



## (12)发明专利申请

(10)申请公布号 CN 107153775 A

(43)申请公布日 2017.09.12

(21)申请号 201710444355.3

(22)申请日 2017.06.13

(71)申请人 京东方科技集团股份有限公司

地址 100015 北京市朝阳区酒仙桥路10号

(72)发明人 张超 张振中

(74)专利代理机构 北京中博世达专利商标代理有限公司 11274

代理人 申健

(51)Int.Cl.

G06F 19/00(2011.01)

权利要求书3页 说明书13页 附图2页

### (54)发明名称

一种智能分诊方法及装置

### (57)摘要

本发明实施例提供一种智能分诊方法及装置,涉及医疗技术领域,用以对患者实现智能分诊来缓解医院的分诊压力。该方法包括:获取患者的病情特征数据;根据患者的病情特征数据、以及病例数据库中每个病种下各病例的病情特征数据,确定患者患病例数据库中每个病种的可能性;根据患者患病例数据库中每个病种的可能性,输出患者的分诊结果。本发明应用于医疗分诊。



1. 一种智能分诊方法,其特征在于,所述方法包括:

获取患者的病情特征数据;

根据所述患者的病情特征数据、以及病例数据库中每个病种下各病例的病情特征数据,确定所述患者患所述病例数据库中每个病种的可能性;

根据所述患者患所述病例数据库中每个病种的可能性,输出所述患者的分诊结果。

2. 根据权利要求1所述的方法,其特征在于,所述病情特征数据包括病情症状信息和/或体征检测参数。

3. 根据权利要求1或2所述的方法,其特征在于,所述获取患者的病情特征数据,具体包括:

根据预设的矩阵元素的位置与矩阵元素所表示的病情特征的对应关系,将输入的患者病情特征转换成患者的病情特征分布矩阵,所述患者的病情特征分布矩阵中的每个元素用于指示所述输入的患者病情特征中是否出现该元素所在位置对应的病情特征。

4. 根据权利要求3所述的方法,其特征在于,所述根据所述患者的病情特征数据、以及病例数据库中每个病种下各病例的病情特征数据,确定所述患者患所述病例数据库中每个病种的可能性,具体包括:

将所述患者的病情特征数据代入到关系模型中,得到满足所述关系模型的所有相关系数矩阵 $X$ ;

从所述所有相关系数矩阵 $X$ 中确定出唯一的相关系数矩阵 $X_0$ ;

根据所述唯一的相关系数矩阵 $X_0$ ,确定所述病例数据库中每个病种的可能性;

其中,所述关系模型为: $h=DX$ ,所述 $h$ 为所述患者的病情特征分布矩阵,所述 $D$ 为由所述病例数据库中每个病种下各病例的病情特征分布矩阵组成的矩阵,所述 $D=[D_1, D_2, \dots, D_i, \dots, D_M]$ ,其中: $D_i=[D_{i,1}, D_{i,2}, \dots, D_{i,j}, \dots, D_{i,K}]$ ,所述 $D_{i,j}$ 为所述病例数据库中病种 $i$ 的第 $j$ 个病例的病情特征分布矩阵,所述 $K$ 用于表示所述病例数据库中病种 $i$ 包括 $K$ 个病例,所述 $M$ 用于表示病例数据库中包括 $M$ 种疾病。

5. 根据权利要求4所述的方法,其特征在于,所述方法还包括:

根据预设的矩阵元素的位置与矩阵元素所表示的病情特征的对应关系,将每个病种下各病例的病情特征转换成病例病情特征分布矩阵;所述病例的病情特征分布矩阵中的每个元素用于指示所述病例的病情特征中是否出现该元素所在位置对应的病情特征。

6. 根据权利要求4所述的方法,其特征在于,所述从所述所有相关系数矩阵 $X$ 中确定出唯一的相关系数矩阵 $X_0$ ,具体包括:

从所述所有相关系数矩阵 $X$ 确定出满足第一预设条件的相关系数矩阵 $X$ ;

在满足第一预设条件的相关系数矩阵 $X$ 中确定出满足第二预设条件的唯一的相关系数矩阵 $X_0$ ;

其中,所述第一预设条件为: $\|DX-h\|_2 \leq \epsilon$ ,所述第二预设条件为: $x^* = \arg \min \|X_0\|_1$ ,其中:所述 $\|\cdot\|_1$ 是L1范式,所述 $\|\cdot\|_2$ 是L2范式,所述 $\epsilon$ 为预设参数, $x^*$ 为目标函数。

7. 根据权利要求4所述的方法,其特征在于,所述根据所述唯一的相关系数矩阵 $X_0$ ,确定所述病例数据库中每个病种的可能性,具体包括:

从所述相关系数矩阵 $X_0$ 中确定出所述病例数据库中每个病种的相关系数矩阵 $\delta_i(X_0)$ ;

将所述病例数据库中每个病种的相关系数矩阵 $\delta_i(X_0)$ 代入到概率计算公式中,得到所

述患者患所述病例数据库中每个病种的概率；

其中,所述概率计算公式为: 
$$C_i = \frac{\|h_i\|_2^2}{\sum_{j=1}^M \|h_j\|_2^2 + \|\eta\|_2^2}$$
, 所述 $C_i$ 用于表示所述患者患所述病例

数据库中的病种 $i$ 的概率,所述 $h_i = D * \delta_i(X_0)$ , 所述 $\delta_i(X_0)$ 为所述病例数据库中病种 $i$ 的相关系数矩阵,所述 $\sum_{j=1}^M \|h_j\|_2^2$ 中的 $M$ 用于表示所述病例数据库中的 $M$ 个病种,所述 $\eta$ 为误差矩阵,所述 $\|\cdot\|_2^2$ 是L2范式的平方。

8. 根据权利要求1所述的方法,其特征在于,所述根据所述病例数据库中每个病种的可能性,输出所述患者的分诊结果,具体包括:

输出每个病种的可能性中的最大的对应的所述患者的分诊结果;

或者,将每个病种的可能性中不为零的,按照可能性大小输出所述患者的分诊结果。

9. 一种智能分诊装置,其特征在于,所述智能分诊装置包括:

获取模块,用于获取患者的病情特征数据;

处理模块,用于根据所述患者的病情特征数据、以及病例数据库中每个病种下各病例的病情特征数据,确定所述患者患所述病例数据库中每个病种的可能性;

输出模块,用于根据所述患者患所述病例数据库中每个病种的可能性,输出所述患者的分诊结果。

10. 根据权利要求9所述的装置,其特征在于,所述病情特征数据包括病情症状信息和/或体征检测参数。

11. 根据权利要求9或10所述的装置,其特征在于,所述获取模块具体用于:

根据预设的矩阵元素的位置与矩阵元素所表示的病情特征的对应关系,将输入的患者病情特征转换成患者的病情特征分布矩阵,所述患者的病情特征分布矩阵中的每个元素用于指示所述输入的患者病情特征中是否出现该元素所在位置对应的病情特征。

12. 根据权利要求11所述的装置,其特征在于,所述处理模块具体用于:

将所述患者的病情特征数据代入到关系模型中,得到满足所述关系模型的所有相关系数矩阵 $X$ ;

从所述所有相关系数矩阵 $X$ 中确定出唯一的相关系数矩阵 $X_0$ ;

根据所述唯一的相关系数矩阵 $X_0$ ,确定所述病例数据库中每个病种的可能性;

其中,所述关系模型为: $h = DX$ ,所述 $h$ 为所述患者的病情特征分布矩阵,所述 $D$ 为由所述病例数据库中每个病种下各病例的病情特征分布矩阵组成的矩阵,所述 $D = [D_1, D_2, \dots, D_i, \dots, D_M]$ ,其中: $D_i = [D_{i,1}, D_{i,2}, \dots, D_{i,j}, \dots, D_{i,K}]$ ,所述 $D_{i,j}$ 为所述病例数据库中病种 $i$ 的第 $j$ 个病例的病例病情特征分布矩阵,所述 $K$ 用于表示所述病例数据库中病种 $i$ 包括 $K$ 个病例,所述 $M$ 用于表示病例数据库中包括 $M$ 种疾病。

