



(12) 发明专利

(10) 授权公告号 CN 112269907 B

(45) 授权公告日 2021.09.03

(21) 申请号 202011201941.3

(22) 申请日 2020.11.02

(65) 同一申请的已公布的文献号
申请公布号 CN 112269907 A

(43) 申请公布日 2021.01.26

(73) 专利权人 山东万里红信息技术有限公司
地址 250000 山东省济南市槐荫区西客站
东广场北绿地缤纷城综合楼1-1208室

(72) 发明人 刘瑞景 罗远哲 李玉琼 薛瑞亭
任光远 刘志明 吕雪萍 冯建业
陈思杰

(74) 专利代理机构 济南文衡创服知识产权代理
事务所(普通合伙) 37323
代理人 刘真

(51) Int. Cl.
G06F 16/906 (2019.01)
G06N 3/08 (2006.01)
G06Q 10/06 (2012.01)
G16Y 40/50 (2020.01)
H04L 12/24 (2006.01)
H04L 29/08 (2006.01)

(56) 对比文件

CN 109346141 A, 2019.02.15
CN 104239970 A, 2014.12.24
CN 109918505 A, 2019.06.21
CN 110321521 A, 2019.10.11
CN 109738438 A, 2019.05.10
CN 107817404 A, 2018.03.20
CN 111083172 A, 2020.04.28
US 10169683 B2, 2019.01.01
WO 2015105239 A1, 2015.07.16
蒋毅等.基于数据挖掘的现场作业风险态势评估方法.《计算机与现代化》.2020,(第296期),78-84.
范学满等.基于AdaBoost的潜射防鱼雷诱饵干扰效果预测研究.《指挥控制与仿真》.2019,第41卷(第3期),52-56.

Qazi Sami ul Haq 等.Neural network based adaboosting approach for hyperspectral data classification.《Proceedings of 2011 International Conference on Computer Science and Network Technology》.2011,241-245.

审查员 李若晨

权利要求书2页 说明书7页 附图1页

(54) 发明名称

一种物联网健康大数据的处理方法

(57) 摘要

本发明公开了物联网健康大数据的处理方法,包括以下步骤,首先确定物联网健康医疗大数据的安全数据格式要求,接着对采集到的数据由分级引擎模块进行风险分级处理,然后采用神经网络特征提取并学习的分类算法同时引入 AdaBoost集成学习方法神经网络特征提取到的物联网健康大数据态势点集进行迭代优化,生成物联网健康大数据网络态势模型。本数据模型更加精确,大大促进了物联网医疗健康大数据态势预测、风险排查、治理,隐私数据安全采用风险分级管控模式可以更好的对隐私数据进行等级保

护,可以辅助安全相关部门及早做好应对措施。



CN 112269907 B

1. 一种物联网健康大数据的处理方法,其特征在于,包括以下步骤:

S1、确定物联网健康医疗大数据的安全数据格式要求,采用数据采集模块收集物联网健康医疗大数据,采用统一格式对通信通道中被篡改以及丢失的异常数据进行数据采集和预处理后统一编码;

S2、对步骤S1采集到的数据由分级引擎模块进行风险分级处理,分级方法为对在节点采集到的数据进行特征提取,并利用不同的特征子集对物联网健康大数据样本进行分类;

S3、采用神经网络特征提取并学习的分类算法同时引入AdaBoost集成学习方法对神经网络特征提取到的物联网健康大数据态势点集进行迭代优化,生成物联网健康大数据网络态势模型;

在所述步骤S3中,分类算法采用神经网络对物联网健康大数据进行特征学习的方法,建立物联网健康大数据网络态势初始模型,再采用AdaBoost集成学习方法对初始模型进行迭代优化,并对异常数据进行排查,形成物联网健康大数据网络态势模型;机器学习采用神经网络对物联网健康大数据进行特征学习的过程由正向传播和误差反向传播两个环节组成,所述正向传播是利用t时期的数据特征值与t-1时期的数据特征值的差值和t时期的数据特征值与t-1时期的时间间隔组成的向量作为输入,经网络连接权值和偏差的求和计算传递到隐含层,并通过传递函数计算得到隐含层的输出,然后传送到输出层;所述反向传播是将误差信号按原来连接通路反向计算,调整各层神经元之间的连接权值和偏差以提高准确性。

2. 根据权利要求1所述物联网健康大数据的处理方法,其特征在于,在所述步骤S2,通过安全风险分级管理机制和风险排查治理机制相结合的模式对物联网医疗健康大数据网络安全领域实现高精度的等级保护。

