



(12) 发明专利

(10) 授权公告号 CN 109614615 B

(45) 授权公告日 2022. 04. 22

(21) 申请号 201811474215.1

G06F 16/35 (2019.01)

(22) 申请日 2018.12.04

G06Q 30/02 (2012.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 109614615 A

(56) 对比文件

CN 105354199 A, 2016.02.24

CN 106897568 A, 2017.06.27

CN 108694201 A, 2018.10.23

CN 108595708 A, 2018.09.28

US 2016350366 A1, 2016.12.01

(43) 申请公布日 2019.04.12

(73) 专利权人 联想(北京)有限公司

地址 100085 北京市海淀区上地信息产业  
基地创业路6号

谭咏梅 等. 结合实体链接与实体聚类的命名实体消歧.《北京邮电大学学报》.2014, 第37卷(第5期),

(72) 发明人 杨帆 张成松

审查员 吴姝泓

(74) 专利代理机构 北京集佳知识产权代理有限公司 11227

代理人 王宝筠

(51) Int. Cl.

G06F 40/295 (2020.01)

G06F 40/30 (2020.01)

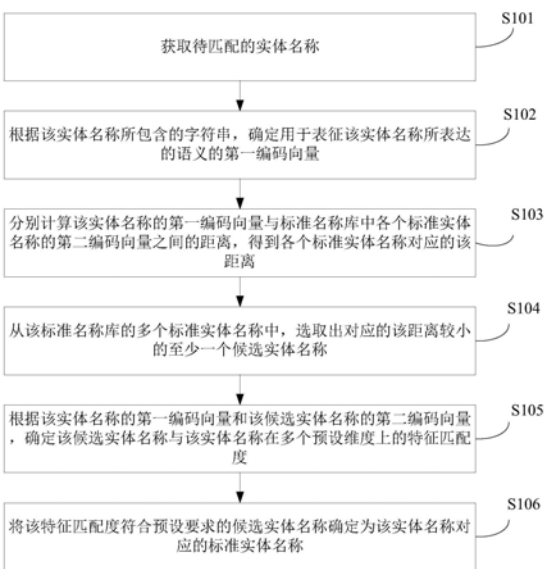
权利要求书3页 说明书14页 附图6页

(54) 发明名称

实体匹配方法、装置及电子设备

(57) 摘要

本申请公开了一种实体匹配方法、装置及电子设备,该方法包括:获取待匹配的实体名称;根据该实体名称所包含的字符串,确定用于表征该实体名称所表达的语义的第一编码向量;分别计算该实体名称的第一编码向量与标准名称库中各个标准实体名称的第二编码向量之间的距离;从该标准名称库的多个标准实体名称中,选取对应的该距离较小的至少一个候选实体名称;根据该实体名称的第一编码向量和该候选实体名称的第二编码向量,确定该候选实体名称与该实体名称在多个预设维度上的特征匹配度;将该特征匹配度符合预设要求的候选实体名称确定为该实体名称对应的标准实体名称。本申请可以提高实体匹配的精准度,并减少人力资源消耗。



1. 一种实体匹配方法,包括:

获取待匹配的实体名称;

根据所述实体名称所包含的字符串,确定用于表征所述实体名称所表达的语义的第一编码向量;

分别计算所述实体名称的第一编码向量与标准名称库中各个标准实体名称的第二编码向量之间的距离,得到各个标准实体名称对应的所述距离;所述标准实体名称的第二编码向量为用于表征所述标准实体名称所表达的语义的编码向量;

从所述标准名称库的多个标准实体名称中,选取出对应的所述距离较小的至少一个候选实体名称;

根据所述实体名称的第一编码向量和所述候选实体名称的第二编码向量,确定所述候选实体名称与所述实体名称在多个预设维度上的特征匹配度;

将所述特征匹配度符合预设要求的候选实体名称确定为所述实体名称对应的标准实体名称;

所述根据所述实体名称的第一编码向量和所述候选实体名称的第二编码向量,确定所述候选实体名称与所述实体名称在多个预设维度上的特征匹配度,包括:

将所述实体名称的第一编码向量和所述候选实体名称的第二编码向量合并为一个多维度的向量,并将合并出的向量确定为表征候选实体名称与实体名称之间特征关系的特征向量,所述特征向量的维度为所述第一编码向量和所述第二编码向量的维度之和;

根据所述特征向量,并利用预先通过机器学习算法训练出的二分类模型,确定所述候选实体名称与所述实体名称之间的特征匹配度。

2. 根据权利要求1所述的实体匹配方法,所述根据所述实体名称所包含的字符串,确定用于表征所述实体名称所表达的语义的第一编码向量,包括:

对所述实体名称分词,得到由所述实体名称分词出的至少一个词构成的第一词序列;

依次确定所述第一词序列中各个词的编码向量;

根据所述第一词序列中各个词的编码向量,构建出用于表征所述实体名称所表达的语义的第一编码向量。

3. 根据权利要求2所述的实体匹配方法,所述根据所述第一词序列中各个词的编码向量,构建出用于表征所述实体名称所表达的语义的第一编码向量,包括:

将所述第一词序列中各个词的编码向量输入到预先训练出的向量转换模型中,并获取所述向量转换模型输出的第一编码向量,所述向量转换模型为利用多份正样本和多份负样本训练得到,其中,每份正样本包括一对正确标注的实体名称样本和实体标准名称样本,每份负样本包括一对错误标注的实体样本名称和实体标准名称样本。

4. 根据权利要求1所述的实体匹配方法,所述二分类模型通过如下方式训练得到:

获取多份正样本和多份负样本,每份正样本包括一对正确标注的实体名称样本和标准实体名称样本;每份负样本包括一对错误标注的实体名称样本和标准实体名称样本;

确定所述正样本中实体名称样本的第一编码向量和标准实体名称样本的第二编码向量,并利用正样本中实体名称样本的第一编码向量和标准实体名称样本的第二编码向量构建所述正样本的特征向量;

确定所述负样本中实体名称样本的第一编码向量和标准实体名称样本的第二编码向

量,并利用负样本中实体名称样本的第一编码向量和标准实体名称样本的第二编码向量构建所述负样本的特征向量;

依次将所述多份正样本的特征向量以及所述多份负样本的特征向量输入到待训练的二分类模型中,并利用机器学习算法对所述二分类模型训练,直至所述二分类模型的预测精准度符合预设要求。

5.一种实体匹配装置,包括:

别名获取单元,用于获取待匹配的实体名称;

名称编码单元,用于根据所述实体名称所包含的字符串,确定用于表征所述实体名称所表达的语义的第一编码向量;

距离计算单元,用于分别计算所述实体名称的第一编码向量与标准名称库中各个标准实体名称的第二编码向量之间的距离,得到各个标准实体名称对应的所述距离;所述标准实体名称的第二编码向量为用于表征所述标准实体名称所表达的语义的编码向量;

候选选取单元,用于从所述标准名称库的多个标准实体名称中,选取出对应的所述距离较小的至少一个候选实体名称;

匹配度确定单元,用于根据所述实体名称的第一编码向量和所述候选实体名称的第二编码向量,确定所述候选实体名称与所述实体名称在多个预设维度上的特征匹配度;

标准匹配单元,用于将所述特征匹配度符合预设要求的候选实体名称确定为所述实体名称对应的标准实体名称;

所述匹配度确定单元,包括:

特征构建单元,用于利用所述实体名称的第一编码向量和所述候选实体名称的第二编码向量,构建表征候选实体名称与实体名称之间特征关系的特征向量;

模型匹配单元,用于根据所述特征向量,并利用预先通过机器学习算法训练出的二分类模型,确定所述候选实体名称与所述实体名称之间的特征匹配度;

所述利用所述实体名称的第一编码向量和所述候选实体名称的第二编码向量,构建表征候选实体名称与实体名称之间特征关系的特征向量,包括:

将所述实体名称的第一编码向量和所述候选实体名称的第二编码向量合并为一个多维度的向量,并将合并出的向量确定为表征候选实体名称与实体名称之间特征关系的特征向量,所述特征向量的维度为所述第一编码向量和所述第二编码向量的维度之和。

6.根据权利要求5所述的实体匹配装置,所述名称编码单元,包括:

分词子单元,用于对所述实体名称分词,得到由所述实体名称分词出的至少一个词构成的第一词序列;

词编码子单元,用于依次确定所述第一词序列中各个词的编码向量;

名称编码子单元,用于根据所述第一词序列中各个词的编码向量,构建出用于表征所述实体名称所表达的语义的第一编码向量。

7.一种电子设备,包括:

处理器和存储器;

其中,所述处理器用于,获取待匹配的实体名称;根据所述实体名称所包含的字符串,确定用于表征所述实体名称所表达的语义的第一编码向量;分别计算所述实体名称的第一编码向量与标准名称库中各个标准实体名称的第二编码向量之间的距离,得到各个标准实

体名称对应的所述距离;所述标准实体名称的第二编码向量为用于表征所述标准实体名称所表达的语义的编码向量;从所述标准名称库的多个标准实体名称中,选取对应的所述距离较小的至少一个候选实体名称;根据所述实体名称的第一编码向量和所述候选实体名称的第二编码向量,确定所述候选实体名称与所述实体名称在多个预设维度上的特征匹配度;将所述特征匹配度符合预设要求的候选实体名称确定为所述实体名称对应的标准实体名称;所述根据所述实体名称的第一编码向量和所述候选实体名称的第二编码向量,确定所述候选实体名称与所述实体名称在多个预设维度上的特征匹配度,包括:将所述实体名称的第一编码向量和所述候选实体名称的第二编码向量合并为一个多维度的向量,并将合并出的向量确定为表征候选实体名称与实体名称之间特征关系的特征向量,所述特征向量的维度为所述第一编码向量和所述第二编码向量的维度之和;根据所述特征向量,并利用预先通过机器学习算法训练出的二分类模型,确定所述候选实体名称与所述实体名称之间的特征匹配度;

存储器,用于存储所述处理器执行以上操作所需的程序。

## 实体匹配方法、装置及电子设备

### 技术领域

[0001] 本发明涉及通信技术领域,更具体地说,涉及一种实体匹配方法、装置及电子设备。

### 背景技术

[0002] 实体匹配已经被应用到多个不同领域,其是指找出属于同一个实体的两个对象。如,一种比较常见的实体匹配的应用场景为,根据待匹配的实体名称,找出待匹配的实体名称对应的实体标准名称。比如,在销售流通领域中,可能需要统计和跟踪不同公司之间的商品流通,但是由于销售流水记录的数据来源不一,使得销售流水中记录的同一个公司的公司名称并不标准,因此,为了实现对销售流水的跟踪和统计,就需要将销售流水中这些不标准的公司名称匹配到相应的公司标准名称。

[0003] 目前为了确定待匹配的实体名称所对应的实体标准名称,需要人工预先观察一些不标准的实体名称与实体标准名称所构成的样本数据的特点,并总结出一些规则逻辑,然后依据这些规则逻辑确定实体名称所对应的实体标准名称。然而,人工观察的样本数据较为有限,因此,基于人工观察所制定的规则逻辑并不能覆盖所有情况,从而导致一些不标准的实体名称无法匹配到相应的实体标准名称。而且,通过人工观察样本数据总结规则逻辑的方式也对于人工的经验要求较高,且需要耗费较多的人力资源。

### 发明内容

[0004] 本发明的目的是提供一种实体匹配方法、装置及电子设备,以增大实体匹配所适用的情况,提高实体匹配的精准度,并减少人力资源消耗。

[0005] 为实现上述目的,本发明提供了如下技术方案:

[0006] 一种实体匹配方法,包括:

[0007] 获取待匹配的实体名称;