13. 根据权利要求12所述的装置,其特征在于,所述处理模块在从所述所有相关系数矩阵 $X$ 中确定出唯一的相关系数矩阵 $X_0$ 时,具体用于:

从所述所有相关系数矩阵 $X$ 确定出满足第一预设条件的相关系数矩阵 $X$ ;

在满足第一预设条件的相关系数矩阵 $X$ 中确定出满足第二预设条件的唯一的相关系数

矩阵 $X_0$ ;

其中,所述第一预设条件为: $\|DX-h\|_2 \leq \varepsilon$ ,所述第二预定条件为: $x^* = \arg \min \|X_0\|_1$ ,其中:所述 $\|\cdot\|_1$ 是L1范式,所述 $\|\cdot\|_2$ 是L2范式,所述 $\varepsilon$ 为预设参数, $x^*$ 为目标函数。

14. 根据权利要求12所述的装置,其特征在于,所述处理模块在根据所述唯一的相关系数矩阵 $X_0$ ,确定所述病例数据库中每个病种的可能性时,具体用于:

从所述相关系数矩阵 $X_0$ 中确定出所述病例数据库中每个病种的相关系数矩阵 $\delta_i(X_0)$ ;

将所述病例数据库中每个病种的相关系数矩阵 $\delta_i(X_0)$ 代入到概率计算公式中,得到所述患者患所述病例数据库中每个病种的概率;

其中,所述概率计算公式为: $C_i = \frac{\|h_i\|_2^2}{\sum_{j=1}^M \|h_j\|_2^2 + \|\eta\|_2^2}$ ,所述 $C_i$ 用于表示所述患者患所述病例

数据库中的病种 $i$ 的概率,所述 $h_i = D * \delta_i(X_0)$ ,所述 $\delta_i(X_0)$ 为所述病例数据库中病种 $i$ 的相关系数矩阵,所述 $\sum_{j=1}^M \|h_j\|_2^2$ 中的 $M$ 用于表示所述病例数据库中的 $M$ 个病种,所述 $\eta$ 为误差矩阵,所述 $\|\cdot\|_2^2$ 是L2范式的平方。

15. 根据权利要求9所述的装置,其特征在于,所述输出模块具体用于:

输出每个病种的可能性中的最大的对应的所述患者的分诊结果;

或者,将每个病种的可能性中不为零的,按照可能性大小输出所述患者的分诊结果。

## 一种智能分诊方法及装置

### 技术领域

[0001] 本发明涉及医疗技术领域,尤其涉及一种智能分诊方法及装置。

### 背景技术

[0002] 随着人民生活条件的不断改善,对于健康的需求也越来越旺盛。近年来,各大医院的门急诊量急剧增长。据不完全统计,北京各大医院的门诊量在过去的几年中增长高达一倍多,由此造成医疗分诊压力大,流程滞后、等候时间长,继而带来医疗质量难以保证,医患矛盾增加等一系列问题。部分患者为能快速就诊,不论病情缓急直接前往急诊,统计表明我国急诊科的非急症病人的比例达到了32%,这又进一步降低了医院的救治效率,造成恶性循环。

[0003] 因此,如何实现智能分诊是本领域的技术人员目前亟需解决的技术问题。

### 发明内容

[0004] 本发明的实施例提供一种智能分诊方法及装置,用以对患者实现智能分诊来缓解医院的分诊压力。

[0005] 为达到上述目的,本发明的实施例采用如下技术方案:

[0006] 本发明实施例的第一方面,提供一种智能分诊方法,所述方法包括:

[0007] 获取患者的病情特征数据;

[0008] 根据所述患者的病情特征数据、以及病例数据库中每个病种下各病例的病情特征数据,确定所述患者患所述病例数据库中每个病种的可能性;

[0009] 根据所述患者患所述病例数据库中每个病种的可能性,输出所述患者的分诊结果。

[0010] 示例性的,所述病情特征数据包括病情症状信息和/或体征检测参数。

[0011] 优选的,所述获取患者的病情特征数据,具体包括:

[0012] 根据预设的矩阵元素的位置与矩阵元素所表示的病情特征的对应关系,将输入的患者的病情特征转换成患者的病情特征分布矩阵,所述患者的病情特征分布矩阵中的每个元素用于指示所述输入的患者的病情特征中是否出现该元素所在位置对应的病情特征。

[0013] 进一步优选的,所述根据所述患者的病情特征数据、以及病例数据库中每个病种下各病例的病情特征数据,确定所述患者患所述病例数据库中每个病种的可能性,具体包括:

[0014] 将所述患者的病情特征数据代入到关系模型中,得到满足所述关系模型的所有相关系数矩阵 $X$ ;

[0015] 从所述所有相关系数矩阵 $X$ 中确定出唯一的相关系数矩阵 $X_0$ ;

[0016] 根据所述唯一的相关系数矩阵 $X_0$ ,确定所述病例数据库中每个病种的可能性;

[0017] 其中,所述关系模型为: $h=DX$ ,所述 $h$ 为所述患者的病情特征分布矩阵,所述 $D$ 为由所述病例数据库中每个病种下各病例的病情特征分布矩阵组成的矩阵,所述 $D=[D_1, D_2, \dots,$

$D_i, \dots, D_M]$ , 其中:  $D_i = [D_{i,1}, D_{i,2}, \dots, D_{i,j}, \dots, D_{i,K}]$ , 所述  $D_{i,j}$  为所述病例数据库中病种  $i$  的第  $j$  个病例的病例病情特征分布矩阵, 所述  $K$  用于表示所述病例数据库中病种  $i$  包括  $K$  个病例, 所述  $M$  用于表示病例数据库中包括  $M$  种疾病。

[0018] 进一步优选的, 所述方法还包括:

[0019] 根据预设的矩阵元素的位置与矩阵元素所表示的病情特征的对应关系, 将每个病种下各病例的病情特征转换成病例病情特征分布矩阵; 所述病例的病情特征分布矩阵中的每个元素用于指示所述病例的病情特征中是否出现该元素所在位置对应的病情特征。

[0020] 优选的, 所述从所述所有相关系数矩阵  $X$  中确定出唯一的相关系数矩阵  $X_0$ , 具体包括:

[0021] 从所述所有相关系数矩阵  $X$  确定出满足第一预设条件的相关系数矩阵  $X$ ;

[0022] 在满足第一预设条件的相关系数矩阵  $X$  中确定出满足第二预设条件的唯一的相关系数矩阵  $X_0$ ;

[0023] 其中, 所述第一预设条件为:  $\|DX - h\|_2 \leq \varepsilon$ , 所述第二预定条件为:  $x^* = \arg \min \|X_0\|_1$ , 其中: 所述  $\|\cdot\|_1$  是  $L1$  范式, 所述  $\|\cdot\|_2$  是  $L2$  范式, 所述  $\varepsilon$  为预设参数,  $x^*$  为目标函数。

[0024] 优选的, 所述根据所述唯一的相关系数矩阵  $X_0$ , 确定所述病例数据库中每个病种的可能性, 具体包括:

[0025] 从所述相关系数矩阵  $X_0$  中确定出所述病例数据库中每个病种的相关系数矩阵  $\delta_i(X_0)$ ;

[0026] 将所述病例数据库中每个病种的相关系数矩阵  $\delta_i(X_0)$  代入到概率计算公式中, 得到所述患者患所述病例数据库中每个病种的概率;

[0027] 其中, 所述概率计算公式为: 
$$C_i = \frac{\|h_i\|_2^2}{\sum_{j=1}^M \|h_j\|_2^2 + \|\eta\|_2^2}$$
 所述  $C_i$  用于表示所述患者患所述

