



(12) 发明专利

(10) 授权公告号 CN 113744851 B

(45) 授权公告日 2024. 07. 02

(21) 申请号 202010463204.4

G16H 10/60 (2018.01)

(22) 申请日 2020.05.27

(56) 对比文件

(65) 同一申请的已公布的文献号  
申请公布号 CN 113744851 A

Edward Choi et al..RETAIN: An Interpretable Predictive Model for Healthcare using Reverse Time Attention Mechanism.《Advances in Neural Information Processing Systems (NIPS2016)》.2016,第29卷第1-9页.

(43) 申请公布日 2021.12.03

(73) 专利权人 阿里巴巴集团控股有限公司  
地址 英属开曼群岛大开曼资本大厦一座四层847号邮箱

审查员 何华兵

(72) 发明人 贺勇 李楠 曹徽

(74) 专利代理机构 北京太合九思知识产权代理有限公司 11610  
专利代理师 张爱 刘戈

(51) Int. Cl.

G16H 40/20 (2018.01)

G06F 40/30 (2020.01)

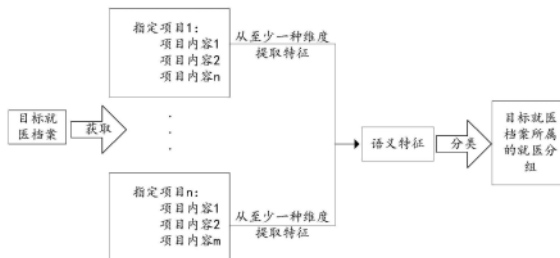
权利要求书4页 说明书15页 附图3页

(54) 发明名称

一种就医分组方法、设备及存储介质

(57) 摘要

本申请实施例提供一种就医分组方法、设备及存储介质,其中,所述方法包括:从目标就医档案中获取至少一个指定项目下的项目内容;根据至少一个指定项目下的项目内容,从至少一种维度,确定目标就医档案对应的语义特征;基于语义特征与就医分组之间的映射关系以及目标就医档案对应的语义特征,确定目标就医档案所属的就医分组。据此,本申请实施例中,对目标就医档案进行就医分组的过程不再偏向于某个医生的经验,而是从数据角度出发,客观地判断目标就医档案所属的就医分组,这样可综合考虑若干医生的经验而确定分组结果,从而提高分组的准确性和合理性。



1. 一种就医分组方法,其特征在于,包括:

从目标就医档案中获取至少一个指定项目下的项目内容;

根据所述至少一个指定项目下的项目内容,分别确定所述至少一个指定项目在至少一种维度下的特征向量;

将所述至少一个指定项目各自在所述至少一种维度下的特征向量进行组合,以确定所述目标就医档案对应的语义特征;

基于语义特征与就医分组之间的映射关系以及所述目标就医档案对应的语义特征,确定所述目标就医档案所属的就医分组;

其中,所述至少一种维度包括项目内维度,计算任一指定项目下的多个项目内容之间的相关度,并根据所述相关度构建所述指定项目在所述项目内维度下的特征向量。

2. 根据权利要求1所述的方法,其特征在于,所述计算任一指定项目下的多个项目内容之间的相关度,包括:

分别对所述至少一个指定项目下的项目内容进行向量化,以获得所述至少一个指定项目各自对应的编码矩阵;

采用内部注意力self-attention机制,对第一项目对应的编码矩阵进行处理,以获得用于表征所述第一项目下多个项目内容之间的相关度的第一矩阵;

其中,所述第一项目为任一指定项目。

3. 根据权利要求2所述的方法,其特征在于,所述根据所述相关度,构建所述指定项目在所述项目内维度下的特征向量,包括:

对所述第一矩阵执行池化处理,以将所述第一矩阵转换为所述第一项目在所述项目内维度下的特征向量。

4. 根据权利要求1所述的方法,其特征在于,所述至少一种维度包括项目间维度,所述根据所述至少一个指定项目下的项目内容,分别确定所述至少一个指定项目各自在所述至少一种维度下的特征向量,包括:

在所述指定项目为多个的情况下,计算所述多个指定项目之间的相关度;

根据所述多个指定项目之间的相关度,构建所述多个指定项目各自在所述项目间维度下的特征向量。

5. 根据权利要求4所述的方法,其特征在于,所述计算所述多个指定项目之间的相关度,包括:

分别对第一项目和第二项目下的项目内容分别进行向量化,以获得所述第一项目和所述第二项目各自对应的编码矩阵;

基于所述第一项目和所述第二项目各自对应的编码矩阵,计算所述第一项目下的各项目内容与所述第二项目下的各项目内容之间的相关度;

根据所述第一项目下的各项目内容与所述第二项目下的各项目内容之间的相关度,构建所述第一项目在所述项目间维度下的特征向量;

其中,所述第一项目和所述第二项目为所述多个指定项目中的任意两个。

6. 根据权利要求5所述的方法,其特征在于,所述基于所述第一项目和第二项目各自对应的编码矩阵,计算所述第一项目下的各项目内容与所述第二项目下的各项目内容之间的相关度,包括:

将所述第一项目对应的编码矩阵乘以预置的查询参数矩阵,以获得所述第一项目对应的查询矩阵;

将所述第二项目对应的编码矩阵乘以预置的键参数矩阵和值参数矩阵,以获得所述第二项目下对应的键矩阵和值矩阵;

基于所述第一项目对应的查询矩阵以及所述第二项目对应的键矩阵和值矩阵,计算用于表征所述第一项目和所述第二项目之间的相关度的第二矩阵。

7. 根据权利要求6所述的方法,其特征在于,所述基于所述第一项目对应的查询矩阵以及所述第二项目对应的键矩阵和值矩阵,计算用于表征所述第一项目和所述第二项目之间的相关度的第二矩阵,包括:

从所述第一项目对应的查询矩阵中获取第*i*个项目内容对应的目标查询向量,从所述第二项目对应的键矩阵和值矩阵中分别获取第*j*个项目内容对应的目标键向量和目标值向量,*i*和*j*为整数;

将所述目标查询向量和所述目标键向量进行点乘,以获得所述第一项目下第*i*个项目内容和所述第二项目下第*j*个项目内容之间的初步相关度因子;

将所述初步相关度因子构建为一中间向量;

将所述中间向量与所述目标值向量进行相乘,以获得所述第一项目下第*i*个项目内容与所述第二项目下各项目内容之间相关度的表征向量;

根据所述第一项目下各项目内容对应的表征向量,构建用于表征所述第一项目和所述第二项目之间的相关度的第二矩阵。

8. 根据权利要求7所述的方法,其特征在于,所述根据所述多个指定项目之间的相关度,构建所述多个指定项目各自在所述项目间维度下的特征向量,包括:

对所述第二矩阵进行池化处理,以获得所述第一项目在所述项目间维度下的特征向量。

9. 根据权利要求1所述的方法,其特征在于,所述至少一种维度包括知识维度,所述根据所述至少一个指定项目下的项目内容,分别确定所述至少一个指定项目各自在所述至少一种维度下的特征向量,包括:

计算第一项目对应的接口向量,所述接口向量中包含读向量;

从知识矩阵中,选择与所述读向量的相关度符合第一预设要求的目标行;

根据所述目标行,生成一输出向量,作为所述第一项目在所述知识维度下的特征向量。

10. 根据权利要求9所述的方法,其特征在于,所述知识矩阵存储在动态外部存储器中。

11. 根据权利要求9所述的方法,其特征在于,所述从知识矩阵中,选择与所述读向量的相关度符合第一预设要求的目标行,包括:

计算所述读向量和所述知识矩阵中的各行之间的点积;

对点积结果进行归一化,以获得所述读向量和所述知识矩阵中的各行之间的相关度;

将相关度最大的*N*行,作为所述目标行,*N*为正整数。

12. 根据权利要求9所述的方法,其特征在于,所述根据所述目标行,生成一输出向量,包括:

若所述目标行为多个,则基于所述多个目标行各自与所述读向量之间的相关度,对所述多个目标行进行加权求和,以获得所述输出向量。

13. 根据权利要求9所述的方法,其特征在於,所述接口向量中还包含写向量和用于表征所述第一项目中所携带的知识的写入值,在基于所述读向量,从预置的知识矩阵中,选择与所述第一项目的相关度符合预设要求的目标行之前,还包括:

基于所述写向量,从所述知识矩阵中,选择与所述写向量的相关度符合第二预设要求的待更新行;

根据所述写入值,对所述待更新行中的指定元素进行更新;

将更新后的知识矩阵,作为确定所述第一项目在该知识维度下的特征向量的依据。

14. 根据权利要求9所述的方法,其特征在於,所述计算第一项目对应的接口向量,所述接口向量中包含读向量,包括:

获取用于表征所述第一项目下各项目内容之间的相关度的第一矩阵;

将所述第一矩阵与预置的接口转换参数矩阵相乘,以获得所述接口向量;

按照预设的拆分规则,从所述接口向量中拆分出读向量、写向量和用于表征所述第一项目中所携带的知识的写入值。

15. 根据权利要求1-14任一项所述的方法,其特征在於,所述指定项目包括诊断项目、手术操作项目、过敏项目、遗传病项目、手术史项目中的一个或多个。

16. 根据权利要求1所述的方法,其特征在於,所述从目标就医档案中获取至少一个指定项目下的项目内容,包括:

将所述目标就医档案输入分组模型;

利用所述分组模型,从所述目标就医档案中获取至少一个指定项目下的项目内容;

所述根据所述至少一个指定项目下的项目内容,从至少一种维度,确定所述目标就医档案对应的语义特征,包括:

利用所述分组模型,从至少一种维度,确定所述目标就医档案对应的语义特征;

所述基于语义特征与就医分组之间的映射关系以及所述目标就医档案对应的语义特征,确定所述目标就医档案所属的就医分组,包括:

利用所述分组模型,基于语义特征与就医分组之间的映射关系以及所述目标就医档案对应的语义特征,确定所述目标就医档案所属的就医分组。

17. 根据权利要求16所述的方法,其特征在於,还包括:

获取训练就医档案样本以及所述训练就医档案样本所属的就医分组;

利用所述训练就医档案样本以及所述训练就医档案样本所属的就医分组训练所述分组模型。

18. 根据权利要求1所述的方法,其特征在於,在确定所述目标就医档案所属的就医分组之后,还包括:

基于就医分组与医保支付指标之间的对应关系,确定所述目标就医档案所属的就医分组对应的目标医保支付指标;

按照所述目标医保支付指标,计算所述目标就医档案对应的医保支付金额。

19. 一种计算设备,其特征在於,包括存储器和处理器;

所述存储器用于存储一条或多条计算机指令;

所述处理器与所述存储器耦合,用于执行所述一条或多条计算机指令,以用于:

从目标就医档案中获取至少一个指定项目下的项目内容;

根据所述至少一个指定项目下的项目内容,分别确定所述至少一个指定项目在至少一种维度下的特征向量;

将所述至少一个指定项目各自在所述至少一种维度下的特征向量进行组合,以确定所述目标就医档案对应的语义特征;

基于语义特征与就医分组之间的映射关系以及所述目标就医档案对应的语义特征,确定所述目标就医档案所属的就医分组;

其中,所述至少一种维度包括项目内维度,计算任一指定项目下的多个项目内容之间的相关度,并根据所述相关度构建所述指定项目在所述项目内维度下的特征向量。

20.一种存储计算机指令的计算机可读存储介质,其特征在于,当所述计算机指令被一个或多个处理器执行时,致使所述一个或多个处理器执行权利要求1-18任一项所述的就医分组方法。

## 一种就医分组方法、设备及存储介质

### 技术领域

[0001] 本申请涉及数据处理技术领域,尤其涉及一种就医分组方法、设备及存储介质。

### 背景技术

[0002] DRGs(Diagnosis-related Groups,诊断相关组)是医疗管理领域应用广泛的一种病例组合系统。所谓DRGs付费方式,是将临床过程相近、费用消耗相似的病例分到同一个DRG分组中进行管理的体系,并以组为单位制定医药费用标准进行付费。

[0003] 目前,医生需要基于自身的经验来确定就医档案所属的DRG分组,这过度依赖医生的自身经验。而不同医生的水平、经验是不一样的,甚至由于某些原因带来一些主观的判断,导致分组不合理,进而影响费用消耗。

### 发明内容

[0004] 本申请的多个方面提供一种就医分组方法、设备及存储介质,用以改善就医分组的准确性。

[0005] 本申请实施例提供一种就医分组方法,包括:

[0006] 从目标就医档案中获取至少一个指定项目下的项目内容;

[0007] 根据所述至少一个指定项目下的项目内容,从至少一种维度,确定所述目标就医档案对应的语义特征;

[0008] 基于语义特征与就医分组之间的映射关系以及所述目标就医档案对应的语义特征,确定所述目标就医档案所属的就医分组。

[0009] 本申请实施例还提供一种计算设备,包括存储器和处理器;

[0010] 所述存储器用于存储一条或多条计算机指令;

[0011] 所述处理器与所述存储器耦合,用于执行所述一条或多条计算机指令,以用于:

[0012] 从目标就医档案中获取至少一个指定项目下的项目内容;

[0013] 根据所述至少一个指定项目下的项目内容,从至少一种维度,确定所述目标就医档案对应的语义特征;

[0014] 基于语义特征与就医分组之间的映射关系以及所述目标就医档案对应的语义特征,确定所述目标就医档案所属的就医分组。

[0015] 本申请实施例还提供一种存储计算机指令的计算机可读存储介质,当所述计算机指令被一个或多个处理器执行时,致使所述一个或多个处理器执行前述的就医分组方法。

[0016] 在本申请实施例中,可从目标就医档案中获取至少一个指定项目下的项目内容;根据至少一个指定项目下的项目内容,从至少一种维度,确定目标就医档案对应的语义特征;基于语义特征与就医分组之间的映射关系以及目标就医档案对应的语义特征,确定目标就医档案所属的就医分组。据此,本申请实施例中,可基于目标就医档案中至少一个指定项目下的项目内容,挖掘出目标就医档案对应的语义特征,并基于挖掘出的语义特征确定目标就医档案所属的就医分组。这使得对目标就医档案进行就医分组的过程不再偏向于某

个医生的经验,而是从数据角度出发,客观地判断目标就医档案所属的就医分组,这样可综合考虑若干医生的经验而确定分组结果,从而提高分组的准确性和合理性。

### 附图说明

[0017] 此处所说明的附图用来提供对本申请的进一步理解,构成本申请的一部分,本申请的示意性实施例及其说明用于解释本申请,并不构成对本申请的不当限定。在附图中:

[0018] 图1为本申请一示例性实施例提供的一种就医分组方法的流程示意图;

[0019] 图2为本申请一示例性实施例提供的一种就医分组方法的逻辑示意图;

[0020] 图3为本申请一示例性实施例提供的一种语义特征确定方案的逻辑示意图;

[0021] 图4为本申请另一示例性实施例提供的一种计算设备的结构示意图。

### 具体实施方式

[0022] 为使本申请的目的、技术方案和优点更加清楚,下面将结合本申请具体实施例及相应的附图对本申请技术方案进行清楚、完整地描述。显然,所描述的实施例仅是本申请一部分实施例,而不是全部的实施例。基于本申请中的实施例,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,都属于本申请保护的范围。

[0023] 目前,通常依赖单个医生的经验确定就医档案的分组,导致分组不合理,进而影响费用消耗。为改善这些技术问题,本申请的一些实施例中:可从目标就医档案中获取至少一个指定项目下的项目内容;根据至少一个指定项目下的项目内容,从至少一种维度,确定目标就医档案对应的语义特征;基于语义特征与就医分组之间的映射关系以及目标就医档案对应的语义特征,确定目标就医档案所属的就医分组。据此,本申请实施例中,可基于目标就医档案中至少一个指定项目下的项目内容,挖掘出目标就医档案对应的语义特征,并基于挖掘出的语义特征确定目标就医档案所属的就医分组。这使得对目标就医档案进行就医分组的过程不再偏向于某个医生的经验,而是从数据角度出发,客观地判断目标就医档案所属的就医分组,这样可综合考虑若干医生的经验而确定分组结果,从而提高分组的准确性和合理性。

[0024] 以下结合附图,详细说明本申请各实施例提供的技术方案。

[0025] 图1为本申请一示例性实施例提供的一种就医分组方法的流程示意图。图2为本申请一示例性实施例提供的一种就医分组方法的逻辑示意图。本实施例提供的就医分组方法可以由一就医分组装置来执行,该就医分组装置可以实现为软件或实现为软件和硬件的组合,该就医分组装置可集成设置在计算设备中。如图1所示,该就医分组方法包括:

[0026] 步骤100、从目标就医档案中获取至少一个指定项目下的项目内容;

[0027] 步骤101、根据至少一个指定项目下的项目内容,从至少一种维度,确定目标就医档案对应的语义特征;

[0028] 步骤102、基于语义特征与就医分组之间的映射关系以及目标就医档案对应的语义特征,确定目标就医档案所属的就医分组。

[0029] 本实施例提供的就医分组方法可应用于各种需要对就医档案进行分组的场景中,例如,医保支付场景、住院管理场景等等,本实施例对应用场景不做限定。

[0030] 另外,本实施例输出的就医分组结果,可以提供给保险公司或医疗卫生管理机构

用于对费用的监管等。当然,这里只是提供了输出的就医分组结果的示例性应用,本实施例并不限于此,本实施例输出的就医分组结果,还可应用于当前或者未来可能需要以就医分组为基础的任何其它情况。

[0031] 本实施例中,病人的每次就医,都会对应一个就医档案。实际应用中,就医档案可由医生或护理人员等进行制备。

[0032] 本实施例中,可在病人每次就医时,按照当次就医的就医档案进行就医分组。也即是,对于病人的每次就医,都会进行一次就医分组。

[0033] 在步骤100中,可从目标就医档案中获取至少一个指定项目下的项目内容。目标就医档案可以是目标病人本次的就医档案。

[0034] 其中,就医档案可以是指按规范记录病人疾病表现和诊疗情况等信息的档案。就医档案中通常包括多个项目,而单个项目下可包含一条或多条项目内容。

[0035] 针对不同的项目,项目下的项目内容可能不完全相同。例如,针对诊断项目,其下的项目内容可能包括高血压、高血脂、糖尿病等等,而针对手术操作项目,其下的项目内容可能包括腹腔镜手术、血管支架手术等等。本实施例对单个项目下的项目内容的数量及类型不作限定。

[0036] 本实施例中,指定项目可以是诊断项目、手术操作项目、过敏项目、遗传病项目或手术史项目等,当然,这些仅是示例性的,本实施例并不限于此,指定项目可以是就医档案中的任意项目。另外,本实施例中,指定项目的数量可以是一个或多个。

[0037] 实际应用中,可预先确定指定项目,例如,可指定诊断项目和手术操作项目作为指定项目。据此,在步骤100中,针对不同的就医档案,可统一获取指定项目下的项目内容,以保持分组依据的一致性。

[0038] 在步骤101中,可根据至少一个指定项目下的项目内容,从至少一种维度,确定目标就医档案对应的语义特征。

[0039] 其中,语义特征用于表征目标就医档案中包含的语义信息。实际应用中,语义特征可采用向量作为表征形式,当然,本实施例并不限于此。

[0040] 本实施例中,语义特征中可包含至少一种维度的语义信息。另外,维度可以是项目内维度、项目间维度或知识维度等。

[0041] 例如,在指定项目为一个的情况下,可根据该指定项目下的项目内容,从项目内维度和知识维度,确定该目标就医档案对应的语义特征。而在指定项目为多个的情况下,则可根据该指定项目下的项目内容,从项目内维度、项目间维度和知识维度,分别确定多个目标就医档案各自对应的语义特征。

[0042] 当然,这仅是示例性的,本实施例中,在确定目标就医档案对应的语义特征的过程中,可根据实际需要确定所采用的维度,本实施例对此不作限定。

[0043] 基于此,在步骤102中,可基于语义特征与就医分组之间的映射关系以及目标就医档案对应的语义特征,确定目标就医档案所属的就医分组。

[0044] 本实施例中,可预先构建语义特征与就医分组之间的映射关系。实际应用中,可采用机器学习模型,从大量的训练病例样本中学习到语义特征与就医分组之间的映射关系,这可有效保证分组结果的客观性。

[0045] 不同就医分组对应的语义特征可能不完全相同,这样,在确定出目标就医档案对



应的语义特征的基础上,可查找该语义特征对应的就医分组,从而可将超找到的就医分组,作为目标就医档案所属的就医分组。

[0046] 本实施例中,可采用有限数量的预置就医分组,例如,可继承DRGs中的DRG分组作为本实施例中的就医分组,当然,还可采用其它方式预置就医分组,例如,可预先采用无监督学习等方式,对大量的训练病例样本进行聚类,并产生的若干类作为就医分组,等。本实施例并不限于此。

[0047] 本实施例中,在确定出目标就医档案所属的就医分组后,还可基于该就医分组进行后续环节的处理。在不同的应用场景中,后续环节可能不完全相同。

[0048] 例如,在医保支付场景中,可基于就医分组与医保支付指标之间的对应关系,确定目标就医档案所属的就医分组对应的目标医保支付指标;按照目标医保支付指标,计算目标就医档案对应的医保支付金额。

[0049] 综上,本实施例中,可从目标就医档案中获取至少一个指定项目下的项目内容;根据至少一个指定项目下的项目内容,从至少一种维度,确定目标就医档案对应的语义特征;基于语义特征与就医分组之间的映射关系以及目标就医档案对应的语义特征,确定目标就医档案所属的就医分组。据此,本申请实施例中,可基于目标就医档案中至少一个指定项目下的项目内容,挖掘出目标就医档案对应的语义特征,并基于挖掘出的语义特征确定目标就医档案所属的就医分组。这使得对目标就医档案进行就医分组的过程不再偏向于某个医生的经验,而是从数据角度出发,客观地判断目标就医档案所属的就医分组,这样可综合考虑若干医生的经验而确定分组结果,从而提高分组的准确性和合理性。

[0050] 在上述或下述实施例中,可根据至少一个指定项目下的项目内容,分别确定至少一个指定项目各自在至少一种维度下的特征向量;将至少一个指定项目各自在至少一种维度下的特征向量进行组合,以生成目标就医档案对应的语义特征。

[0051] 图3为本申请一示例性实施例提供的一种语义特征确定方案的逻辑示意图。

[0052] 参考图3,可将至少一个指定项目各自在至少一种维度下的特征向量进行拼接,以生成目标就医档案对应的语义特征。例如,可采用concat操作,将至少一个指定项目各自在至少一种维度下的特征向量进行拼接。

[0053] 当然,本实施例中特征向量的组合方式并不限于拼接,还可采用其它组合方式。另外,多个特征向量之间的组合顺序也不做限定。

[0054] 实际应用中,可预先配置特征向量之间的组合方式及组合顺序。据此,在执行组合操作的过程中,针对不同的就医档案,可采用固定的组合方式及组合顺序,以保证语义特征的格式一致性。

[0055] 另外,值得说明的是,图3中示出了两个指定项目(诊断集合对应于诊断项目,手术操作集合对应于手术操作项目)的情况,并将两个指定项目各自在至少一种维度下的特征向量进行组合后,生成目标就医档案对应的语义特征。但应当理解的是,指定项目的数量并不限于图3中示出的两个,指定项目的数量还可以更多。

[0056] 由于本实施例中针对至少一个指定项目的特征向量的确定方案相似,为便于描述,以下将以第一项目为例进行特征向量的确定过程的描述,应当理解的是,第一项目可以是至少一个指定项目中的任意一个。

[0057] 本实施例中,可根据第一项目下的项目内容,确定第一项目在至少一个维度下的

特征向量。

[0058] 正如前文提及的,维度可以是项目内维度、项目间维度或知识维度。以下将以这几种示例性维度为例,分别说明不同维度下的特征向量的确定方案。

[0059] 项目内维度

[0060] 对于项目内维度,可在第一项目下的项目内容为多个的情况下,计算第一项目下的多个项目内容之间的相关度;根据第一项目下的多个项目内容之间的相关度,构建第一项目在项目内维度下的特征向量。