3. 根据权利要求2所述物联网健康大数据的处理方法,其特征在于,在所述步骤S2包括如下具体步骤:

S21、对物联网健康大数据集采用规范化矢量表示,即, $V(d) = \{d_1, w_1(d), d_2, w_2(d), \dots, d_n, w_n(d)\}$,其中, d_n 为物联网健康大数据集D中的一条数据, $w_n(d)$ 为 d_n 在D中的权值,权值累计为1;

S22、对物联网健康大数据集进行评价函数构建。

4. 根据权利要求3所述物联网健康大数据的处理方法,其特征在于,在所述步骤S22包括如下具体步骤:

S221、假设有n个评价对象,m个数据指标;

S222、建立评价对象因素指标集 $U = \{u_1, u_2, \dots, u_n\}$;

S223、建立评价集 $V = \{v_1, v_2, \dots, v_m\}$;

S224、建立单因素评价,即建立一个从U到F(V)的映射;

$\gamma : \rightarrow F(V)$

$$u_i \rightarrow \gamma(u_i) = \frac{r_{i1}}{v_1} + \dots + \frac{r_{im}}{v_m}, 0 \leq r_{ij} \leq 1, \sum_{j=1}^m r_{ij} = 1, 1 \leq i \leq n, 1 \leq j \leq m$$

由 γ 可以诱导出基本关系,得到相关矩阵:

$$R = \begin{pmatrix} r_{11} & \cdots & r_{1m} \\ \vdots & & \vdots \\ r_{n1} & \cdots & r_{nm} \end{pmatrix}, R \text{ 为单因素评价矩阵;}$$

S225、确定评价函数，

设 $A = (a_1, \dots, a_m)$ 表示各因素权重，则评价函数为：

$$B = A \circ R = (a_1, \dots, a_m) \circ \begin{pmatrix} r_{11} & \cdots & r_{1m} \\ \vdots & & \vdots \\ r_{n1} & \cdots & r_{nm} \end{pmatrix} = (b_1, \dots, b_m)$$

采用算子 $(\cdot, +)$ 计算， $b_j = \sum_{i=1}^n a_i r_{ij}, (j=1, 2, \dots, m), 0 \leq a_i \leq 1$ 且 $\sum_{i=1}^n a_i = 1$ 对向量 (b_1, \dots, b_m)

去模糊化， m 个评价依次记为 $1, 2, \dots, m$ ，整体评价分数为：

$$y = b_1 + 2b_2 + \dots + mb_m = \sum_{j=1}^m \sum_{i=1}^n j a_i r_{ij}$$

5. 根据权利要求1所述物联网健康大数据的处理方法，其特征在于，在所述步骤S3包括如下具体步骤：

S31、假设 m 个初始学习的特征样本集为： $s = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$ ，其中 x_m 为特征样本， y_m 表示不同的风险问题分类，各个样本初始权重 d_1, d_2, \dots, d_m 均设置为 $\frac{1}{m}$ ；算法的最大迭代次数为 T ，初始迭代次数为 1 ；

S32、针对 m 个训练集，通过AdaBoost集成学习方法对神经网络的连接点权重值进行优化，得到最优的连接点的权重值；

S33、利用优化后的神经网络对 m 个训练集进行训练，获得第 t 次的物联网健康大数据网络态势模型 h_t ；

S34、计算并保存第 t 次的物联网健康大数据网络态势模型 h_t 的权重 ω_t ，根据物联网健康大数据网络态势模型 h_t ，对 m 个训练集的预测误差绝对值和小于设定值或达到最大迭代次数，得到最终的物联网健康大数据网络态势模型： $h = \sum_{i=1}^T \omega_i h_i$ 。

6. 根据权利要求1所述物联网健康大数据的处理方法，其特征在于，在所述步骤S1中，所述数据采集模块包括数据过滤模块和数据预处理模块，所述数据采集模块用于筛除物联网健康医疗大数据中的错误数据和无效数据。

7. 根据权利要求1所述物联网健康大数据的处理方法，其特征在于，在所述步骤S2中，采用分级引擎模块对物联网健康医疗大数据的特征进行提取、分类和分级，所述分级引擎模块包括特征提取模块和分级模型模块。