病例数据库中的病种  $i$  的概率, 所述  $h_i = D * \delta_i(X_0)$ , 所述  $\delta_i(X_0)$  为所述病例数据库中病种  $i$  的相关系数矩阵, 所述  $\sum_{j=1}^M \|h_j\|_2^2$  中的  $M$  用于表示所述病例数据库中的  $M$  个病种, 所述  $\eta$  为误差矩阵, 所述  $\|\cdot\|_2^2$  是  $L2$  范式的平方。

[0028] 优选的, 所述根据所述病例数据库中每个病种的可能性, 输出所述患者的分诊结果, 具体包括:

[0029] 输出每个病种的可能性中的最大的对应的所述患者的分诊结果;

[0030] 或者, 将每个病种的可能性中不为零的, 按照可能性大小输出所述患者的分诊结果。

[0031] 本发明实施例的第二方面, 提供一种智能分诊装置, 所述装置包括:

[0032] 获取模块, 用于获取患者的病情特征数据;

[0033] 处理模块, 用于根据所述患者的病情特征数据、以及病例数据库中每个病种下各病例的病情特征数据, 确定所述患者患所述病例数据库中每个病种的可能性;

[0034] 输出模块, 用于根据所述患者患所述病例数据库中每个病种的可能性, 输出所述

患者的分诊结果。

[0035] 示例性的,所述病情特征数据包括病情症状信息和/或体征检测参数。

[0036] 优选的,所述获取模块具体用于:

[0037] 根据预设的矩阵元素的位置与矩阵元素所表示的病情特征的对应关系,将输入的患者的病情特征转换成患者的病情特征分布矩阵,所述患者的病情特征分布矩阵中的每个元素用于指示所述输入的患者的病情特征中是否出现该元素所在位置对应的病情特征。

[0038] 进一步优选的,所述处理模块具体用于:

[0039] 将所述患者的病情特征数据代入到关系模型中,得到满足所述关系模型的所有相关系数矩阵X;

[0040] 从所述所有相关系数矩阵X中确定出唯一的相关系数矩阵 $X_0$ ;

[0041] 根据所述唯一的相关系数矩阵 $X_0$ ,确定所述病例数据库中每个病种的可能性;

[0042] 其中,所述关系模型为: $h=DX$ ,所述h为所述患者的病情特征分布矩阵,所述D为由所述病例数据库中每个病种下各病例的病情特征分布矩阵组成的矩阵,所述 $D=[D_1, D_2, \dots, D_i, \dots, D_M]$ ,其中: $D_i=[D_{i,1}, D_{i,2}, \dots, D_{i,j}, \dots, D_{i,K}]$ ,所述 $D_{i,j}$ 为所述病例数据库中病种i的第j个病例的病例病情特征分布矩阵,所述K用于表示所述病例数据库中病种i包括K个病例,所述M用于表示病例数据库中包括M种疾病。

[0043] 进一步优选的,其特征在于,所述处理模块在从所述所有相关系数矩阵X中确定出唯一的相关系数矩阵 $X_0$ 时,具体用于:

[0044] 从所述所有相关系数矩阵X确定出满足第一预设条件的相关系数矩阵X;

[0045] 在满足第一预设条件的相关系数矩阵X中确定出满足第二预设条件的唯一的相关系数矩阵 $X_0$ ;

[0046] 其中,所述第一预设条件为: $\|DX-h\|_2 \leq \varepsilon$ ,所述第二预定条件为: $x^* = \arg \min \|X_0\|_1$ ,其中:所述 $\|\cdot\|_1$ 是L1范式,所述 $\|\cdot\|_2$ 是L2范式,所述 $\varepsilon$ 为预设参数, $x^*$ 为目标函数。

[0047] 进一步优选的,所述处理模块在根据所述唯一的相关系数矩阵 $X_0$ ,确定所述病例数据库中每个病种的可能性时,具体用于:

[0048] 从所述相关系数矩阵 $X_0$ 中确定出所述病例数据库中每个病种的相关系数矩阵 $\delta_i(X_0)$ ;

[0049] 将所述病例数据库中每个病种的相关系数矩阵 $\delta_i(X_0)$ 代入到概率计算公式中,得到所述患者患所述病例数据库中每个病种的概率;

[0050] 其中,所述概率计算公式为:
$$C_i = \frac{\|h_i\|_2^2}{\sum_{j=1}^M \|h_j\|_2^2 + \|\eta\|_2^2}$$
,所述 $C_i$ 用于表示所述患者患所述

病例数据库中的病种i的概率,所述 $h_i = D * \delta_i(X_0)$ ,所述 $\delta_i(X_0)$ 为所述病例数据库中病种i的相关系数矩阵,所述 $\sum_{j=1}^M \|h_j\|_2^2$ 中的M用于表示所述病例数据库中的M个病种,所述 $\eta$ 为误差矩阵,所述 $\|\cdot\|_2^2$ 是L2范式的平方。

[0051] 优选的,所述输出模块具体用于:

- [0052] 输出每个病种的可能性中的最大的对应的所述患者的分诊结果；
- [0053] 或者，将每个病种的可能性中不为零的，按照可能性大小输出所述患者的分诊结果。
- [0054] 本发明实施例提供的智能分诊方法及装置，首先，通过获取患者的病情特征数据；然后，根据患者的病情特征数据、以及病例数据库中每个病种下各病例的病情特征数据，确定患者患病例数据库中每个病种的可能性；最后，根据患者患病例数据库中每个病种的可能性，输出患者的分诊结果，从而实现了了对患者的智能分诊，以减少医院的分诊压力。

## 附图说明

- [0055] 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性的前提下，还可以根据这些附图获得其他的附图。
- [0056] 图1为本发明实施例提供的一种智能分诊方法的方法流程图；
- [0057] 图2为本发明实施例提供的一种语义空间的示意图；
- [0058] 图3为本发明实施例提供的一种智能分诊装置的结构示意图。

## 具体实施方式

- [0059] 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。
- [0060] 为了便于清楚描述本发明实施例的技术方案，在本发明的实施例中，采用了“第一”、“第二”等字样对功能或作用基本相同的相同项或相似项进行区分，本领域技术人员可以理解“第一”、“第二”等字样并不对数量和执行次序进行限定。
- [0061] 本文中术语“和/或”，仅仅是一种描述关联对象的关联关系，表示可以存在三种关系，例如，A和/或B，可以表示：单独存在A，同时存在A和B，单独存在B这三种情况。另外，本文中字符“/”，一般表示前后关联对象是一种“或”的关系。
- [0062] 本发明实施例提供的智能分诊方法的执行主体可以为智能分诊装置，示例的，该智能分诊装置可以是用于执行上述智能分诊方法的终端，也可以是用于执行上述智能分诊方法的处理器。其中：该终端可以为计算机、智能手机、平板电脑、笔记本电脑、UMPC (Ultra-mobile Personal Computer, 超级移动个人计算机)、上网本、PDA (Personal Digital Assistant, 个人数字助理) 等终端设备，且不限于此。
- [0063] 本发明实施例提供一种智能分诊方法，如图1所示，该方法包括：
- [0064] 101、获取患者的病情特征数据。
- [0065] 该患者的病情特征数据是用于表示患者具有的病情特征的数据。示例性的，患者或者其他人(例如护士)可以通过计算机上安装的智能分诊系统的输入界面，输入病情特征文本；还可以通过计算机的语音采集模块(例如麦克风)，采集患者声音，并且由计算机上安装的语音识别系统识别出患者口述信息，计算机上安装的智能分诊系统从识别出的患者口



述信息中选择出、或模糊匹配出病情特征关键词(是指病情特征数据库中所存在的病情特征);计算机上安装的智能分诊系统还可通过计算机的身份识别功能,识别患者身份信息(例如扫描身份证或就诊卡等),根据患者身份信息从医院检查数据库(用于存储体征检测参数的数据库,体征检测参数包括患者做身体检查的项目和结果)中调取该患者的体征检测参数。