[0061] 本实施例中,基于项目内维度,可挖掘出第一项目下的项目内容之间可能存在的相互作用信息。这些信息可能影响目标就医档案的分组结果。

[0062] 例如,对于诊断项目,其下的心脏增大可能是高血压的并发症,而其下的视网膜病变可能是糖尿病的症状等等。

[0063] 本实施例中,可计算第一项目下的多个项目内容之间的相关度,以利用相关度表征第一项目下的项目内容之间可能存在的相互作用信息。

[0064] 实际应用中,可分别对至少一个指定项目下的项目内容进行向量化,以获得至少一个指定项目各自对应的编码矩阵;采用内部注意力self-attention机制,对第一项目对应的编码矩阵进行处理,以获得用于表征第一项目下多个项目内容之间的相关度的第一矩阵。

[0065] 其中,参考图3,向量化的过程可以是:采用Embedding技术,对第一项目下的每个项目内容进行d维向量化编码,从而得到第一项目对应的编码矩阵。例如,可将项目内容进行100维向量化编码。如图3所示,可获的诊断项目对应的编码矩阵Embedding1,手术操作项目对应的编码矩阵Embedding2。

[0066] 在第一项目下包含n个项目内容的情况下,可通过向量化操作获得 $n*d$ 的编码矩阵。

[0067] self-attention(又名intra-attention),是一种求同列表元素之间的相关度的注意力模型,可针对列表中每一个元素,计算其与同列表中其他元素之间的相关度。

[0068] 本实施例中,基于前述的向量化操作,可获得第一项目下的各项目内容对应的向量化编码x,针对每一个项目内容可分别乘三个预置的参数矩阵:查询参数矩阵 $W_q$ 、键参数矩阵 $W_k$ 和值参数矩阵 $W_v$ ,得到查询向量q、键向量k和值向量v。其中,查询向量q、键向量k和值向量v均来自同一输入,也即第一项目对应的编码矩阵。

[0069] 值得说明的是,三个预置的参数矩阵 $W_q$ 、 $W_k$ 和 $W_v$ 以及查询向量q、键向量k和值向量v为注意力attention机制下的基础概念,在此不再展开详述。

[0070] 据此,第一项目下的每个项目内容均可得到查询向量q、键向量k和值向量v,进而可构建第一项目对应的查询矩阵Q、键矩阵K和值矩阵V。

[0071] 之后,可基于第一项目对应的查询矩阵、键矩阵和值矩阵,计算用于表征第一项目下多个项目内容之间的相关度的第一矩阵。

[0072] 实际应用中,可按照如下公式计算出用于表征第一项目下多个项目内容之间的相关度的第一矩阵 $Z_1$ :

[0073]  $Z_1 = \text{Softmax} \left( \frac{Q * K^T}{\sqrt{d_k}} * V \right)$ ,其中,Q为第一项目对应的查询矩阵,K为第一项目对

应的键矩阵,  $V$ 为第一项目对应的值矩阵,  $\sqrt{d_k}$ 为根据第一项目对应的编码矩阵而确定的常数。

[0074] 以上是从矩阵的计算维度阐述了第一矩阵 $Z_1$ 的确定过程。

[0075] 以下将从项目内容对应的查询向量 $q$ 、键向量 $k$ 和值向量 $v$ 的计算维度说明第一矩阵 $Z_1$ 的确定过程。

[0076] 以第一项目下的第 $i$ 个项目内容和第 $j$ 个项目内容为例, 其中,  $i = (1, 2, \dots, n)$ ,  $j = (1, 2, \dots, n)$ ,  $n$ 为第一项目下的项目内容的数量:

[0077] 首先, 可计算 $q_i$ 与 $k_j$ 的点积, 然后除以一个常数 $\sqrt{d_k}$ 得到初步相关度因子 $s_{ij} = q_i * k_j$ ;

[0078] 这样, 可得到了第 $i$ 个项目内容与其他所有项目内容(包括第 $i$ 个项目内容自身)之间的初步相关度因子, 将这些初步相关度因子构建为向量 $s_i = \{s_{i1}, s_{i2}, \dots, s_{in}\}$ , 并可对 $s_i$ 进行归一化, 得到 $new-s_i$ , 将 $new-s_i$ 作为中间向量;

[0079] 之后, 可将 $new-s_i$ 与第 $i$ 个项目内容对应的值向量相乘, 得到用于表征第 $i$ 个项目内容与其它项目内容之间的相关度的表征向量 $X_i$ ;

[0080] 最后, 可将 $n$ 个项目内容的表征向量组合(例如, 相加), 以获得第一矩阵 $Z_1$ 。

[0081] 至此, 可计算出第一项目下多个项目内容之间的相似度, 并表征为第一矩阵 $Z_1$ 。

[0082] 在此基础上, 可对第一矩阵执行池化处理, 以将第一矩阵转换为第一项目在项目内维度下的特征向量。例如, 可对第一矩阵中的每一列进行pooling处理, 以将第一矩阵转换为一维向量, 作为第一项目在项目内维度下的特征向量。

[0083] 参考图3, 用于表征诊断项目内多个项目内容之间相关度的第一矩阵为 $X_{11}$ , 诊断项目在项目内维度下的特征向量为 $V_{11}$ , 用于表征手术操作项目内多个项目内容之间相关度的第一矩阵为 $X_{22}$ , 手术操作项目在项目内维度下的特征向量为 $V_{22}$ 。

[0084] 据此, 在项目内维度下, 可挖掘单个指定项目内各项目内容之间的相互作用信息, 将这些信息携带在语义特征中, 可有效提高分组结果的准确性和合理性。

[0085] 项目间维度

[0086] 对于项目间维度, 可在指定项目为多个的情况下, 计算多个指定项目之间的相关度; 根据多个指定项目之间的相关度, 构建多个指定项目各自在项目间维度下的特征向量。

[0087] 本实施例中, 基于项目间维度, 可挖掘出指定项目之间可能存在的相互作用信息。这些信息可能影响目标就医档案的分组结果。

[0088] 例如, 对于手术操作项目, 其下的射频消融手术可能是诊断项目下的肿瘤的必要处理, 等等。

[0089] 本实施例中, 可计算多个指定项目之间的相关度, 以利用相关度表征多个指定项目之间可能存在的相互作用信息。

[0090] 可选地, 可分别对第一项目和第二项目下的项目内容分别进行向量化, 以获得第一项目和第二项目各自对应的编码矩阵; 基于第一项目和第二项目各自对应的编码矩阵, 计算第一项目下的各项目内容与第二项目下的各项目内容之间的相关度。

[0091] 其中, 第一项目和第二项目可以是多个指定项目中的任意两个。值得说明的是, 第二项目可以是第一项目本身。

[0092] 参考图3,向量化的过程可以是:采用Embedding技术,对第一项目和第二项目下的每个项目内容进行d维向量化编码,从而得到第一项目和第二项目各自对应的编码矩阵。例如,可将项目内容进行100维向量化编码。参考图3,可获得诊断项目对应的编码矩阵Embedding1,手术操作项目对应的编码矩阵Embedding2。

[0093] 在第一项目下包含n个项目内容,的情况下,可通过向量化操作获得 $n*d$ 的编码矩阵。在第二项目下包含m个项目内容,的情况下,可通过向量化操作获得 $m*d$ 的编码矩阵。

[0094] 本实施例中,基于前述的向量化操作,可获得第一项目下的各项目内容对应的向量化编码以及第二项目下的各项目内容对应的向量化编码,针对每一个项目内容可分别乘三个预置的参数矩阵:查询参数矩阵 $W_q$ 、键参数矩阵 $W_k$ 和值参数矩阵 $W_v$ ,得到查询向量q、键向量k和值向量v。

[0095] 值得说明的是,三个预置的参数矩阵 $W_q$ 、 $W_k$ 和 $W_v$ 以及查询向量q、键向量k和值向量v为注意力attention机制下的基础概念,在此不再展开详述。

[0096] 据此,第一项目下的每个项目内容均可得到查询向量q、键向量k和值向量v,进而可构建第一项目对应的查询矩阵 $Q_1$ 、键矩阵 $K_1$ 和值矩阵 $V_1$ 。第二项目下的每个项目内容均可得到查询向量q、键向量k和值向量v,进而可构建第二项目对应的查询矩阵 $Q_2$ 、键矩阵 $K_2$ 和值矩阵 $V_2$ 。