[0008] 根据所述实体名称所包含的字符串,确定用于表征所述实体名称所表达的语义的第一编码向量;

[0009] 分别计算所述实体名称的第一编码向量与标准名称库中各个标准实体名称的第二编码向量之间的距离,得到各个标准实体名称对应的所述距离;所述标准实体名称的第二编码向量为用于表征所述标准实体名称所表达的语义的编码向量;

[0010] 从所述标准名称库的多个标准实体名称中,选取出对应的所述距离较小的至少一个候选实体名称;

[0011] 根据所述实体名称的第一编码向量和所述候选实体名称的第二编码向量,确定所述候选实体名称与所述实体名称在多个预设维度上的特征匹配度;

[0012] 将所述特征匹配度符合预设要求的候选实体名称确定为所述实体名称对应的标准实体名称。

[0013] 优选地,所述根据所述实体名称所包含的字符串,确定用于表征所述实体名称所

表达的语义的第一编码向量,包括:

[0014] 对所述实体名称分词,得到由所述实体名称分词出的至少一个词构成的第一词序列;

[0015] 依次确定所述第一词序列中各个词的编码向量;

[0016] 根据所述第一词序列中各个词的编码向量,构建出用于表征所述实体名称所表达的语义的第一编码向量。

[0017] 优选地,所述根据所述第一词序列中各个词的编码向量,构建出用于表征所述实体名称所表达的语义的第一编码向量,包括:

[0018] 将所述第一词序列中各个词的编码向量输入到预先训练出的向量转换模型中,并获取所述向量转换模型输出的第一编码向量,所述向量转换模型为利用多份正样本和多份负样本训练得到,其中,每份正样本包括一对正确标注的实体名称样本和实体标准名称样本,每份负样本包括一对错误标注的实体样本名称和实体标准名称样本。

[0019] 优选地,所述根据所述实体名称的第一编码向量和所述候选实体名称的第二编码向量,确定所述候选实体名称与所述实体名称在多个预设维度上的特征匹配度,包括:

[0020] 利用所述实体名称的第一编码向量和所述候选实体名称的第二编码向量,构建表征候选实体名称与实体名称之间特征关系的特征向量;

[0021] 根据所述特征向量,并利用预先通过机器学习算法训练出的二分类模型,确定所述候选实体名称与所述实体名称之间的特征匹配度。

[0022] 优选地,所述利用所述实体名称的第一编码向量和所述候选实体名称的第二编码向量,构建表征候选实体名称与实体名称之间特征关系的特征向量,包括:

[0023] 将所述实体名称的第一编码向量和所述候选实体名称的第二编码向量合并为一个多维度的向量,并将合并出的向量确定为表征候选实体名称与实体名称之间特征关系的特征向量,所述特征向量的维度为所述第一编码向量和所述第二编码向量的维度之和。

[0024] 优选地,所述二分类模型通过如下方式训练得到:

[0025] 获取多份正样本和多份负样本,每份正样本包括一对正确标注的实体名称样本和标准实体名称样本;每份负样本包括一对错误标注的实体名称样本和标准实体名称样本;

[0026] 确定所述正样本中实体名称样本的第一编码向量和标准实体名称样本的第二编码向量,并利用正样本中实体名称样本的第一编码向量和标准实体名称样本的第二编码向量构建所述正样本的特征向量;

[0027] 确定所述负样本中实体名称样本的第一编码向量和标准实体名称样本的第二编码向量,并利用负样本中实体名称样本的第一编码向量和标准实体名称样本的第二编码向量构建所述负样本的特征向量;

[0028] 依次将所述多份正样本的特征向量以及所述多份负样本的特征向量输入到待训练的二分类模型中,并利用机器学习算法对所述二分类模型训练,直至所述二分类模型的预测精准度符合预设要求。

[0029] 又一方面,本申请还提供了一种实体匹配装置,包括:

[0030] 别名获取单元,用于获取待匹配的实体名称;

