118] 步骤y,获取用户账号的当前数据,并对所述当前数据进行离散数据编码,以获取所述当前账号的当前特征值。

[0119] 当企业通过用户账号登录平台后,会自动获取用户账号的当前数据,并对这些当前数据进行离散数据编码,以得到当前账号的当前特征值。如对于像纳税人等级的有序序列{A→B→C→D}转换为有序数列如{1、2、3、4};对于行业等无序序列,采用one-hot编码,即假设有3个行业{金融业、采矿业、文化娱乐},则金融业编码为{1,0,0},采矿业编码为{0,1,0},文化娱乐编码为{0,0,1}。既会增加特征向量,得到当前账号的当前特征值。

[0120] 在本实施例中,通过对当前数据进行离散数据编码,得到当前账号的当前特征值,从而保障了获取到的当前特征值的准确性。

[0121] 本发明实施例还提供一种资源信息推荐装置,参照图3,所述资源信息推荐装置包括:

[0122] 获取模块,用于获取各资源信息的历史数据,并基于特征选择在各所述历史数据中获取与用户账号具有相关性的多个历史特征值;

[0123] 计算模块,用于计算用户账号的当前特征值与各所述历史特征值之间的相似度值,并在各所述相似度值中获取预设数量的目标相似度值;

[0124] 确定模块,用于基于各所述资源信息确定各所述目标相似度值对应的目标资源信息,并获取各所述目标资源信息中的参数信息;

[0125] 推荐模块,用于基于预设推荐算法对各所述目标相似度值和各所述参数信息进行计算,以获取各所述资源信息对应的推荐指数,并在各所述推荐指数中确定数值最高的目标推荐指数,对所述目标推荐指数对应的资源信息进行推荐。

[0126] 可选地,所述推荐模块,还用于:

[0127] 依次遍历各所述目标相似度值,并确定当前遍历的当前目标相似度值对应的参数信息;

[0128] 根据预设推荐算法对所述当前目标相似度值和所述当前目标相似度值对应的参数信息进行计算,以获取所述当前目标相似度值对应的资源信息的推荐指数,直至各所述目标相似度值遍历完成。

- [0129] 可选地,所述计算模块,还用于:
- [0130] 获取所述当前特征值对应的当前标准化特征值和各所述历史特征值对应的历史标准化特征值;
- [0131] 根据所述当前标准化特征值和各所述历史标准化特征值确定所有相似度值。
- [0132] 可选地,所述计算模块,还用于:
- [0133] 基于预设数据处理方式对所述当前特征值进行标准化,以获取所述当前特征值对应的当前标准化特征值;
- [0134] 基于所述预设数据处理方式对各所述历史特征值进行标准化,以获取各所述历史特征值对应的历史标准化特征值。
- [0135] 可选地,所述计算模块,还用于:
- [0136] 依次遍历各所述历史标准化特征值,并计算当前遍历的历史标准化特征值和所述当前标准化特征值之间的欧式距离;
- [0137] 基于所述欧式距离确定当前遍历的历史标准化特征值和当前标准化特征值的相似度值,直至各所述历史标准化特征值遍历完成。
- [0138] 可选地,所述获取模块,还用于:
- [0139] 基于预设校验算法对各所述历史数据进行离散化,以获取各所述历史数据对应的初级特征值;
- [0140] 基于所述预设校验算法对各所述初级特征值进行评分,并基于评分结果获取与用户账号具有相关性的多个历史特征值。
- [0141] 所述资源信息推荐装置,还包括:
- [0142] 获取用户账号的当前数据,并对所述当前数据进行离散数据编码,以获取所述当前账号的当前特征值。
- [0143] 上述各程序模块所执行的方法可参照本发明资源信息推荐方法各个实施例,此处不再赘述。
- [0144] 本发明还提供一种计算机存储介质。
- [0145] 本发明计算机存储介质上存储有资源信息推荐程序,所述资源信息推荐程序被处理器执行时实现如上所述的资源信息推荐方法的步骤。
- [0146] 其中,在所述处理器上运行的资源信息推荐程序被执行时所实现的方法可参照本发明资源信息推荐方法各个实施例,此处不再赘述。
- [0147] 需要说明的是,在本文中,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者系统不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者系统所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括该要素的过程、方法、物品或者系统中还存在另外的相同要素。
- [0148] 上述本发明实施例序号仅仅为了描述,不代表实施例的优劣。
- [0149] 通过以上的实施方式的描述,本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现,当然也可以通过硬件,但很多情况下前者是更佳的实施方式。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品存储在如上所述的一个

存储介质(如ROM/RAM、磁碟、光盘)中,包括若干指令用以使得一台终端设备(可以是手机,计算机,服务器,空调器,或者网络设备等)执行本发明各个实施例所述的方法。

[0150] 以上仅为本发明的优选实施例,并非因此限制本发明的专利范围,凡是利用本发明说明书及附图内容所作的等效结构或等效流程变换,或直接或间接运用在其他相关的技术领域,均同理包括在本发明的专利保护范围内。