[0066] 示例性的,上述的病情特征数据包括:病情症状信息和/或体征检测参数,其中,病情症状信息为观察到患者的症状或患者感受到的症状,例如可以是患者的口述症状或输入的症状文本等,例如:心悸气短、肢体麻木、耳鸣等。而体征检测参数包括患者的各项指标检测值,例如血压值、血糖值等,其反映出的病情特征可以是血压微高、血压过高等。

[0067] 当然,病情特征数据除了可以是上述类型以外,还可以是表示病情特征数据的数字集合,例如可以是矩阵。此时,步骤101具体包含以下内容:

[0068] 101a、根据预设的矩阵元素的位置与矩阵元素所表示的病情特征的对应关系,将输入的患者的病情特征转换成患者的病情特征分布矩阵,患者的病情特征分布矩阵中的每个元素用于指示输入的患者的病情特征中是否出现该元素所在位置对应的病情特征。

[0069] 具体的,假设数据库中有Q个病情特征,该预设的矩阵元素的位置与矩阵元素所表示的病情特征的对应关系用集合I表示,该集合I为 $Q \times 1$ 的矩阵, $I = [I_1, I_2, \dots, I_Q]^T$ 。其中, $I_j (1 \leq j \leq Q)$ 表示第j个位置处的病情特征,从而集合I表示的是从第1个位置处的病情特征到第Q个位置处的病情特征。

[0070] 示例的,假设数据库中有1000个病情特征,上述的集合I为: $I = [I_1, I_2, \dots, I_{1000}]^T$ 。其中, $I_{500}$ 为第500个位置处的病情特征,从而集合I表示从第1个位置处的病情特征到第1000个位置处的病情特征。

[0071] 102、根据患者的病情特征数据、以及病例数据库中每个病种下各病例的病情特征数据,确定患者患病例数据库中每个病种的可能性。

[0072] 示例性的,上述的患者患病例数据库中每个病种的可能性可以是指患者患每个病种的概率,用0至1间的数值进行表示。或者是患者患每个病种的可能性对应的数值(可以为包含大于1的数值),数值越大表示可能性越大。

[0073] 示例性的,上述的病情特征数据为病情特征文本时,例如:患者的病情特征文本为眩晕、恶心以及心悸气短;上述的步骤102中确定患者患病例数据库中每个病种的可能性具体过程参照以下内容:这里病例数据库中的病种个数以3个为例,分别为病种A、病种B以及病种C,其中:病种A以包含3个病例为例,病种B以包含4个病例为例,病种C以包含5个病例为例,而患者所具有的病情特征以3个为例。将患者的病情特征与病例数据库中的每个病种下各病例中的病情特征文本进行匹配,若患者的3个病情特征均出现在病例数据库中病种A下的同一个病例中,且该患者的3个病情特征没有全部出现在其他病种下的病情特征文本中,则该患者患病种A的可能性最大;若患者的2个病情特征出现在病例数据库中病种A下的第一病例中,剩下的1个病情特征出现在病种A下的第二病例中,且该患者的3个病情特征没有全部出现在其他病种下的病情特征文本中,则该患者患病种A的可能性相对于上面的结果较小。若患者的第1个病情特征出现在病例数据库中病种A下的第一病例中,第2个病情特征出现在病种A下的第二病例中,第3个病情特征出现在病种A下的第三病例中,且该患者的3个病情特征没有全部出现在其他病种下的病情特征文本中,则该患者患病种A的可能性相

对于上面的两种结果是最小的。当然,通过匹配病情特征得到患者患每个病种的可能性大小的规则,可以根据实际需要进行设置。

[0074] 示例性的,上述的病情特征数据为矩阵时,上述的步骤102具体包括以下内容:

[0075] 102a、将患者的病情特征数据代入到关系模型中,得到满足关系模型的所有相关系数矩阵X。

[0076] 102b、从所有相关系数矩阵X中确定出唯一的相关系数矩阵 $X_0$ 。

[0077] 102c、根据唯一的相关系数矩阵 $X_0$ ,确定病例数据库中每个病种的可能性。

[0078] 其中,上述的关系模型为: $h=DX$ , $h$ 为患者的病情特征分布矩阵, $D$ 为由病例数据库中每个病种下各病例的病情特征分布矩阵组成的矩阵, $D=[D_1, D_2, \dots, D_i, \dots, D_M]$ ,其中: $D_i=[D_{i,1}, D_{i,2}, \dots, D_{i,j}, \dots, D_{i,K}]$ , $D_{i,j}$ 为病例数据库中病种 $i$ 的第 $j$ 个病例的病例病情特征分布矩阵, $K$ 用于表示病例数据库中病种 $i$ 包括 $K$ 个病例, $M$ 用于表示病例数据库中包括 $M$ 种疾病。

[0079] 基于上述的内容,可选的,在上述的步骤102之前,该方法还包括以下内容:

[0080] A1、根据预设的矩阵元素的位置与矩阵元素所表示的病情特征的对应关系,将每个病种下各病例的病情特征转换成病例病情特征分布矩阵。

[0081] 其中,上述的病例的病情特征分布矩阵中的每个元素用于指示病例的病情特征中是否出现该元素所在位置对应的病情特征。

[0082] 具体的,若病例数据库中有 $Q$ 个病情特征,则病例数据库中每个病种下各病例的病情特征分布矩阵中包含 $Q$ 个元素。

[0083] 示例性的,病种A下的任一个病例的病情特征分布矩阵为 $D_{A1}=[I_1, I_2, \dots, I_Q]^T$ 。由于数据库中的病情特征的集合 $I$ 为 $Q \times 1$ 的矩阵,相应的这里的病种A下的任一个病例的病情特征分布矩阵也为 $Q \times 1$ 的矩阵。其中, $I_j (1 \leq j \leq Q)$ 表示病种A下的任一个病例的第 $j$ 个位置处的病情特征,从而 $D_{A1}$ 表示病种A下的任一个病例中从第1个位置处的病情特征到第 $Q$ 个位置处的病情特征。

[0084] 示例性的,上述的关系模型可以是预先建立好的,也可以是根据需要实时进行建立的,对于上述的关系模型的建立过程可以参考以下的内容:

[0085] 由于本发明是基于医学大数据进行疾病的病种预测,因此需要大量的案例(例如各医院历年的确诊病例),这对应着流程图中的病例数据库中每个病种下各病例。本发明使用符号 $D$ 来表示病例数据库中病种集合,假设其中一共包含 $M$ 种疾病(即 $M$ 个病种),则 $D_i (1 \leq i \leq M)$ 表示病例数据库的第 $i$ 种疾病。假设第 $i$ 种疾病中包含 $K$ 个病例,则 $D_{i,j} (1 \leq i \leq M, 1 \leq j \leq K)$ 表示第 $i$ 种疾病中的第 $j$ 个病例。每一个病例由一系列对应的特征向量(如症状和体征检测参数)构成,则 $D$ 构成了一个确诊病例的语义空间。

[0086] 对于新来的患者 $h$ (其含义是指:患者的病情特征分布矩阵用 $h$ 表示),假设其患有疾病 $D_i$ ,依据本发明的基本思想:患有同一疾病的患者极有可能出现相似的特征(如症状和体征检测参数),则患者 $h$ 可以表示为 $D_i$ 中包含病例的线性组合,即 $h=\alpha_{i,1} \times D_{i,1} + \alpha_{i,2} \times D_{i,2} + \dots + \alpha_{i,K} \times D_{i,K}$ ,其中, $\alpha_{i,j}$ 是相关系数。例如,对于疾病“高血压”,病例1中的症状有“眩晕、恶心、心悸气短”,病例2中的症状有“心悸气短、耳鸣、肢体麻木”,病例3中的症状有“眩晕、恶心、耳鸣、心悸气短”,新来患者的症状有“心悸气短、肢体麻木”,则有“新来患者=病例1+病例2-病例3”。

[0087] 为了表示简洁和方便,上面的表达形式可以用矩阵表示。假设 $D_i=[D_{i1},$

$D_{i2}, \dots, D_{iK}]$ ,  $X_i = [\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{iK}]^T$ , 其中上标T表示矩阵的转置, 则有  $h = D_i X_i$ 。