[0031] 名称编码单元,用于根据所述实体名称所包含的字符串,确定用于表征所述实体名称所表达的语义的第一编码向量;

[0032] 距离计算单元,用于分别计算所述实体名称的第一编码向量与标准名称库中各个标准实体名称的第二编码向量之间的距离,得到各个标准实体名称对应的所述距离;所述标准实体名称的第二编码向量为用于表征所述标准实体名称所表达的语义的编码向量;

[0033] 候选选取单元,用于从所述标准名称库的多个标准实体名称中,选取对应的所述距离较小的至少一个候选实体名称;

[0034] 匹配度确定单元,用于根据所述实体名称的第一编码向量和所述候选实体名称的第二编码向量,确定所述候选实体名称与所述实体名称在多个预设维度上的特征匹配度;

[0035] 标准匹配单元,用于将所述特征匹配度符合预设要求的候选实体名称确定为所述实体名称对应的标准实体名称。

[0036] 优选地,所述名称编码单元,包括:

[0037] 分词子单元,用于对所述实体名称分词,得到由所述实体名称分词出的至少一个词构成的第一词序列;

[0038] 词编码子单元,用于依次确定所述第一词序列中各个词的编码向量;

[0039] 名称编码子单元,用于根据所述第一词序列中各个词的编码向量,构建出用于表征所述实体名称所表达的语义的第一编码向量。

[0040] 优选地,所述匹配度确定单元,包括:

[0041] 特征构建单元,用于利用所述实体名称的第一编码向量和所述候选实体名称的第二编码向量,构建表征候选实体名称与实体名称之间特征关系的特征向量;

[0042] 模型匹配单元,用于根据所述特征向量,并利用预先通过机器学习算法训练出的二分类模型,确定所述候选实体名称与所述实体名称之间的特征匹配度。

[0043] 又一方面,本申请还提供了一种电子设备,包括:

[0044] 处理器和存储器;

[0045] 其中,所述处理器用于,获取待匹配的实体名称;根据所述实体名称所包含的字符串,确定用于表征所述实体名称所表达的语义的第一编码向量;分别计算所述实体名称的第一编码向量与标准名称库中各个标准实体名称的第二编码向量之间的距离,得到各个标准实体名称对应的所述距离;所述标准实体名称的第二编码向量为用于表征所述标准实体名称所表达的语义的编码向量;从所述标准名称库的多个标准实体名称中,选取对应的所述距离较小的至少一个候选实体名称;根据所述实体名称的第一编码向量和所述候选实体名称的第二编码向量,确定所述候选实体名称与所述实体名称在多个预设维度上的特征匹配度;将所述特征匹配度符合预设要求的候选实体名称确定为所述实体名称对应的标准实体名称;

[0046] 存储器,用于存储所述处理器执行以上操作所需的程序。

[0047] 通过以上方案可知,本申请实施例中,在获取到待匹配的实体名称之后,会根据该实体名称包含的字符串,确定用于表征该实体名称所表达语义的第一编码向量,并依据该第一编码向量与标准名称库中各个标准实体名称的距离,从标准名称库中选取具有该距离较小的至少一个候选实体名称,实现对标准名称库的全面筛选,以筛选出与实体名称匹配度相对较高的至少一个候选实体名称,提高匹配的覆盖率。而且,为了提高匹配精准度,本申请还会根据实体名称的第一编码向量以及候选实体名称的第二编码向量,进一步确定候选实体名称与实体名称在多个预设维度上的特征匹配度,并将特征匹配度符合要求的候

选实体名称确定为该实体名称的标准实体名称。可见,本申请在不需要依靠人工通过经验观察得到的规则逻辑,便可以为待匹配的实体名称匹配到标准实体名称,而且,匹配过程的覆盖率较高,并有利于提高精准度。

## 附图说明

[0048] 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。

[0049] 图1为本申请实施例提供的一种实体匹配方法的一种流程示意图;

[0050] 图2为本申请实施例中确定实体名称的第一编码向量的一种流程示意图;