一种物联网健康大数据的处理方法

技术领域

[0001] 本发明涉及物联网健康大数据安全领域,具体地,涉及一种物联网健康大数据的处理方法。

背景技术

[0002] 近年来,大数据在健康领域的不断应用和发展,已为广大群众提供了优质、便捷的服务,也为健康领域的科技发展带来动力。目前健康医疗大数据平台汇集了医疗服务、疾病防治、妇幼健康、人口家庭、电子健康档案、家医签约服务、行业监管等方面的数据资源,主要包括门诊住院信息,住院病案信息,人口个案信息,电子健康档案,电子健康卡注册使用信息,数据量都达到了万亿条以上,健康医疗大数据资源体系已初具规模。

[0003] 目前的医疗健康大数据安全方面采用加密、数据库审计、防篡改等技术,网络安全方面通过等级保护实现基本安全,但是隐私数据安全还是没有有一个妥善的机制。

[0004] 采用神经网络特征提取并学习的分类算法同时引入AdaBoost集成学习方法神经网络特征提取到的物联网健康大数据态势点集进行迭代优化,生成物联网健康大数据网络态势模型,并通过安全风险分级管理机制和风险排查治理机制相结合的模式对物联网医疗健康大数据网络安全领域可实现高精度的等级保护,隐私数据安全也能及早预测和及时治理。

发明内容

[0005] 为了解决如何对物联网健康大数据多种数据样本进行安全风险、风险态势预测、排查、治理的问题,本发明提供一种物联网健康大数据的处理方法。

[0006] 根据本发明实施例的物联网健康大数据的处理方法,包括如下步骤:

[0007] S1、确定物联网健康医疗大数据的安全数据格式要求,采用数据采集模块收集物联网健康医疗大数据,采用统一格式对通信通道中被篡改以及丢失的数据等异常数据进行数据采集和预处理后统一编码;

[0008] S2、对步骤S1采集到的数据由分级引擎模块进行风险分级处理,分级方法为对在节点采集到的数据进行特征提取,并利用不同的特征子集对物联网健康大数据样本进行分类;

[0009] S3、采用神经网络特征提取并学习的分类算法同时引入AdaBoost集成学习方法神经网络特征提取到的物联网健康大数据态势点集进行迭代优化,生成物联网健康大数据网络态势模型。

[0010] 根据本发明的一些实施例,在所述步骤S2,通过安全风险分级管理机制和风险排查治理机制相结合的模式对物联网医疗健康大数据网络安全领域实现高精度的等级保护。

[0011] 根据本发明的一些实施例,在所述步骤S2包括如下具体步骤:

[0012] S21、对物联网健康大数据集采用规范化矢量表示,即, $V(d) = \{d_1, w_1(d), d_2, w_2(d), \dots, d_n, w_n(d)\}$, 其中, d_n 为物联网健康大数据集 D 中的一条数据, $w_n(d)$ 为 d_n 在 D 中的权

值,权值累计为1;

[0013] S22、对物联网健康大数据集进行评价函数构建。

[0014] 根据本发明的一些实施例,在所述步骤S22包括如下具体步骤:

[0015] S221、假设有n个评价对象,m个数据指标;

[0016] S222、建立评价对象因素指标集 $U = \{u_1, u_2, \dots, u_n\}$;

[0017] S223、建立评价集 $V = \{v_1, v_2, \dots, v_m\}$;

[0018] S224、建立单因素评价,即建立一个从U到F(V)的映射;

[0019] $\gamma : \rightarrow F(V)$

[0020] $u_i \rightarrow \gamma(u_i) = \frac{r_{i1}}{v_1} + \dots + \frac{r_{im}}{v_m}, 0 \leq r_{ij} \leq 1, \sum_{j=1}^m r_{ij} = 1, 1 \leq i \leq n, 1 \leq j \leq m$

[0021] 由 γ 可以诱导出基本关系,得到相关矩阵:

[0022] $R = \begin{pmatrix} r_{11} & \dots & r_{1m} \\ \vdots & & \vdots \\ r_{n1} & \dots & r_{nm} \end{pmatrix}$, R为单因素评价矩阵;

[0023] S225、确定评价函数,

[0024] 设 $A = (a_1, \dots, a_m)$ 表示各因素权重,则评价函数为:

[0025] $B = A \circ R = (a_1, \dots, a_m) \circ \begin{pmatrix} r_{11} & \dots & r_{1m} \\ \vdots & & \vdots \\ r_{n1} & \dots & r_{nm} \end{pmatrix} = (b_1, \dots, b_m)$,