[0088] 通过上面的讨论, 可以看到每一个患者可以表示成由其包含的已知病种下病例所构成的语义子空间, 属于该病种的某一病例可以由相应子空间 (病情特征) 的线性组合构成。

[0089] 上面所讨论的是新来的患者h, 假设其患有疾病 $D_i$ 所做的讨论, 那么对于新来的患者在不知道所患病种的前提下, 类比于上述的过程, 当给定病种矩阵D, 可以通过寻找患者h在D中的语义子空间来确定其所患疾病。令  $D = [D_1, D_2, \dots, D_M]$ , 则患者h与给定病例间的关系模型为:  $h = DX$ 。

[0090] 具体的, 对于上面的  $D = [D_1, D_2, \dots, D_M]$ , 由于病种 $D_1, D_2, \dots, D_M$ 这M个病种的每个病种下可能包含多个病例, 因此, 这里 $D_1, D_2, \dots, D_M$ 中的 $D_i$  ( $1 \leq i \leq M$ ) 为由第i个病种下所包含的各病例的病情特征分布矩阵构成的集合。例如, 假设M个病种的每个病种下均包含两个病例, 则  $D = [D_{11}, D_{12}, D_{21}, D_{22}, \dots, D_{M1}, D_{M2}]$ 。

[0091] 示例性的, 假设病例数据库中有3个病种, 分别为病种A、病种B以及病种C, 该病种A包含3个病例, 病种B包含2个病例, 病种C包含2个病例, 则  $D = [D_{A1}, D_{A2}, D_{B1}, D_{B2}, D_{B3}, D_{C1}, D_{C2}]$ 。

[0092] 基于上面的病种矩阵D,  $D = [D_{A1}, D_{A2}, D_{B1}, D_{B2}, D_{B3}, D_{C1}, D_{C2}]$ , 假设基于上述的3个病种的7个病例统计出的病情特征数据有100个, 那么, 上面的患者h与给定病例间的关系模型:  $h = DX$  中的病种矩阵D是一个  $100 \times 7$  的矩阵, 而相关系数矩阵X为  $7 \times 1$  的矩阵, 用  $X = [\alpha_{A1}, \alpha_{A2}, \alpha_{B1}, \alpha_{B2}, \alpha_{B3}, \alpha_{C1}, \alpha_{C2}]^T$  来表示。

[0093] 需要说明的是, 在实际的应用中, 上述的病例数据库中的病种的个数为成百上千个, 而每个病种下的病例相应的可能也是成百上千乃至更多, 基于每个病种下的病例所抽出的病情特征可能是成千上万个, 因此, 上面的内容仅仅是一种示例, 用于解释说明本方案, 并不进行限定。

[0094] 示例性的, 上述的步骤102b中确定相关系数矩阵 $X_0$ , 可以是: 从所有相关系数矩阵X中任意选择一个相关系数矩阵 $X_0$ 。然后基于该任意选择的一个关系系数矩阵 $X_0$ 确定病例数据库中每个病种的可能性。

[0095] 示例性的, 上述的步骤102b中确定相关系数矩阵 $X_0$ , 可以是采用算法来确定, 具体包括以下内容:

[0096] 102b1、从所有相关系数矩阵X确定出满足第一预设条件的相关系数矩阵X。

[0097] 102b2、在满足第一预设条件的相关系数矩阵X中确定出满足第二预设条件的唯一的相关系数矩阵 $X_0$ 。

[0098] 其中, 上述的第一预设条件为:  $\|DX - h\|_2 \leq \epsilon$ , 上述的第二预设条件为:  $x^* = \arg \min \|X_0\|_1$ , 其中:  $\|\cdot\|_1$  是L1范式,  $\|\cdot\|_2$  是L2范式,  $\epsilon$  为预设参数,  $x^*$  为目标函数。其中, 该  $x^* = \arg \min \|X_0\|_1$  表示的是  $x^*$  取最小值时, 对应的 $X_0$ 为唯一的相关系数矩阵。

[0099] 需要说明的是, 上述的L1范式  $\|\cdot\|_1$  的运算是: 范式中变量所包含的每个元素的绝对值之和, 例如, 若  $X = [\alpha_{11}, \alpha_{12}, \dots, \alpha_{MK}]$ , 则  $\|X\|_1 = |\alpha_{11}| + |\alpha_{12}| + \dots + |\alpha_{MK}|$ 。而上述的L2范式  $\|\cdot\|_2$  的运算是: 范式中变量所包含的每个元素的平方之和, 例如, 若  $X = [\alpha_{11}, \alpha_{12}, \dots, \alpha_{MK}]$ , 则  $\|X\|_2 = \alpha_{11}^2 + \alpha_{12}^2 + \dots + \alpha_{MK}^2$ 。

[0100] 上述的步骤102b1以及102b2中的第一预设条件和第二预设条件中所采用的是稀疏解法, 即使用最少的病例去重构患者h的病情特征, 采用稀疏解法能够降低“噪音”数据的

影响,使得上述的关系模型 $h=DX$ 具有良好的鲁棒性。

[0101] 示例性的,当步骤102c中的可能性用概率来表示时,上述的步骤102c具体包括以下内容:

[0102] 102c1、从唯一的相关系数矩阵 $X_0$ 中确定出病例数据库中每个病种的相关系数矩阵 $\delta_i(X_0)$ 。

[0103] 其中,将 $X_0$ 中第 $i$ 个病种下各病例的相关系数保留,其他元素置为0,得到 $\delta_i(X_0)$ 。

[0104] 示例性的,假设唯一的相关系数矩阵 $X_0=[\alpha_{A1},\alpha_{A2},\alpha_{B1},\alpha_{B2},\alpha_{B3},\alpha_{C1},\alpha_{C2}]^T$ ,则病种A的相关系数矩阵为: $\delta_A(X_0)=[\alpha_{A1},\alpha_{A2},0,0,0,0,0]^T$ ;病种B的相关系数矩阵为: $\delta_B(X_0)=[0,0,\alpha_{B1},\alpha_{B2},\alpha_{B3},0,0]^T$ ;病种C的相关系数矩阵为: $\delta_C(X_0)=[0,0,0,0,0,\alpha_{C1},\alpha_{C2}]^T$ 。

[0105] 102c2、将病例数据库中每个病种的相关系数矩阵 $\delta_i(X_0)$ 代入到概率计算公式中,得到患者患病例数据库中每个病种的概率。

[0106] 示例性的,概率计算公式为:

$$[0107] \quad C_i = \frac{\|h_i\|_2^2}{\sum_{j=1}^M \|h_j\|_2^2 + \|\eta\|_2^2}, \quad (\text{公式 1})$$

[0108] 其中, $C_i$ 用于表示患者患病例数据库中的病种 $i$ 的概率, $h_i=D*\delta_i(X_0)$ , $\delta_i(X_0)$ 为病例数据库中病种 $i$ 的相关系数矩阵, $\sum_{j=1}^M \|h_j\|_2^2$ 中的 $M$ 用于表示病例数据库中的 $M$ 个病种, $\eta$ 为误差矩阵,该 $\eta=h-h_1-h_2-\dots-h_M$ ,其中: $h$ 为新来患者的病情特征分布矩阵, $h_i$ 为病例数据库中的病种 $i$  ( $1 \leq i \leq M$ ) 的病情特征分布矩阵,而该 $h_i$ 是由病种 $i$ 下的所有病例的病情特征分布矩阵组成的。 $\|\cdot\|_2^2$ 是L2范式的平方。

[0109] 具体的,对于上述的 $h_i=D*\delta_i(X_0)$ ,仍然以上文所列举的例子进行说明。假设病例数据库中有3个病种,分别为病种A、病种B以及病种C,该病种A包含3个病例,病种B包含2个病例,病种C包含2个病例,则 $D=[D_{A1},D_{A2},D_{B1},D_{B2},D_{B3},D_{C1},D_{C2}]$ 。假设基于上述的3个病种的7个病例统计出的病情特征数据有100个,且假设唯一的相关系数矩阵 $X_0=[\alpha_{A1},\alpha_{A2},\alpha_{B1},\alpha_{B2},\alpha_{B3},\alpha_{C1},\alpha_{C2}]^T$ 。