[0051] 图3为本申请实施例中一种训练向量转换模型的一种流程示意图;

[0052] 图4为本申请实施例中一种训练二分类模型的一种流程示意图;

[0053] 图5为本申请实施例中一种实体匹配方法又一种流程示意图;

[0054] 图6为本申请实施例中一种实体匹配装置的一种组成示意图;

[0055] 图7为本申请实施例中一种电子设备的组成架构示意图。

[0056] 说明书和权利要求书及上述附图中的术语“第一”、“第二”、“第三”“第四”等(如果存在)是用于区别类似的部分,而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换,以便这里描述的本申请的实施例能够以除了在这里图示的以外的顺序实施。

## 具体实施方式

[0057] 本申请的方案适用于对待匹配的实体名称匹配标准名称的场景,其在不需要人工观察样本数据得到匹配规则的前提下,可以提高标准实体名称的召回率和精准度。

[0058] 本申请实施例的方案适用于具备数据计算能力的电子设备,如个人计算机或者服务器等等。

[0059] 下面结合附图对本申请实施例的方案进行详细介绍。

[0060] 如,参见图1,其示出了本申请一种实体匹配方法一个实施例的流程示意图,本实施例的方法可以用于上面提到的电子设备,本实施例的方法可以包括:

[0061] S101,获取待匹配的实体名称。

[0062] 待匹配的实体名称也就是需要确定所对应的标准实体名称的实体名称。在实际应用中,为了与标准实体名称相区分,也将待匹配的实体名称称为实体别名。但是由于用户并不确定实体别名是否为标准的实体名称,因此,实体别名匹配出的标准实体名称也有可能是实体别名本身。

[0063] 如,获取用户输入的待匹配的实体名称。又如,从多个待匹配标准实体名称的实体名称中确定当前待匹配的实体名称。当然,在实际应用中,在需要匹配标准实体名称的过程中,获取待匹配的实体名称的方式可以有多种,本申请对此不加以限制。

[0064] S102,根据该实体名称所包含的字符串,确定用于表征该实体名称所表达的语义的第一编码向量。



[0065] 其中,为了便于区分,将基于实体名称所包含的字符串确定出的编码向量称为第一编码向量,而将后续标准名称库中标准实体名称的编码向量称为第二编码向量。

[0066] 可以理解的是,根据实体名称所包含的字符串,确定该实体名称对应的第一编码向量的方式可以有多种,如,在一种可能的实现方式中,可以采用已知的编码向量模型来确定该实体名称的第一编码向量。

[0067] 又如,在又一种可能的实现方式中,还可以根据正确标注的实体名称样本与标准实体名称样本对,以及错误标注的实体名称样本与标准实体名称样本,训练一个用于确定字符串编码向量的编码模型,这样,将该实体名称的字符串输入到该编码模型,便可以输出该实体名称对应的第一编码向量。

[0068] 其中,依据正确标注的实体名称样本与标准实体名称样本对,以及错误标注的实体名称样本与标准实体名称样本,训练一个用于确定字符串编码向量的编码模型的原理可以以为:对于正确标注的实体名称样本与标准实体名称样本而言,该实体名称样本的编码向量与标准实体名称样本的编码向量之间的距离较小;而对于错误标注的实体名称样本与标准实体名称样本而言,实体名称样本的编码向量与标准实体名称的编码向量之间的距离相对较大,基于此,通过不断调整待训练的编码模型,可以输出向量预测的精准度较高的编码向量。

[0069] 当然,根据实体名称的字符串,确定实体名称的第一编码向量还可以有其他实现方式,本申请对此不加以限制。

[0070] S103,分别计算该实体名称的第一编码向量与标准名称库中各个标准实体名称的第二编码向量之间的距离,得到各个标准实体名称对应的该距离。

[0071] 其中,标准实体名称的第二编码向量为用于表征标准实体名称所表达的语义的编码向量。

[0072] 可以理解的是,为了能够便捷的计算第一编码向量与第二编码向量之间的距离,该第一编码向量的维度与该第二编码向量的维度相同。

[0073] 在本申请实施例,可以在获取到待匹配的实体名称之后,分别计算该标准名称库中各个标准实体名称的第二编码向量,如,在执行该步骤S102的同时,计算标准名称库中各个标准实体名称的编码向量;或者是,在执行了步骤S102之后,计算标准名称库中各个标准实体名称的编码向量。