[0026] 采用算子 $(\cdot, +)$ 计算, $b_j = \sum_{i=1}^n a_i r_{ij}, (j=1, 2, \dots, m), 0 \leq a_i \leq 1$ 且 $\sum_{i=1}^n a_i = 1$, 对向量 (b_1, \dots, b_m) 去模糊化,m个评价依次记为1, 2, \dots , m, 整体评价分数为:

[0027] $y = b_1 + 2b_2 + \dots + mb_m = \sum_{j=1}^m \sum_{i=1}^n j a_i r_{ij}$ 。

[0028] 根据本发明的一些实施例,在所述步骤S3中,分类算法采用神经网络对物联网健康大数据进行特征学习的方法,建立物联网健康大数据网络态势初始模型,再采用AdaBoost集成学习方法对初始模型进行迭代优化,并对异常数据进行排查,形成物联网健康大数据网络态势模型。

[0029] 根据本发明的一些实施例,在所述步骤S3中,所述机器学习采用神经网络对物联网健康大数据进行特征学习的过程由正向传播和误差反向传播两个环节组成,所述正向传播是利用t时期的数据特征值与t-1时期的数据特征值的差值(变化率)和t时期的数据特征值与t-1时期的时间间隔组成的向量作为输入,经网络连接权值和偏差的求和计算传递到隐含层,并通过传递函数计算得到隐含层的输出,然后传送到输出层;所述反向传播是将误差信号按原来连接通路反向计算,调整各层神经元之间的连接权值和偏差以提高准确性。

[0030] 根据本发明的一些实施例,在所述步骤S3包括如下具体步骤:

[0031] S31、假设m个初始学习的特征样本集为: $s = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$,其中 x_m

为特征样本, y_m 表示不同的风险问题分类, 各个样本初始权重 d_1, d_2, \dots, d_m 均设置为 $\frac{1}{m}$; 算法的最大迭代次数为 T , 初始迭代次数为 1;

[0032] S32、针对 m 个训练集, 通过 AdaBoost 集成学习方法对神经网络的连接点权重值进行优化, 得到最优的连接点的权重值;

[0033] S33、利用优化后的神经网络对 m 个训练集进行训练, 获得第 t 次的物联网健康大数据网络态势模型 h_t ;

[0034] S34、计算并保存第 t 次的物联网健康大数据网络态势模型 h_t 的权重 ω_t , 根据物联网健康大数据网络态势模型 h_t , 对 m 个训练集的预测误差绝对值和小于设定值或达到最大

迭代次数, 得到最终的物联网健康大数据网络态势模型: $h = \sum_{i=1}^T \omega_i h_i$ 。

[0035] 根据本发明的一些实施例, 在所述步骤 S1 中, 所述数据采集模块包括数据过滤模块和数据预处理模块, 所述数据采集模块用于筛除物联网健康医疗大数据中的错误数据和无效数据。

[0036] 根据本发明的一些实施例, 在所述步骤 S2 中, 采用分级引擎模块对物联网健康医疗大数据的特征进行提取、分类和分级, 所述分级引擎模块包括特征提取模块和分级模型模块。

[0037] 根据本发明实施例的物联网健康大数据的处理方法, 数据模型更加精确, 运用神经网络模式处理数据, 大大促进了物联网医疗健康大数据态势预测、风险排查、治理, 隐私数据安全采用风险分级管控模式可以更好的对隐私数据进行等级保护, 可以辅助安全相关部门及早做好应对措施。

附图说明

[0038] 图1为本发明实施例的物联网健康大数据的处理方法的流程图。

[0039] 图2为本发明实施例的物联网健康大数据的处理方法的结构示意图。

具体实施方式

[0040] 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做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。

[0041] 下面参照图1描述根据本发明实施例的物联网健康大数据的处理方法。

[0042] 根据本发明实施例的物联网健康大数据的处理方法, 包括如下步骤:

[0043] S1、确定物联网健康医疗大数据的安全数据格式要求, 采用数据采集模块收集物联网健康医疗大数据, 采用统一格式对通信通道中被篡改以及丢失的数据等异常数据进行数据采集和预处理后统一编码; 其中, 物联网健康医疗大数据包括: 患者的住院信息、住院病案信息、人口个案信息、电子健康档案和电子健康卡注册使用信息等, 这些信息都记录在各大医院、社区等的服务器数据库, 网络协议/网络IP地址以及连接方式等也都不尽相同, 健康数据的显示/存储方式主要有电子文档、图像、纸质文件、不同尺寸/材质的影像图片