[0097] 之后,可基于第一项目对应的查询矩阵以及第二项目对应的键矩阵和值矩阵,计算用于表征第一项目和第二项目之间的相关度的第二矩阵。

[0098] 这与前述的项目内维度下的self attention机制不同。两者的不同在于,self attention机制使用的是来自同一输入(指定项目对应的编码矩阵)的查询矩阵、键矩阵和值矩阵,而项目间维度下,使用的查询矩阵、键矩阵和值矩阵则并非来自同一输入,而是使用来自第一项目的查询矩阵和来自第二项目的查询矩阵、键矩阵和值矩阵。本文中,将这种注意力机制描述为相互注意inter attention机制。

[0099] 实际应用中,可按照如下公式计算出用于表征第一项目和第二项目之间的相关度的第二矩阵 $Z_2$ :

[0100]  $Z_2 = \text{Softmax} \left( \frac{Q_1 * K_2^T}{\sqrt{d_k}} * V_2 \right)$ ,其中, $Q_1$ 为第一项目对应的查询矩阵, $K_2$ 为第二项目

对应的键矩阵, $V_2$ 为第二项目对应的值矩阵, $\sqrt{d_k}$ 为根据第一项目和第二项目给咱对应的编码矩阵而确定的常数。

[0101] 以上是从矩阵的计算维度阐述了第二矩阵 $Z_2$ 的确定过程。

[0102] 以下将从项目内容对应的查询向量q、键向量k和值向量v的计算维度说明第二矩阵 $Z_2$ 的确定过程。

[0103] 在项目间维度下,可使用第一项目下单个项目内容对应的查询向量q,去与第二项目下的各项目内容对应的键向量k进行比对,以确定双方之间的相关度。

[0104] 以第一项目下的第i个项目内容和第二项目下的第j个项目内容为例,其中, $i = (1, 2, \dots, n)$ , $j = (1, 2, \dots, m)$ ,n为第一项目下的项目内容的数量,m为第二项目下的项目内容的数量:

[0105] 首先,可计算 $q_i$ 与 $k_j$ 的点积,然后除以一个常数 $\sqrt{d_k}$ 得到初步相关度因子 $s_{ij}=q_i*k_j$ ,其中, $q_i$ 是第一项目下第*i*个项目内容对应的查询向量, $k_j$ 是第二项目下第*j*个项目内容对应的键向量;

[0106] 这样,可得到了第一项目下第*i*个项目内容与第二项目下所有项目内容之间的初步相关度因子,将这些初步相关度因子构建为向量 $s_i = \{s_{i1}, s_{i2}, \dots, s_{in}\}$ ,并可对 $s_i$ 进行归一化,得到 $new-s_i$ ,将 $new-s_i$ 作为中间向量;

[0107] 之后,可将 $new-s_i$ 与第二项目下的第*j*个项目内容对应的值向量 $v_j$ 相乘,得到用于表征第一项目下第*i*个项目内容与第二项目下各项目内容之间的相关度的表征向量 $X_i$ ;

[0108] 最后,可将第一项目下*n*个项目内容的表征向量组合(例如,相加),以获得第二矩阵 $Z_2$ 。

[0109] 至此,可计算出第一项目与第二项目之间的相似度,并表征为第一项目对应的第二矩阵。

[0110] 同理,可计算出第二项目与第一项目之间的相似度,并表征为第二项目对应的第二矩阵。

[0111] 另外,以上是针对两个指定项目之间的相似度的计算方案,针对更多指定项目之间的相似度方案可进行适应性扩展,例如,针对第一项目,可分别计算用于表征第一项目与第二项目之间的相似度的矩阵A,用于表征第一项目与第三项目之间相似度的矩阵B,并将A和B组合后,作为用于表征第一项目与第二项目及第三项目之间的相似度的第二矩阵。当然,这仅是示例性的,本实施例并不限于此。

[0112] 在此基础上,可对第一项目对应的第二矩阵执行池化处理,以将第二矩阵转换为第一项目在项目间维度下的特征向量。例如,可对第一项目对应的第二矩阵中的每一列进行pooling处理,以将第一项目对应的第二矩阵转换为一维向量,作为第一项目在项目间维度下的特征向量。

[0113] 参考图3,用于表征诊断项目与手术操作项目之间相关度的第二矩阵为 $X_{12}$ ,诊断项目在项目间维度下的特征向量为 $V_{12}$ ,用于表征手术操作项目与诊断项目之间相关度的第二矩阵为 $X_{21}$ ,手术操作项目在项目间维度下的特征向量为 $V_{21}$ 。

[0114] 据此,在项目间维度下,可挖掘指定项目之间的相互作用信息,将这些信息携带在语义特征中,可有效提高分组结果的准确性和合理性。

[0115] 知识维度

[0116] 对于知识维度,可计算第一项目对应的接口向量。

[0117] 参考图3,本实施例中,可复用前述的项目内维度下,所生成的第一项目对应的第一矩阵;将第一矩阵与预置的接口转换参数矩阵相乘,以获得接口向量。其中,接口转换参数矩阵可从大量的训练就医档案样本中学习而得到。

[0118] 当然,实际应用中,将第一矩阵与预置的接口转换参数矩阵相乘后,还可对相乘结果执行池化等操作,以产生接口向量,本实施例对此不做限定。

[0119] 基于此,可按照预置的拆分规则,从接口向量中拆分出读向量、写向量和用于表征第一项目中所携带的知识的写入值。

[0120] 例如,可在拆分规则中规定接口向量中的哪些维作为读向量,哪些维作为写向量等,从而可根据拆分规则,确定出第一项目对应的读向量、写向量和写入值。

[0121] 本实施例中,可利用知识矩阵承载单个指定项目下的通用或专业知识。其中,知识矩阵可存储在动态外部存储器中,另外,不同指定项目对应的知识矩阵可不完全相同。

[0122] 本实施例中,知识矩阵的每一行可由键向量 $k$ 和值向量 $v$ 组成,基于此,本实施例中,可对知识矩阵采用先写后读的机制,获取第一项目在知识维度下的特征向量。

[0123] 写过程可以是:基于写向量,从知识矩阵中,选择与写向量的相关度符合第二预设要求的待更新行;根据写入值,对待更新行中的指定元素进行更新;将更新后的知识矩阵,作为确定第一项目在知识维度下的特征向量的依据。

[0124] 实际应用中,可使用写向量与知识矩阵的每一行的键向量 $k$ 进行点积,并对点积结果进行归一化,以获得写向量与知识矩阵中每一行之间的相关度;将相关度最大的行作为待更新行,按照写入值,对待更新行进行更新。其中,前述的拆分规则中还可指定至少一个写入位置,此处可按照写入值,在待更新行中的至少一个写入位置进行更新。

[0125] 读过程可以是:,从更新后的知识矩阵中,选择与读向量的相关度符合第一预设要求的目标行;根据目标行,生成一输出向量,作为第一项目在知识维度下的特征向量。

[0126] 实际应用中,可计算读向量和知识矩阵中的各行之间的点积;对点积结果进行归一化,以获得读向量和知识矩阵中的各行之间的相关度;将相关度最大的 $N$ 行,作为目标行, $N$ 为正整数。

[0127] 例如, $N$ 可设置为3。据此,可从更新后的知识矩阵中确定出3个目标行。

[0128] 当然, $N$ 的取值并不限于此。

[0129] 在此基础上,可基于多个目标行各自与读向量之间的相关度,对多个目标行进行加权求和,以获得输出向量。

[0130] 例如,目标行为 $m_1$ 、 $m_2$ 和 $m_3$ ,其中,目标行 $m_1$ 、 $m_2$ 和 $m_3$ 与读向量之间的相关度分别为 $0.4$ 、 $0.2$ 、 $0.1$ ,则输出向量 $=0.4*m_1+0.2*m_2+0.1*m_3$ 。此处的求和可以是将各目标行上相同位置上的元素值进行加权求和,以获得输出向量;也可以是分别对多个目标行上的元素值进行加权,而后再将多个加权后的目标行进行组合(例如,叠放或拼接),以获得输出向量,在此不做限定。

[0131] 参考图3,在知识维度下,可基于诊断集合对应的第一矩阵 $X_{11}$ ,生成接口向量 $interface\ v_1$ ,并可基于接口向量 $interface\ v_1$ ,对诊断集合对应的知识矩阵Dynamic External Memory 1执行先写后读操作,从而获得诊断集合在知识维度下的特征向量 $r_1$ 。