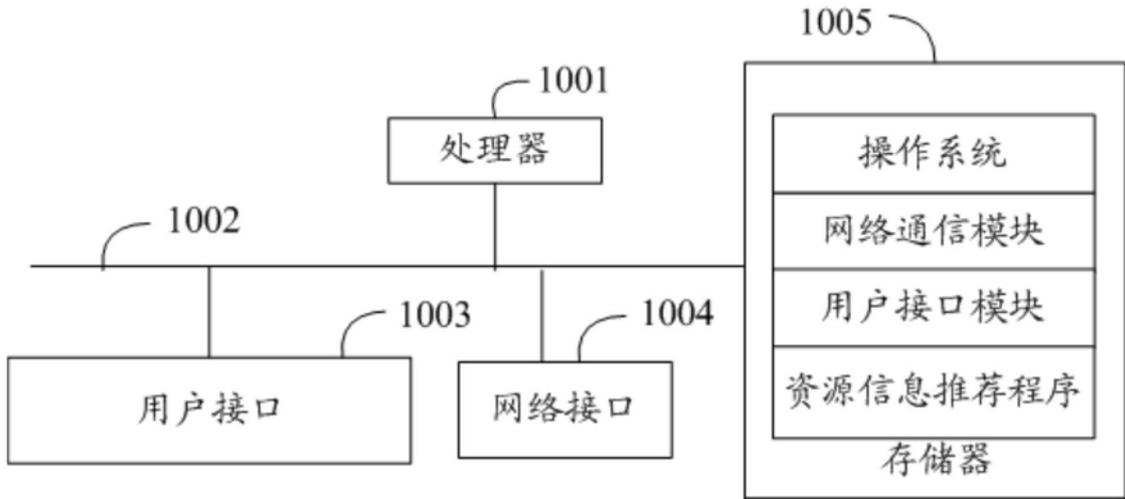


图1

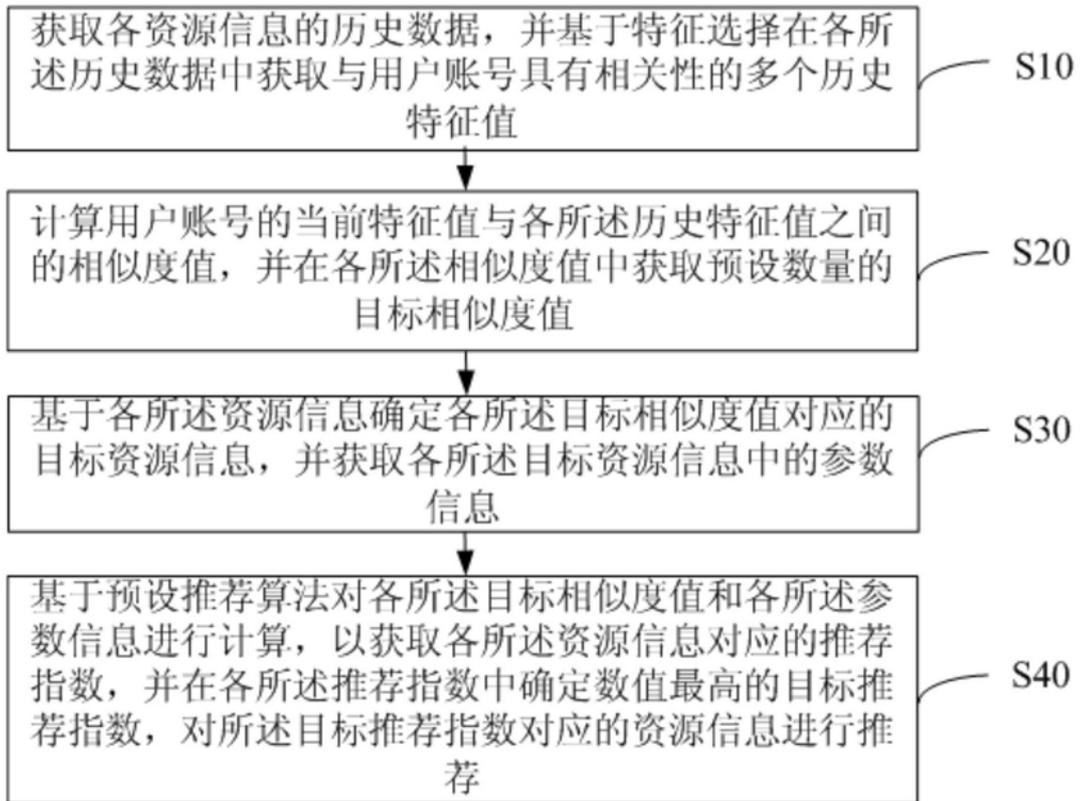


图2



图3

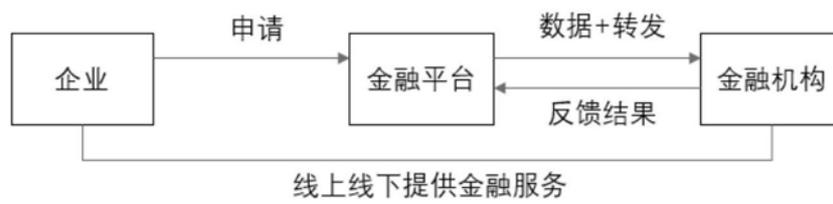


图4

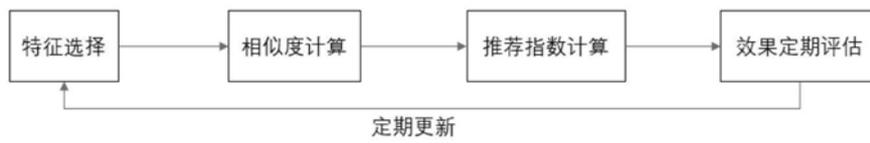


图5