等,针对不同的数据类型进行风险分级,产生不同的分类,并定期进行风险排查的方式进行漏洞扫描,降低预测误差,保证数据安全可靠。

[0044] S2、对步骤S1采集到的数据由分级引擎模块进行风险分级处理,分级方法为对在节点采集到的数据进行特征提取,并利用不同的特征子集对物联网健康大数据样本进行分类;

[0045] S3、采用神经网络特征提取并学习的分类算法同时引入AdaBoost集成学习方法神经网络特征提取到的物联网健康大数据态势点集进行迭代优化,生成物联网健康大数据网络态势模型。

[0046] 根据本发明的一些实施例,在所述步骤S2,通过安全风险分级管理机制和风险排查治理机制相结合的模式对物联网医疗健康大数据网络安全领域实现高精度的等级保护,由此可以提升数据的保护效果,防止数据被篡改、丢失或是泄漏。

[0047] 根据本发明的一些实施例,在所述步骤S2包括如下具体步骤:

[0048] S21、对物联网健康大数据集采用规范化矢量表示,即, $V(d) = \{d_1, w_1(d), d_2, w_2(d), \dots, d_n, w_n(d)\}$,其中, d_n 为物联网健康大数据集D中的一条数据, $w_n(d)$ 为 d_n 在D中的权值,权值累计为1;

[0049] S22、对物联网健康大数据集进行评价函数构建。

[0050] 根据本发明的一些实施例,在所述步骤S22包括如下具体步骤:

[0051] S221、假设有n个评价对象,m个数据指标;

[0052] S222、建立评价对象因素指标集 $U = \{u_1, u_2, \dots, u_n\}$;

[0053] S223、建立评价集 $V = \{v_1, v_2, \dots, v_m\}$;

[0054] S224、建立单因素评价,即建立一个从U到F(V)的映射;

[0055] $\gamma : \rightarrow F(V)$

[0056] $u_i \rightarrow \gamma(u_i) = \frac{r_{i1}}{v_1} + \dots + \frac{r_{im}}{v_m}, 0 \leq r_{ij} \leq 1, \sum_{j=1}^m r_{ij} = 1, 1 \leq i \leq n, 1 \leq j \leq m$

[0057] 由 γ 可以诱导出基本关系,得到相关矩阵:

[0058] $R = \begin{pmatrix} r_{11} & \dots & r_{1m} \\ \vdots & & \vdots \\ r_{n1} & \dots & r_{nm} \end{pmatrix}$, R为单因素评价矩阵;

[0059] S225、确定评价函数,

[0060] 设 $A = (a_1, \dots, a_m)$ 表示各因素权重,则评价函数为:

[0061] $B = A \circ R = (a_1, \dots, a_m) \circ \begin{pmatrix} r_{11} & \dots & r_{1m} \\ \vdots & & \vdots \\ r_{n1} & \dots & r_{nm} \end{pmatrix} = (b_1, \dots, b_m)$,

[0062] 采用算子 $(\cdot, +)$ 计算, $b_j = \sum_{i=1}^n a_i r_{ij}, (j=1, 2, \dots, m), 0 \leq a_i \leq 1$ 且 $\sum_{i=1}^n a_i = 1$,对向量 (b_1, \dots, b_m) 去模糊化,m个评价依次记为1, 2, \dots , m,整体评价分数为:

$$[0063] \quad y = b_1 + 2b_2 + \dots + mb_m = \sum_{j=1}^m \sum_{i=1}^n ja_i r_{ij}。$$

[0064] 根据本发明的一些实施例,在所述步骤S3中,分类算法采用神经网络对物联网健康大数据进行特征学习的方法,建立物联网健康大数据网络态势初始模型,再采用AdaBoost集成学习方法对初始模型进行迭代优化,并对异常数据进行排查,形成物联网健康大数据网络态势模型,神经网络特征学习和AdaBoost集成学习的相结合,可以提升态势模型的学习能力和效果,进而提升数据模型的准确度和模拟的真实性。

[0065] 根据本发明的一些实施例,在所述步骤S3中,所述机器学习采用神经网络对物联网健康大数据进行特征学习的过程由正向传播和误差反向传播两个环节组成,所述正向传播是利用t时期的数据特征值与t-1时期的数据特征值的差值(变化率)和t时期的数据特征值与t-1时期的时间间隔组成的向量作为输入,经网络连接权值和偏差的求和计算传递到隐含层,并通过传递函数计算得到隐含层的输出,然后传送到输出层;所述反向传播是将误差信号按原来连接通路反向计算,调整各层神经元之间的连接权值和偏差以提高准确性。