[0132] 据此,在知识维度下,可根据目标就医档案中单个指定项目携带的知识,更新该指定项目对应的知识矩阵,并从更新后的知识矩阵中抓取与该指定项目相关的知识,表征为知识维度下的特征向量。这可将目标就医档案中各指定项目对应的潜在知识携带在语义特征中,可有效提高分组结果的准确性和合理性。

[0133] 以上,分别从几个示例性维度,描述了特征向量的确定方案。

[0134] 在此基础上,本实施例中,可将各指定项目在各维度下的特征向量进行组合,以生成目标就医档案对应的语义特征。

[0135] 参考图3,可将前述获得的特征向量 $V_{11}$ 、 $V_{12}$ 、 $V_{22}$ 、 $V_{21}$ 、 $r_1$ 和 $r_2$ 进行拼接,以获得语义特征。

[0136] 据此,本实施例中,可从各种维度,提取目标就医档案对应的语义特征,使得语义特征可携带丰富语义信息,使得目标就医档案的数据化表征更加准确,这可有效提高分组

的准确性和合理性。

[0137] 在上述或下述实施例中,可采用预训练的分组模型来执行就医分组方法中的各操作。

[0138] 基于此,可将目标就医档案输入分组模型;

[0139] 利用分组模型,从目标就医档案中获取至少一个指定项目下的项目内容;

[0140] 利用分组模型,从至少一种维度,确定目标就医档案对应的语义特征;

[0141] 利用分组模型,基于语义特征与就医分组之间的映射关系以及目标就医档案对应的语义特征,确定目标就医档案所属的就医分组。

[0142] 这样,在实际应用中,可将目标就医档案输入分组模型,即可由分组模型输出目标就医档案对应的就医分组。

[0143] 在分组模型内部,可按照就医分组方法的各操作,设置对应的处理模块,以分别用于对目标就医档案进行项目内容获取、确定语义特征及分组等处理。

[0144] 另外,可采用端到端的训练方式,对分组模型进行预训练。

[0145] 实际应用中,可获取训练就医档案样本以及训练就医档案样本所属的就医分组;利用训练就医档案样本以及训练就医档案样本所属的就医分组训练分组模型。

[0146] 其中,前文提及的各种需要预置的模型参数,例如,预置的查询参数矩阵、键参数矩阵、值参数矩阵、接口转换参数矩阵以及语义特征与就医分组之间的映射关系等等,均可在预训练阶段中产生。

[0147] 可选地,可对每个训练就医档案样本使用cross-entropy loss技术执行模型训练过程。当然,本实施例并不限于此。

[0148] 需要说明的是,上述实施例所提供方法的各步骤的执行主体均可以是同一设备,或者,该方法也由不同设备作为执行主体。

[0149] 另外,在上述实施例及附图中的描述的一些流程中,包含了按照特定顺序出现的多个操作,但是应该清楚了解,这些操作可以不按照其在本文中出现的顺序来执行或并行执行,操作的序号如100、101等,仅仅是用于区分各个不同的操作,序号本身不代表任何的执行顺序。另外,这些流程可以包括更多或更少的操作,并且这些操作可以按顺序执行或并行执行。需要说明的是,本文中的“第一”、“第二”等描述,是用于区分不同的消息、设备、模块等,不代表先后顺序,也不限定“第一”和“第二”是不同的类型。

[0150] 图4为本申请另一示例性实施例提供的一种计算设备的结构示意图。如图4所示,该计算设备包括存储器40和处理器41;

[0151] 存储器40用于存储一条或多条计算机指令;

[0152] 处理器41与存储器41耦合,用于执行一条或多条计算机指令,以用于:

[0153] 从目标就医档案中获取至少一个指定项目下的项目内容;

[0154] 根据至少一个指定项目下的项目内容,从至少一种维度,确定目标就医档案对应的语义特征;

[0155] 基于语义特征与就医分组之间的映射关系以及目标就医档案对应的语义特征,确定目标就医档案所属的就医分组。

[0156] 在一可选实施例中,处理器41在根据至少一个指定项目下的项目内容,从至少一种维度,确定目标就医档案对应的语义特征时,用于:

[0157] 根据至少一个指定项目下的项目内容,分别确定至少一个指定项目各自在至少一种维度下的特征向量;

[0158] 将至少一个指定项目各自在至少一种维度下的特征向量进行组合,以生成目标就医档案对应的语义特征。

[0159] 在一可选实施例中,至少一种维度包括项目内维度,处理器41在根据至少一个指定项目下的项目内容,分别确定至少一个指定项目各自在至少一种维度下的特征向量时,用于:

[0160] 在第一项目下的项目内容为多个的情况下,计算第一项目下的多个项目内容之间的相关度;

[0161] 根据第一项目下的多个项目内容之间的相关度,构建第一项目在项目内维度下的特征向量;

[0162] 其中,第一项目为至少一个指定项目中的任意一个。

[0163] 在一可选实施例中,处理器41在计算第一项目下的多个项目内容之间的相关度时,用于:

[0164] 分别对至少一个指定项目下的项目内容进行向量化,以获得至少一个指定项目各自对应的编码矩阵;

[0165] 采用内部注意力self-attention机制,对第一项目对应的编码矩阵进行处理,以获得用于表征第一项目下多个项目内容之间的相关度的第一矩阵。

[0166] 在一可选实施例中,处理器41在根据第一项目下的多个项目内容之间的相关度,构建第一项目在项目内维度下的特征向量时,用于:

[0167] 对第一矩阵执行池化处理,以将第一矩阵转换为第一项目在项目内维度下的特征向量。

[0168] 在一可选实施例中,至少一种维度包括项目间维度,处理器41在根据至少一个指定项目下的项目内容,分别确定至少一个指定项目各自在至少一种维度下的特征向量时,用于:

[0169] 在指定项目为多个的情况下,计算多个指定项目之间的相关度;

[0170] 根据多个指定项目之间的相关度,构建多个指定项目各自在项目间维度下的特征向量。

[0171] 在一可选实施例中,处理器41在计算多个指定项目之间的相关度时,用于:

[0172] 分别对第一项目和第二项目下的项目内容分别进行向量化,以获得第一项目和第二项目各自对应的编码矩阵;

[0173] 基于第一项目和第二项目各自对应的编码矩阵,计算第一项目下的各项目内容与第二项目下的各项目内容之间的相关度;

[0174] 根据第一项目下的各项目内容与第二项目下的各项目内容之间的相关度,构建第一项目在项目间维度下的特征向量;

[0175] 其中,第一项目和第二项目为多个指定项目中的任意两个。

[0176] 在一可选实施例中,处理器41在基于第一项目和第二项目各自对应的编码矩阵,计算第一项目下的各项目内容与第二项目下的各项目内容之间的相关度时,用于:

[0177] 将第一项目对应的编码矩阵乘以预置的查询参数矩阵,以获得第一项目对应的查



询矩阵；

[0178] 将第二项目对应的编码矩阵乘以预置的键参数矩阵和值参数矩阵,以获得第二项目下对应的键矩阵和值矩阵；

[0179] 基于第一项目对应的查询矩阵以及第二项目对应的键矩阵和值矩阵,计算用于表征第一项目和第二项目之间的相关度的第二矩阵。

[0180] 在一可选实施例中,处理器41在基于第一项目对应的查询矩阵以及第二项目对应的键矩阵和值矩阵,计算用于表征第一项目和第二项目之间的相关度的第二矩阵时,用于:

[0181] 从第一项目对应的查询矩阵中获取第*i*个项目内容对应的目标查询向量,从第二项目对应的键矩阵和值矩阵中分别获取第*j*个项目内容对应的目标键向量和目标值向量,*i*和*j*为整数；

[0182] 将目标查询向量和目标键向量进行点乘,以获得第一项目下第*i*个项目内容和第二项目下第*j*个项目内容之间的初步相关度因子；

[0183] 将初步相关度因子构建为一中间向量；

[0184] 将中间向量与目标值向量进行相乘,以获得第一项目下第*i*个项目内容与第二项目下各项目内容之间相关度的表征向量；

[0185] 根据第一项目下各项目内容对应的表征向量,构建用于表征第一项目和第二项目之间的相关度的第二矩阵。

[0186] 在一可选实施例中,处理器41在根据多个指定项目之间的相关度,构建多个指定项目各自在项目间维度下的特征向量时,用于:

[0187] 对第二矩阵进行池化处理,以获得第一项目在项目间维度下的特征向量。

[0188] 在一可选实施例中,至少一种维度包括知识维度,处理器41在根据至少一个指定项目下的项目内容,分别确定至少一个指定项目各自在至少一种维度下的特征向量时,用于:

[0189] 计算第一项目对应的接口向量,接口向量中包含读向量；

[0190] 从知识矩阵中,选择与读向量的相关度符合第一预设要求的目标行；

[0191] 根据目标行,生成一输出向量,作为第一项目在知识维度下的特征向量。

[0192] 在一可选实施例中,知识矩阵存储在动态外部存储器中。

[0193] 在一可选实施例中,处理器41在从知识矩阵中,选择与读向量的相关度符合第一预设要求的目标行时,用于:

[0194] 计算读向量和知识矩阵中的各行之间的点积；

[0195] 对点积结果进行归一化,以获得读向量和知识矩阵中的各行之间的相关度；

[0196] 将相关度最大的*N*行,作为目标行,*N*为正整数。

[0197] 在一可选实施例中,处理器41在根据目标行,生成一输出向量时,用于:

[0198] 若目标行为多个,则基于多个目标行各自与读向量之间的相关度,对多个目标行进行加权求和,以获得输出向量。

[0199] 在一可选实施例中,接口向量中还包含写向量和用于表征第一项目中所携带的知识的写入值,处理器41在基于读向量,从预置的知识矩阵中,选择与第一项目的相关度符合预设要求的目标行之前,还用于:

[0200] 基于写向量,从知识矩阵中,选择与写向量的相关度符合第二预设要求的待更新

行;

- [0201] 根据写入值,对待更新行中的指定元素进行更新;
- [0202] 将更新后的知识矩阵,作为确定第一项目在知识维度下的特征向量的依据。
- [0203] 在一可选实施例中,处理器41在计算第一项目对应的接口向量,接口向量中包含读向量时,用于:
- [0204] 获取用于表征第一项目下各项目内容之间的相关度的第一矩阵;
- [0205] 将第一矩阵与预置的接口转换参数矩阵相乘,以获得接口向量;
- [0206] 按照预设的拆分规则,从接口向量中拆分出读向量、写向量和用于表征第一项目中所携带的知识的写入值。
- [0207] 在一可选实施例中,指定项目包括诊断项目、手术操作项目、过敏项目、遗传病项目、手术史项目中的一个或多个。
- [0208] 在一可选实施例中,处理器41在从目标就医档案中获取至少一个指定项目下的项目内容时,用于:
- [0209] 将目标就医档案输入分组模型;
- [0210] 利用分组模型,从目标就医档案中获取至少一个指定项目下的项目内容;
- [0211] 根据至少一个指定项目下的项目内容,从至少一种维度,确定目标就医档案对应的语义特征时,用于:
- [0212] 利用分组模型,从至少一种维度,确定目标就医档案对应的语义特征;
- [0213] 基于语义特征与就医分组之间的映射关系以及目标就医档案对应的语义特征,确定目标就医档案所属的就医分组时,用于:
- [0214] 利用分组模型,基于语义特征与就医分组之间的映射关系以及目标就医档案对应的语义特征,确定目标就医档案所属的就医分组。
- [0215] 在一可选实施例中,处理器41还用于:
- [0216] 获取训练就医档案样本以及训练就医档案样本所属的就医分组;
- [0217] 利用训练就医档案样本以及训练就医档案样本所属的就医分组训练分组模型。
- [0218] 在一可选实施例中,处理器41在确定目标就医档案所属的就医分组之后,还用于:
- [0219] 基于就医分组与医保支付指标之间的对应关系,确定目标就医档案所属的就医分组对应的目标医保支付指标;
- [0220] 按照目标医保支付指标,计算目标就医档案对应的医保支付金额。
- [0221] 值得说明的是,上述关于计算设备的各实施例中的技术细节,可参考前述的就医分组方法相关实施例中的描述,为节省篇幅,在此不再赘述,但这不应造成对本申请保护范围的损失。
- [0222] 进一步,如图4所示,该计算设备还包括:通信组件42、电源组件43等其它组件。图4中仅示意性给出部分组件,并不意味着计算设备只包括图4所示组件。
- [0223] 相应地,本申请实施例还提供一种存储有计算机程序的计算机可读存储介质,计算机程序被执行时能够实现上述方法实施例中可由计算设备执行的各步骤。
- [0224] 其中,图4中的存储器,用于存储计算机程序,并可被配置为存储其它各种数据以支持在计算平台上的操作。这些数据的示例包括用于在计算平台上操作的任何应用程序或方法的指令,联系人数据,电话簿数据,消息,图片,视频等。存储器可以由任何类型的易失

性或非易失性存储设备或者它们的组合实现,如静态随机存取存储器(SRAM),电可擦除可编程只读存储器(EEPROM),可擦除可编程只读存储器(EPROM),可编程只读存储器(PROM),只读存储器(ROM),磁存储器,快闪存储器,磁盘或光盘。

[0225] 其中,图4中的通信组件被配置为便于通信组件所在设备和其他设备之间有线或无线方式的通信。通信组件所在设备可以接入基于通信标准的无线网络,如WiFi,2G、3G、4G/LTE、5G等移动通信网络,或它们的组合。在一个示例性实施例中,通信组件经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中,所述通信组件还包括近场通信(NFC)模块,以促进短程通信。例如,在NFC模块可基于射频识别(RFID)技术,红外数据协会(IrDA)技术,超宽带(UWB)技术,蓝牙(BT)技术和其他技术来实现。

[0226] 其中,图4中的电源组件,为电源组件所在设备的各种组件提供电力。电源组件可以包括电源管理系统,一个或多个电源,及其他与为电源组件所在设备生成、管理和分配电力相关联的组件。

[0227] 本领域内的技术人员应明白,本申请的实施例可提供为方法、系统、或计算机程序产品。因此,本申请可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0228] 本申请是参照根据本申请实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0229] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品,该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

[0230] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上,使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理,从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0231] 在一个典型的配置中,计算设备包括一个或多个处理器(CPU)、输入/输出接口、网络接口和内存。

[0232] 内存可能包括计算机可读介质中的非永久性存储器,随机存取存储器(RAM)和/或非易失性内存等形式,如只读存储器(ROM)或闪存(flash RAM)。内存是计算机可读介质的示例。

[0233] 计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。

计算机的存储介质的例子包括,但不限于相变内存 (PRAM)、静态随机存取存储器 (SRAM)、动态随机存取存储器 (DRAM)、其他类型的随机存取存储器 (RAM)、只读存储器 (ROM)、电可擦除可编程只读存储器 (EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器 (CD-ROM)、数字多功能光盘 (DVD) 或其他光学存储、磁盒式磁带,磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质,可用于存储可以被计算设备访问的信息。按照本文中的界定,计算机可读介质不包括暂存电脑可读媒体 (transitory media),如调制的数据信号和载波。

[0234] 还需要说明的是,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、商品或者设备不仅包括那些要素,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、商品或者设备所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括所述要素的过程、方法、商品或者设备中还存在另外的相同要素。

[0235] 以上所述仅为本申请的实施例而已,并不用于限制本申请。对于本领域技术人员来说,本申请可以有各种更改和变化。凡在本申请的精神和原理之内所作的任何修改、等同替换、改进等,均应包含在本申请的权利要求范围之内。