[0110] 基于上面的内容,考虑到上述的3个病种的7个病例统计出的病情特征数据有100个,则对应的 $D$ 为 $100*7$ 的矩阵,所确定出 $h_A=D*\delta_A(X_0)$ 中的 $\delta_A(X_0)$ 为 $7*1$ 的矩阵, $\delta_A(X_0)=[\alpha_{A1},\alpha_{A2},0,0,0,0,0]^T$ ;  $h_B=D*\delta_B(X_0)$ 中的 $\delta_B(X_0)$ 为 $7*1$ 的矩阵, $\delta_B(X_0)=[0,0,\alpha_{B1},\alpha_{B2},\alpha_{B3},0,0]^T$ ;  $h_C=D*\delta_C(X_0)$ 中的 $\delta_C(X_0)$ 为 $7*1$ 的矩阵, $\delta_C(X_0)=[0,0,0,0,0,\alpha_{C1},\alpha_{C2}]^T$ 。这样上述的 $h_A$ 、 $h_B$ 以及 $h_C$ 中的矩阵运算才满足矩阵乘法的准则。然后,将的 $h_A$ 、 $h_B$ 以及 $h_C$ 的内容带入到上面的公式1中可以得到患者患病种A、病种B以及病种C的概率。

[0111] 示例性的,上述的 $C=[C_1,C_2,\dots,C_M,C_n]$ ,由上述的公式1中可以得知 $C_i$ 满足 $C_1+C_2+\dots+C_M+C_n=1$ ,其中, $C_n$ 的计算公式如下:

$$[0112] \quad C_n = \frac{\|\eta\|_2^2}{\sum_{j=1}^M \|h_j\|_2^2 + \|\eta\|_2^2} \quad (\text{公式 2})$$

[0113] 通过上述的公式1和公式2可以得知, $C_i$ 反映了患者 $h$ 属于病种 $D_i$ 可能性的 $C_n$

反映了患者h不属于前面任一病种D<sub>1</sub>-DM的可能性)。这是因为C<sub>i</sub>越大,表明构成患者h的病情特征分布矩阵中包含属于病种D<sub>i</sub>的病例越多,即患者h位于D<sub>i</sub>语义子空间的部分越多,则属于病种D<sub>i</sub>的可能性越大。例如,参考图2中给出的语义空间示意图,假设已知病例中一共有三个病种,图2中的不同的形状代表不同的病种,相同的形状的个数代表该病种下的病例个数,其中:圆点是新来的患者,则明显可以得到对于图2中的(a)有 $C=[1,0,0,0]$ ,即患者可能患有四角星所代表的病种。对于图2中的(b)有 $C=[0.25,0.375,0.375,0]$ ,则很难准确判断出患者患有何种疾病。

[0114] 103、根据患者患病例数据库中每个病种的可能性,输出患者的分诊结果。

[0115] 其中,上述的分诊结果可以包括为患者分配的科室、进一步还可以包括分诊流程、为患者所分配的医生、以及可参考的治疗指南等。

[0116] 示例性的,上述的步骤103可以采用以下任一种方式实现:

[0117] 方式A、输出每个病种的可能性中的最大的对应的患者的分诊结果。例如,计算患者患每个病种的可能性大小,且所确定出患者患病种A的可能性最大,则在智能分诊系统的界面上显示出患者患病种A所对应的分诊结果。

[0118] 方式B、将每个病种的可能性中不为零的,按照可能性大小输出患者的分诊结果。例如,计算出患者患每个病种的可能性,并将每个病种的可能性中不为零的按从大到小的顺序排序,则在智能分诊系统的界面上从大到小依次显示出患者患每个病种的可能性的分诊结果。

[0119] 本发明实施例提供的智能分诊方法,首先,通过获取患者的病情特征数据;然后,根据患者的病情特征数据、以及病例数据库中每个病种下各病例的病情特征数据,确定患者患病例数据库中每个病种的可能性;最后,根据患者患病例数据库中每个病种的可能性,输出患者的分诊结果,从而实现了患者的智能分诊,以减少医院的分诊压力。

[0120] 下面将基于图1对应的智能分诊方法的实施例中的相关描述对本发明实施例提供的一种智能分诊装置进行介绍。以下实施例中与上述实施例相关的技术术语、概念等的说明可以参照上述的实施例,这里不再赘述。

[0121] 本发明实施例提供一种智能分诊装置,如图3所示,该装置包括:获取模块31、处理模块32以及确定模块33,其中:

[0122] 获取模块31,用于获取患者的病情特征数据。

[0123] 处理模块32,用于根据患者的病情特征数据、以及病例数据库中每个病种下各病例的病情特征数据,确定患者患所述病例数据库中每个病种的可能性。

[0124] 输出模块33,用于根据患者患病例数据库中每个病种的可能性,输出患者的分诊结果。

[0125] 示例性的,该患者的病情特征数据可以是病情特征文本,也可以是用于表示该病情特征文本的数据。

[0126] 示例性的,上述的病情特征数据包括:病情症状信息和/或体征检测参数,其中,病情症状信息为观察到患者的症状或患者感受到的症状,例如可以是患者的口述症状或输入的症状文本等,例如:心悸气短、肢体麻木、耳鸣等。而体征检测参数包括患者的各项指标检测值,例如血压值、血糖值等,其反映出的病情特征可以是血压微高、血压过高等。

[0127] 示例性的,上述的患者患病例数据库中每个病种的可能性可以是指患者患每个病

种的概率,用0至1间的数值进行表示。或者是患者患每个病种的可能性对应的数值(可以为包含大于1的数值),数值越大表示可能性越大。

[0128] 示例性的,上述的病情特征数据为病情特征文本时,例如:患者的病情特征文本为眩晕、恶心以及心悸气短;上述的步骤102中确定患者患病例数据库中每个病种的可能性具体过程参照以下内容:这里病例数据库中的病种个数以3个为例,分别为病种A、病种B以及病种C,其中:病种A以包含3个病例为例,病种B以包含4个病例为例,病种C以包含5个病例为例,而患者所具有的病情特征以3个为例。将患者的病情特征与病例数据库中的每个病种下各病例中的病情特征文本进行匹配,若患者的3个病情特征均出现在病例数据库中病种A下的同一个病例中,且该患者的3个病情特征没有全部出现在其他病种下的病情特征文本中,则该患者患病种A的可能性最大;若患者的2个病情特征出现在病例数据库中病种A下的第一病例中,剩下的1个病情特征出现在病种A下的第二病例中,且该患者的3个病情特征没有全部出现在其他病种下的病情特征文本中,则该患者患病种A的可能性相对于上面的结果较小。若患者的第1个病情特征出现在病例数据库中病种A下的第一病例中,第2个病情特征出现在病种A下的第二病例中,第3个病情特征出现在病种A下的第三病例中,且该患者的3个病情特征没有全部出现在其他病种下的病情特征文本中,则该患者患病种A的可能性相对于上面的两种结果是最小的。

[0129] 示例性的,上述的获取模块31具体用于:

[0130] 根据预设的矩阵元素的位置与矩阵元素所表示的病情特征的对应关系,将输入的患者的病情特征转换成患者的病情特征分布矩阵,患者的病情特征分布矩阵中的每个元素用于指示输入的患者的病情特征中是否出现该元素所在位置对应的病情特征。

[0131] 具体的,假设数据库中有Q个病情特征,该预设的矩阵元素的位置与矩阵元素所表示的病情特征的对应关系用集合I表示,该集合I为 $Q \times 1$ 的矩阵, $I = [I_1, I_2, \dots, I_Q]^T$ 。其中, $I_j (1 \leq j \leq Q)$ 表示第j个位置处的病情特征,从而集合I表示的是从第1个位置处的病情特征到第Q个位置处的病情特征。