[0066] 根据本发明的一些实施例,在所述步骤S3包括如下具体步骤:

[0067] S31、假设m个初始学习的特征样本集为: $s = \{(x_1, y_1), (x_2, x_2), \dots, (x_m, y_m)\}$,其中 x_m 为特征样本, y_m 表示不同的风险问题分类,各个样本初始权重 d_1, d_2, \dots, d_m 均设置为 $\frac{1}{m}$;算法

的最大迭代次数为T,初始迭代次数为1;

[0068] S32、针对m个训练集,通过AdaBoost集成学习方法对神经网络的连接点权重值进行优化,得到最优的连接点的权重值;

[0069] S33、利用优化后的神经网络对m个训练集进行训练,获得第t次的物联网健康大数据网络态势模型 h_t ;

[0070] S34、计算并保存第t次的物联网健康大数据网络态势模型 h_t 的权重 ω_t ,根据物联网健康大数据网络态势模型 h_t ,对m个训练集的预测误差绝对值和小于设定值或达到最大

迭代次数,得到最终的物联网健康大数据网络态势模型: $h = \sum_{i=1}^T \omega_i h_i$ 。

[0071] 根据本发明的一些实施例,在所述步骤S1中,不同数据类型的实时网络数据流(包括MAC层、网络层、传输层、应用层),由数据采集模块进行收集进入,所述数据采集模块包括数据过滤模块和数据预处理模块,所述数据采集模块用于筛除物联网健康医疗大数据中的错误数据和无效数据。

[0072] 根据本发明的一些实施例,在所述步骤S2中,采用分级引擎模块对物联网健康医疗大数据的特征进行提取、分类和分级,所述分级引擎模块包括特征提取模块和分级模型模块。

[0073] 具体而言,对步骤S1采集到的数据由分级引擎模块进行风险分级处理,分级引擎模块由两个小模块组成,分别是特征提取模块和分级模型模块。

[0074] 根据本发明的一些实施例,在所述步骤S3中,评价函数的权重来自于外部攻击,权重最高;通信节点的数据量大小,权重按大小分配;通信交叉点越多,权重越大。

[0075] 方法是:对在节点采集到的数据进行特征提取,并利用不同的特征子集对物联网

健康大数据样本进行分类,具体包括如下步骤:

[0076] 对物联网健康大数据集采用规范化矢量表示,

[0077] 即 $V(d) = \{d_1, w_1(d), d_2, w_2(d), \dots, d_n, w_n(d)\}$, 其中 d_n 为物联网健康大数据集 D 中的一条数据, $w_n(d)$ 为 d_n 在 D 中的权值,

[0078] 权重的评价函数如下:

[0079] (1) 外部攻击, 权重最高;

[0080] (2) 通信节点的数据量大小, 权重按大小分配;

[0081] (3) 通信交叉点越多, 权重越大。

[0082] 在所述步骤 S2 中, 对物联网健康大数据集进行评价函数构建, 评价函数过程如下:

[0083] S221、假设有 n 个评价对象, m 个数据指标;

[0084] S222、建立评价对象因素指标集 $U = \{u_1, u_2, \dots, u_n\}$;

[0085] S223、建立评价集 $V = \{v_1, v_2, \dots, v_m\}$;

[0086] S24、建立单因素评价, 即建立一个从 U 到 $F(V)$ 的映射;

[0087] $\gamma: U \rightarrow F(V)$

[0088] $u_i \rightarrow \gamma(u_i) = \frac{r_{i1}}{v_1} + \dots + \frac{r_{im}}{v_m}, 0 \leq r_{ij} \leq 1, \sum_{j=1}^m r_{ij} = 1, 1 \leq i \leq n, 1 \leq j \leq m$

[0089] 由 γ 可以诱导出基本关系, 得到相关矩阵:

[0090] $R = \begin{pmatrix} r_{11} & \dots & r_{1m} \\ \vdots & & \vdots \\ r_{n1} & \dots & r_{nm} \end{pmatrix}$ R 为单因素评价矩阵;

[0091] S225、确定评价函数,

[0092] 设 $A = (a_1, \dots, a_m)$ 表示各因素权重, 则评价函数为:

[0093] $B = A \circ R = (a_1, \dots, a_m) \circ \begin{pmatrix} r_{11} & \dots & r_{1m} \\ \vdots & & \vdots \\ r_{n1} & \dots & r_{nm} \end{pmatrix} = (b_1, \dots, b_m)$,

[0094] 采用算子 $(\cdot, +)$ 计算, $b_j = \sum_{i=1}^n a_i r_{ij}, (j=1, 2, \dots, m), 0 \leq a_i \leq 1$ 且 $\sum_{i=1}^n a_i = 1$, 对向量

(b_1, \dots, b_m) 去模糊化, m 个评价依次记为 $1, 2, \dots, m$, 整体评价分数为:

[0095] $y = b_1 + 2b_2 + \dots + mb_m = \sum_{j=1}^m \sum_{i=1}^n j a_i r_{ij}$ 。

[0096] 根据本发明的一些实施例, 在所述步骤 S3 中, 分类算法采用神经网络对物联网健康大数据进行特征学习的方法, 建立物联网健康大数据网络态势初始模型, 再采用 AdaBoost 集成学习方法, 对这个初始模型进行迭代优化, 也就是进行异常数据的排查, 形成物联网健康大数据网络态势模型, 其中, 机器学习主要包括以下步骤: 首先确定网络层数, 接着选取节点数, 接着选择隐含神经元数目, 最后设计算法。

[0097] 所述机器学习采用神经网络对物联网健康大数据进行特征学习的过程由正向传播和误差反向传播两个环节组成, 所述正向传播是利用 t 时期的数据特征值与 $t-1$ 时期的数

据特征值的差值(也就是变化率)和t时期的数据特征值与t-1时期的时间间隔组成的向量作为输入,经网络连接权值和偏差的求和和计算传递到隐含层,并通过传递函数计算得到隐含层的输出,然后传送到输出层;所述反向传播是将误差信号按原来连接通路反向计算,调整各层神经元之间的连接权值和偏差以提高准确性,体步骤如下:

[0098] S31、假设m个初始学习的特征样本集为: $s = \{(x_1, y_1), (x_2, x_2), \dots (x_m, y_m)\}$,其中 x_m 为特征样本, y_m 表示不同的风险问题分类,各个样本初始权重 d_1, d_2, \dots, d_m 均设置为 $\frac{1}{m}$;算法的最大迭代次数为T,初始迭代次数为1;

[0099] S32、针对m个训练集,通过AdaBoost集成学习方法对神经网络的连接点权重值进行优化,得到最优的连接点的权重值;

[0100] S33、利用优化后的神经网络对m个训练集进行训练,获得第t次的物联网健康大数据网络态势模型 h_t ;

[0101] S34、计算并保存第t次的物联网健康大数据网络态势模型 h_t 的权重 ω_t ,根据物联网健康大数据网络态势模型 h_t ,对m个训练集的预测误差绝对值和小于设定值或达到最大

迭代次数,得到最终的物联网健康大数据网络态势模型: $h = \sum_{i=1}^T \omega_i h_i$ 。

[0102] 以此,可以根据现有数据以及新增数据进行风险排查迭代模拟,提前将影响到物联网医疗健康大数据安全的各项安全隐患设置防护措施,降低风险。

[0103] 根据本发明实施例的物联网健康大数据的处理方法,数据模型更加精确,运用神经网络模式处理数据,大大促进了物联网医疗健康大数据态势预测、风险排查、治理,隐私数据安全采用风险分级管控模式可以更好的对隐私数据进行等级保护,可以辅助安全相关部门及早做好应对措施。

[0104] 对于本领域技术人员而言,显然本发明不限于上述示范性实施例的细节,而且在不背离本发明的精神或基本特征的情况下,能够以其他的具体形式实现本发明。因此,无论从哪一点来看,均应将实施例看作是示范性的,而且是非限制性的,本发明的范围由所附权利要求而不是上述说明限定,因此旨在将落在权利要求的等同要件的含义和范围内的所有变化囊括在本发明内。不应将权利要求中的任何附图标记视为限制所涉及的权利要求。

[0105] 此外,应当理解,虽然本说明书按照实施方式加以描述,但并非每个实施方式仅包含一个独立的技术方案,说明书的这种叙述方式仅仅是为清楚起见,本领域技术人员应当将说明书作为一个整体,各实施例中的技术方案也可以经适当组合,形成本领域技术人员可以理解的其他实施方式。