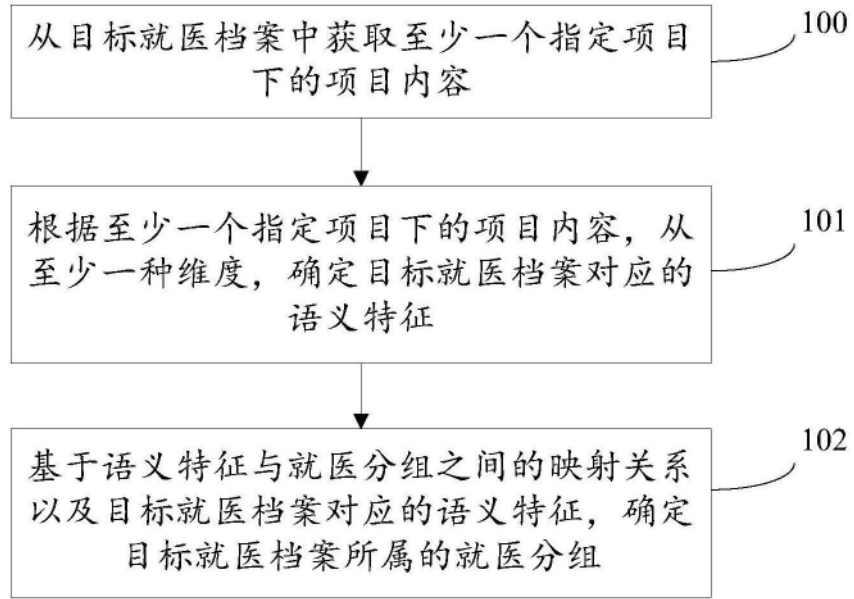


图1

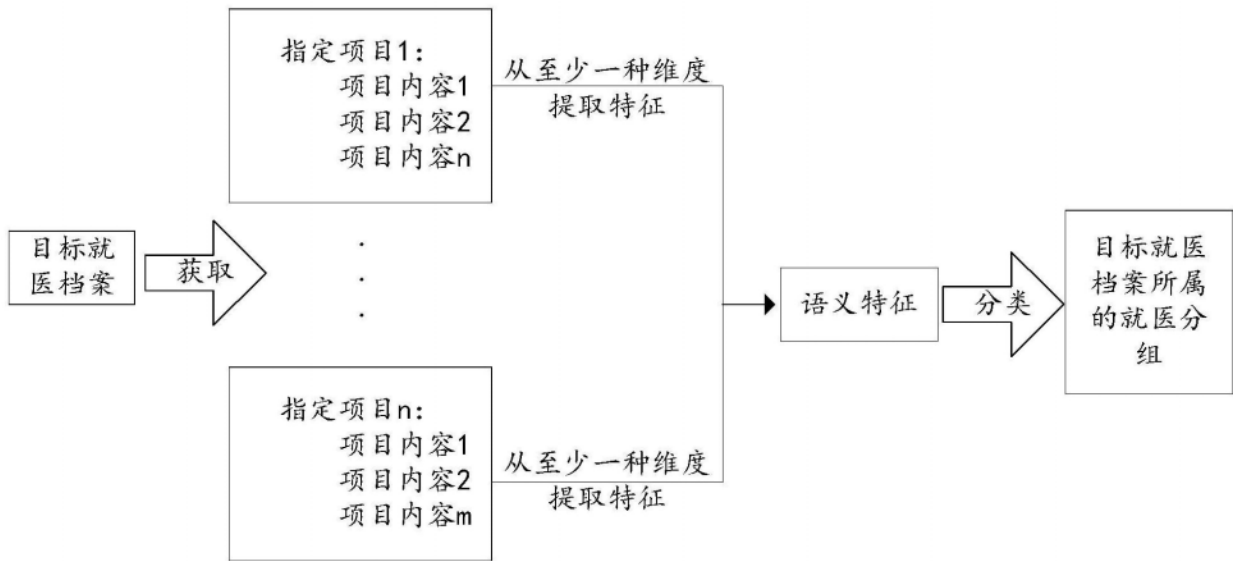


图2

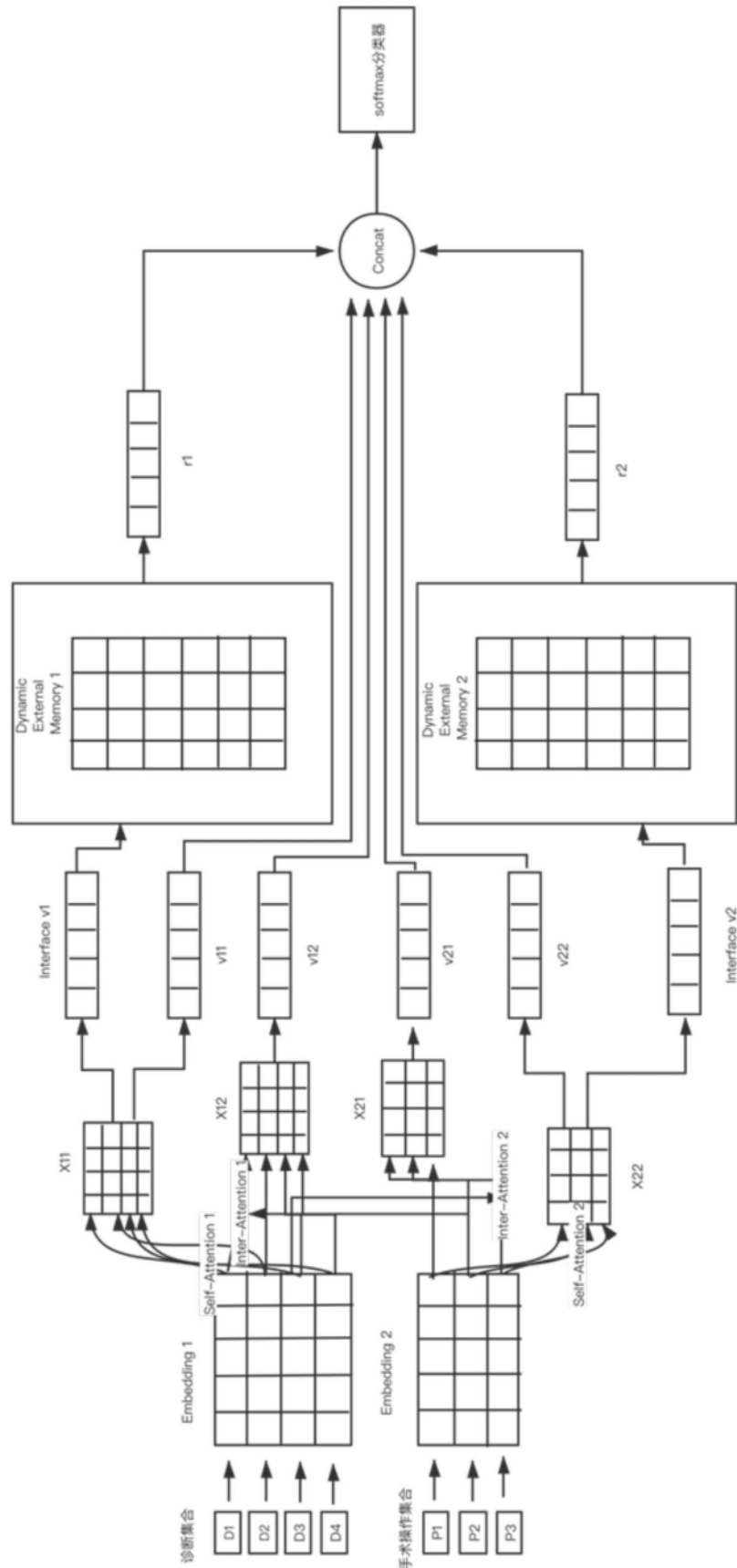


图3

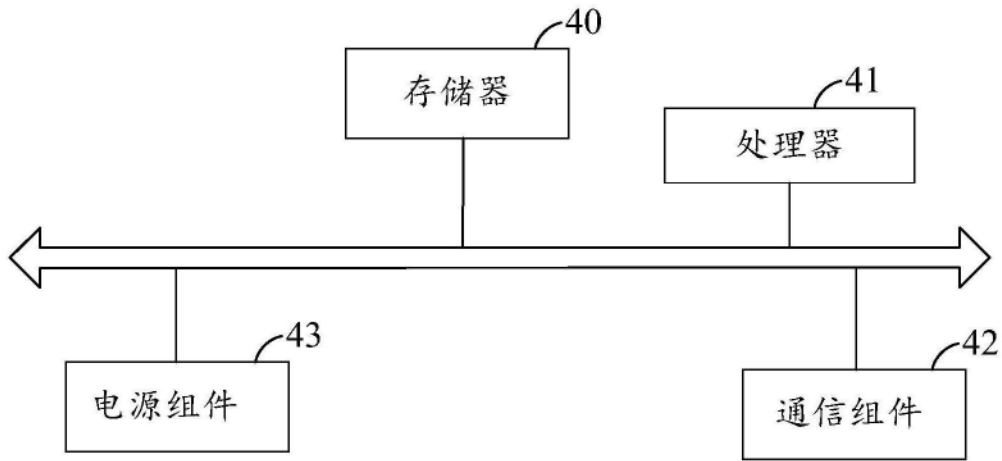


图4