[0132] 示例的,假设数据库中有1000个病情特征,上述的集合I为: $I = [I_1, I_2, \dots, I_{1000}]^T$ 。其中, $I_{500}$ 为第500个位置处的病情特征,从而集合I表示从第1个位置处的病情特征到第1000个位置处的病情特征。

[0133] 示例性的,上述的处理模块32具体用于:

[0134] 将患者的病情特征数据代入到关系模型中,得到满足关系模型的所有相关系数矩阵X。

[0135] 从所有相关系数矩阵X中确定出唯一的相关系数矩阵 $X_0$ 。

[0136] 根据唯一的相关系数矩阵 $X_0$ ,确定病例数据库中每个病种的可能性。

[0137] 其中,关系模型为: $h = DX$ ,h为患者的病情特征分布矩阵,D为由病例数据库中每个病种下各病例的病情特征分布矩阵组成的矩阵, $D = [D_1, D_2, \dots, D_i, \dots, D_M]$ ,其中: $D_i = [D_{i,1}, D_{i,2}, \dots, D_{i,j}, \dots, D_{i,K}]$ , $D_{i,j}$ 为病例数据库中病种i的第j个病例的病例病情特征分布矩阵,K用于表示病例数据库中病种i包括K个病例,M用于表示病例数据库中包括M种疾病。

[0138] 基于上述的内容,可选的,如图3所示,该装置还包括:转换模块34,其中:

[0139] 转换模块34用于根据预设的矩阵元素的位置与矩阵元素所表示的病情特征的对应关系,将每个病种下各病例的病情特征转换成病例病情特征分布矩阵。

[0140] 其中,上述的病例的病情特征分布矩阵中的每个元素用于指示病例的病情特征中是否出现该元素所在位置对应的病情特征。

[0141] 具体的,若病例数据库中有Q个病情特征,则病例数据库中每个病种下各病例的病情特征分布矩阵中包含Q个元素。

[0142] 示例性的,病种A下的任一个病例的病情特征分布矩阵为 $D_{A1} = [I_1, I_2, \dots, I_Q]^T$ 。由于数据库中的病情特征的集合I为 $Q \times 1$ 的矩阵,相应的这里的病种A下的任一个病例的病情特征分布矩阵也为 $Q \times 1$ 的矩阵。其中, $I_j (1 \leq j \leq Q)$ 表示病种A下的任一个病例的第j个位置处的病情特征,从而 $D_{A1}$ 表示的病种A下的任一个病例中从第1个位置处的病情特征到第Q个位置处的病情特征。

[0143] 示例性的,上述的关系模型可以是预先建立好的,也可以是根据需要实时进行建立的,对于上述的关系模型的建立过程具体可以参考方法部分的内容,这里不再详细赘述。

[0144] 示例性的,上输的处理模块32在从所有相关系数矩阵X中确定出唯一的相关系数矩阵 $X_0$ 时,可以从所有相关系数矩阵X中任意选择一个相关系数矩阵 $X_0$ 。然后基于该任意选择的一个关系系数矩阵 $X_0$ 确定病例数据库中每个病种的可能性。

[0145] 示例性的,上述的处理模块32在从所有相关系数矩阵X中确定出唯一的相关系数矩阵 $X_0$ 时,也可以使用算法进行确定该相关系数矩阵 $X_0$ ,上述的处理模块32具体用于:

[0146] 从所有相关系数矩阵X确定出满足第一预设条件的相关系数矩阵X。

[0147] 在满足第一预设条件的相关系数矩阵X中确定出满足第二预设条件的唯一的相关系数矩阵 $X_0$ 。

[0148] 其中,第一预设条件为: $\|DX-h\|_2 \leq \varepsilon$ ,第二预定条件为: $x^* = \arg \min \|X_0\|_1$ ,其中: $\|\cdot\|_1$ 是L1范式, $\|\cdot\|_2$ 是L2范式, $\varepsilon$ 为预设参数, $x^*$ 为目标函数。其中,该 $x^* = \arg \min \|X_0\|_1$ 表示的是 $x^*$ 取最小值时,对应的 $X_0$ 为唯一的相关系数矩阵。

[0149] 需要说明的是,上述的L1范式 $\|\cdot\|_1$ 的运算是:范式中变量所包含的每个元素的绝对值之和,例如,若 $X = [\alpha_{11}, \alpha_{12}, \dots, \alpha_{MK}]$ ,则 $\|X\|_1 = |\alpha_{11}| + |\alpha_{12}| + \dots + |\alpha_{MK}|$ 。而上述的L2范式 $\|\cdot\|_2$ 的运算是:范式中变量所包含的每个元素的平方之和,例如,若 $X = [\alpha_{11}, \alpha_{12}, \dots, \alpha_{MK}]$ ,则 $\|X\|_2 = \alpha_{11}^2 + \alpha_{12}^2 + \dots + \alpha_{MK}^2$ 。

[0150] 上述的处理模块32在确定唯一的相关系数矩阵 $X_0$ 时所涉及的第一预设条件和第二预设条件中所采用的是稀疏解法,即使用最少的病例去重构患者h的病情特征,采用稀疏解法能够降低“噪音”数据的影响,使得上述的关系模型 $h=DX$ 具有良好的鲁棒性。

[0151] 示例性的,当上述的处理模块32在根据唯一的相关系数矩阵 $X_0$ ,确定病例数据库中每个病种的可能性用概率来表示时,该处理模块32具体用于:

[0152] 从相关系数矩阵 $X_0$ 中确定出病例数据库中每个病种的相关系数矩阵 $\delta_i(X_0)$ 。

[0153] 其中,将 $X_0$ 中第i个病种下各病例的相关系数保留,其他元素置为0,得到 $\delta_i(X_0)$ 。

[0154] 示例性的,假设唯一的相关系数矩阵 $X_0 = [\alpha_{A1}, \alpha_{A2}, \alpha_{B1}, \alpha_{B2}, \alpha_{B3}, \alpha_{C1}, \alpha_{C2}]^T$ ,则病种A的相关系数矩阵为: $\delta_A(X_0) = [\alpha_{A1}, \alpha_{A2}, 0, 0, 0, 0, 0]^T$ ;病种B的相关系数矩阵为: $\delta_B(X_0) = [0, 0, \alpha_{B1}, \alpha_{B2}, \alpha_{B3}, 0, 0]^T$ ;病种C的相关系数矩阵为: $\delta_C(X_0) = [0, 0, 0, 0, 0, \alpha_{C1}, \alpha_{C2}]^T$ 。

[0155] 将病例数据库中每个病种的相关系数矩阵 $\delta_i(X_0)$ 代入到概率计算公式中,得到患者患病例数据库中每个病种的概率。

[0156] 示例性的,上述的概率计算公式为:

$$[0157] \quad C_i = \frac{\|h_i\|_2^2}{\sum_{j=1}^M \|h_j\|_2^2 + \|\eta\|_2^2}, \quad (\text{公式 1})$$

[0158] 其中,  $C_i$  用于表示患者患病例数据库中的病种  $i$  的概率,  $h_i = D * \delta_i(X_0)$ ,  $D_i$  为病例数据库病种  $i$  的病情特征分布矩阵,  $\delta_i(X_0)$  为病例数据库中病种  $i$  的相关系数矩阵,  $\sum_{j=1}^M \|h_j\|_2^2$  中的  $M$  用于表示病例数据库中的  $M$  个病种,  $\eta$  为误差矩阵, 该  $\eta = h - h_1 - h_2 - \dots - h_M$ , 其中:  $h$  为新来患者的病情特征分布矩阵,  $h_i$  为病例数据库中的病种  $i$  ( $1 \leq i \leq M$ ) 的病情特征分布矩阵, 而该  $h_i$  是由病种  $i$  下的所有病例的病情特征分布矩阵组成的。 $\|\cdot\|_2^2$  是 L2 范式的平方。