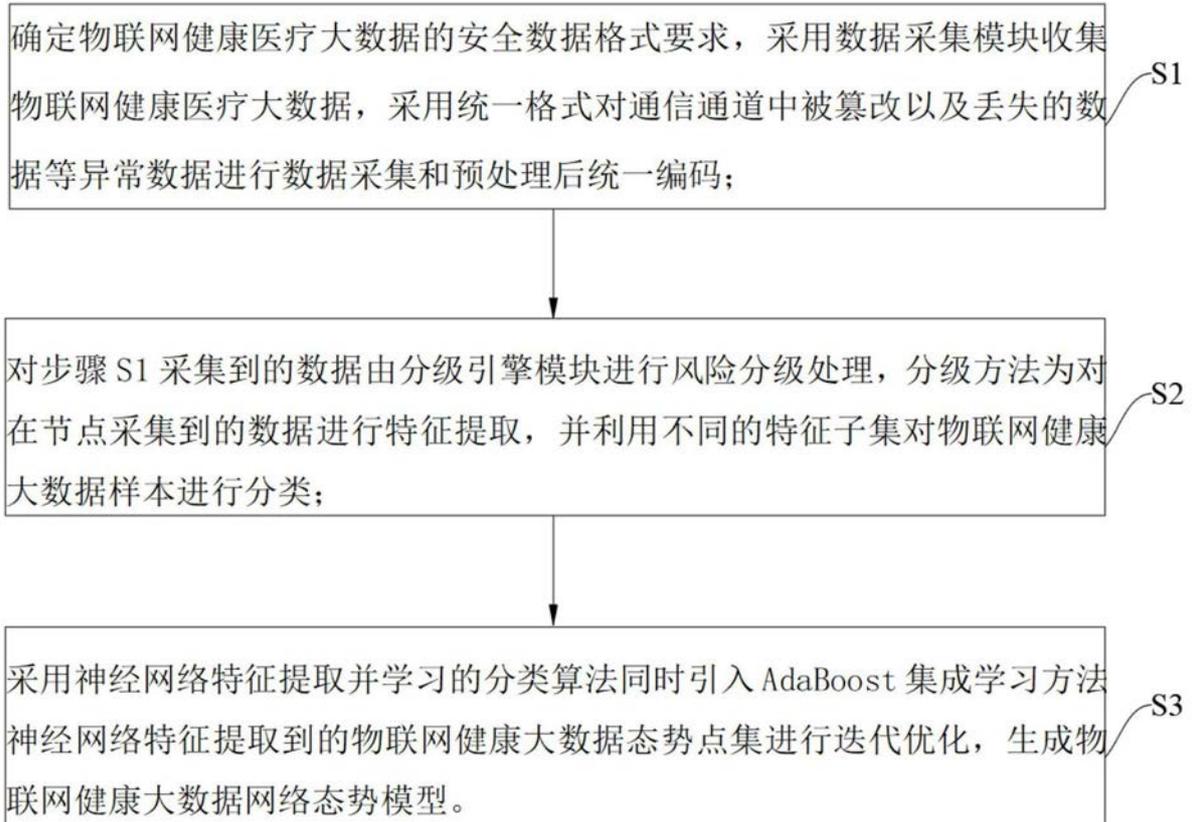


图1

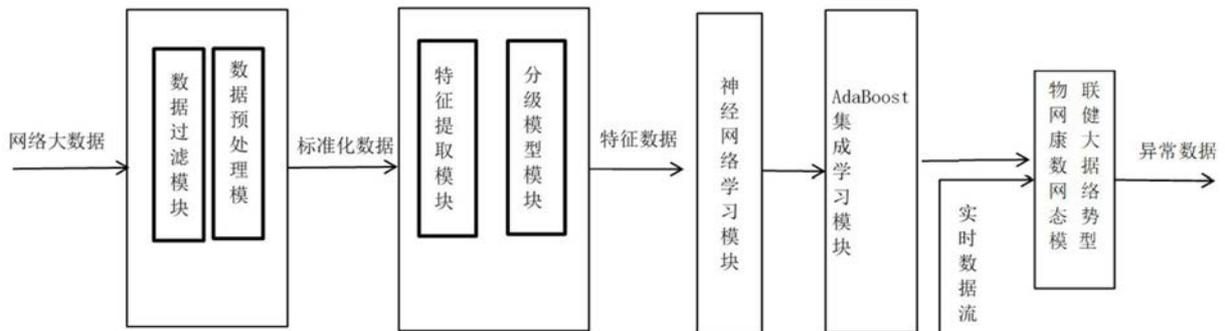


图2