[0159] 示例性的, 上述的  $C = [C_1, C_2, \dots, C_M, C_\eta]$ , 由上述的公式 1 中可以得知  $C_i$  满足  $C_1 + C_2 + \dots + C_M + C_\eta = 1$ , 其中,  $C_\eta$  的计算公式如下:

$$[0160] \quad C_\eta = \frac{\|\eta\|_2^2}{\sum_{j=1}^M \|h_j\|_2^2 + \|\eta\|_2^2} \quad (\text{公式 2})$$

[0161] 通过上述的公式 1 和公式 2 可以得知,  $C_i$  反映了患者  $h$  属于病种  $D_i$  可能性的的大小 ( $C_\eta$  反映了患者  $h$  不属于前面任一病种  $D_1 - D_M$  的可能性)。这是因为  $C_i$  越大, 表明构成患者  $h$  的病情特征分布矩阵中包含属于病种  $D_i$  的病例越多, 即患者  $h$  位于  $D_i$  语义子空间的部分越多, 则属于病种  $D_i$  的可能性越大。例如, 参考图 2 中给出的语义空间示意图, 假设已知病例中一共有三个病种 (分别对应不同的形状), 圆形节点是新来的患者, 则明显可以看到对于左边的图有  $C = [1, 0, 0, 0]$ , 即患者可能患有四角星所代表的病种。对于右边的图有  $C = [0.25, 0.375, 0.375, 0]$ , 则很难准确判断出患者患有何种疾病。

[0162] 示例性的, 上述的分诊结果包括可参考的治疗指南、分诊流程以及所涉及的科室等信息。

[0163] 示例性的, 上述的输出模块 33 具体用于以下任一种方式实现:

[0164] 方式 A、输出每个病种的可能性中的最大的对应的患者的分诊结果。例如, 计算患者患每个病种的可能性大小, 且所确定出患者患病种 A 的可能性最大, 则在智能分诊系统的界面上显示出患者患病种 A 所对应的分诊结果。

[0165] 方式 B、将每个病种的可能性中不为零的, 按照可能性大小输出患者的分诊结果。例如, 计算出患者患每个病种的可能性, 并将每个病种的可能性中不为零的按从大到小的顺序排序, 则在智能分诊系统的界面上从大到小依次显示出患者患每个病种的可能性的分诊结果。

[0166] 本发明实施例提供的智能分诊装置, 首先, 该装置通过获取患者的病情特征数据; 然后, 根据患者的病情特征数据、以及病例数据库中每个病种下各病例的病情特征数据, 确定患者患病例数据库中每个病种的可能性; 最后, 根据患者患病例数据库中每个病种的可能性, 输出患者的分诊结果, 从而实现了患者的智能分诊, 以减少医院的分诊压力。

[0167] 通过以上的实施方式的描述, 所属领域的技术人员可以清楚地了解到, 为描述的方便和简洁, 仅以上述各功能模块的划分进行举例说明, 实际应用中, 可以根据需要而将上述功能分配由不同的功能模块完成, 即将装置的内部结构划分成不同的功能模块, 以完成



以上描述的全部或者部分功能。上述描述的系统,装置和单元的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0168] 在本申请所提供的几个实施例中,应该理解到,所揭露的智能分诊装置,可以通过其它的方式实现。例如,以上所描述的装置的实施例仅仅是示意性的,例如,所述模块或单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口,装置或单元的间接耦合或通信连接,可以是电性,机械或其它的形式。

[0169] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0170] 另外,在本发明各个实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。

[0171] 所述集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用时,可以存储在一个计算机可读取存储介质中。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或者该技术方案的全部或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一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备等)或处理器(processor)执行本发明各个实施例所述方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0172] 以上所述,仅为本发明的具体实施方式,但本发明的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,可轻易想到变化或替换,都应涵盖在本发明的保护范围之内。因此,本发明的保护范围应以所述权利要求的保护范围为准。

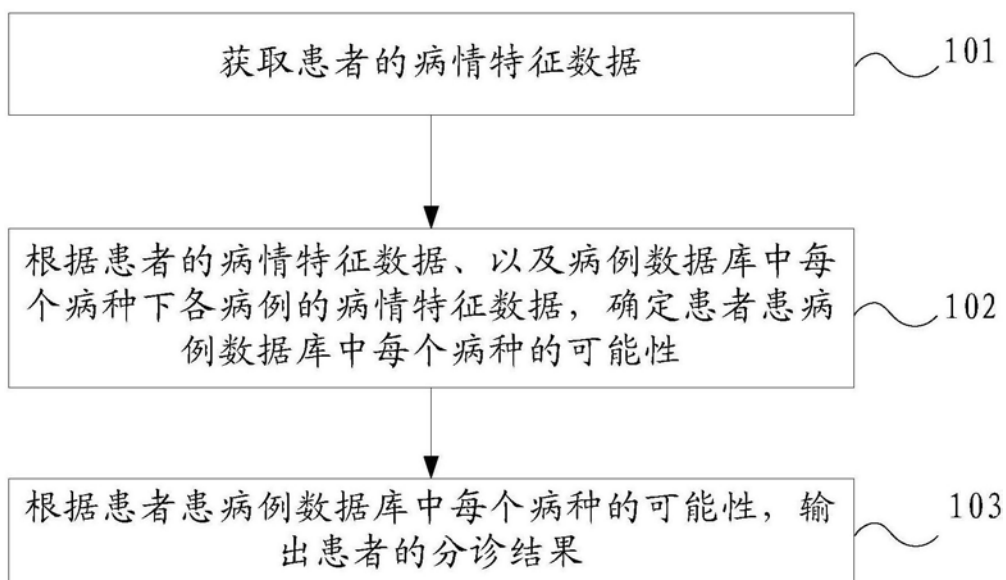


图1

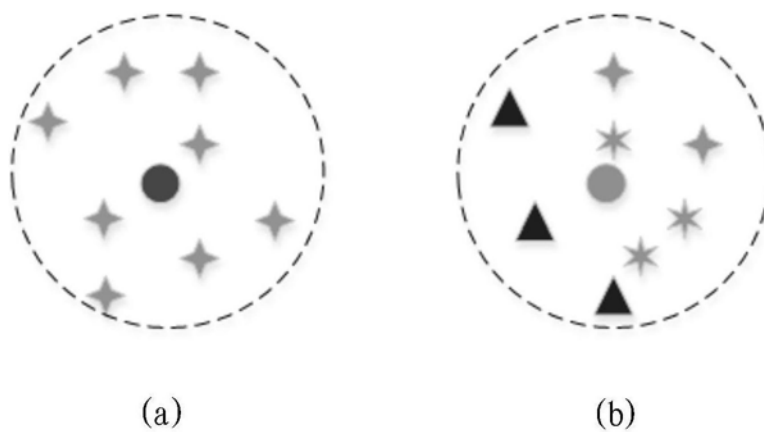


图2

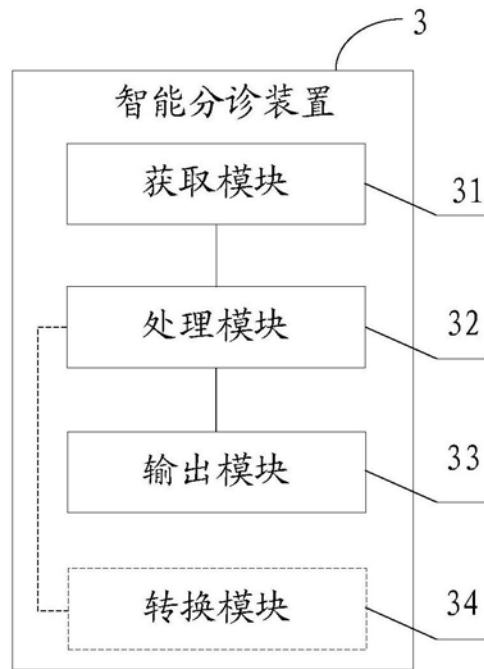


图3