[0074] 可选的,考虑到在获取到待匹配的实体名称之后,计算标准名称库中各个标准实体名称的第二编码向量需要耗费较长时间,因此,在本申请实施例中可以预先计算出标准名称库中各个标准实体名称的第二编码向量,并存储标准名称库中各个标准实体名称各自对应的第二编码向量。相应的,在该步骤S103可以为:分别计算该实体名称的第一编码向量与预先存储的标准名称库中各个标准实体名称的第二编码向量之间的距离。

[0075] 其中,第一编码向量与第二编码向量之间的距离也可以称为向量距离。该向量距离可以有多种形式,如可以计算第一编码向量与第二编码向量之间的欧几里得距离或者曼哈顿距离等等。

[0076] 可以理解的是,对于每个标准实体名称而言,需要计算该实体名称的第一编码向量与该标准实体名称的第二编码向量之间的向量距离,因此,每个标准实体名称对应一个向量距离,而多个标准实体名称对应了多个向量距离。

[0077] S104,从该标准名称库的多个标准实体名称中,选取出对应的该距离较小的至少一个候选实体名称。

[0078] 可以理解的是,如果标准名称库中的标准实体名称的第二编码向量与该待匹配的实体名称的第一编码向量之间的距离较小,则说明该标准实体名称是该实体名称匹配的标准实体名称,因此,需要从标准名称库中筛选出该距离相对较小的至少一个候选实体名称。

[0079] 如,可以选取出对应的该距离较小的前指定数量个候选实体名称,以构成候选实体名称列表。

[0080] 本申请的发明人经研究发现,现有依靠人工观察得到的规则逻辑无法覆盖所有的情况,使得匹配的召回率较高。为了提高实体名称匹配的召回率,本申请的发明人想到了基于搜索引擎对标准名称库进行初步检索,基于搜索引擎对标准名称库进行初步检索具有召回率高的特点,往往都能覆盖正确的标准实体名称。

[0081] 然而,基于搜索引擎对标准实体名称库进行检索的检索效率较低,因此,在本申请实施例中,发明人提出了类似搜索引擎的思想,通过比对待匹配的实体名称的第一编码向量与标准名称库中的标准实体名称的第二编码向量之间的向量距离相似性,由于向量距离计算的复杂度相对较低,从而相对于利用搜索引擎筛选出候选实体名称而言,利用基于向量距离来筛选候选实体名称的方式更为高效。

[0082] 进一步的,考虑到单纯基于向量距离并不能精准的反映候选实体名称与待匹配的实体名称之间匹配程度,因此,为了进一步提高匹配精准度,本申请在筛选出至少一个候选实体名称之后,还会执行后续步骤S105和S106的操作,以对候选实体名称中进一步筛选,并最终确定与实体名称匹配的标准实体名称。

[0083] S105,根据该实体名称的第一编码向量和该候选实体名称的第二编码向量,确定该候选实体名称与该实体名称在多个预设维度上的特征匹配度。

[0084] 可以理解的是,第一编码向量反映的是待匹配的实体名称的字符串所具有的语义特征,而第二编码向量反映出的是候选实体名称的字符串所具有的语义特征,因此,针对每个候选实体名称,需要依据该第一编码向量和第二编码向量可以分别分析该候选实体名称和实体名称在预先设定的多个预设维度所具有的特征匹配度。

[0085] 其中,该多个预设维度可以根据需要设定,如,该多个预设维度可以是反映不同信息类别的多个维度,这样,可以从多个信息类别的角度上,结合第一编码向量和第二编码向量,分析实体名称和候选实体名称的字符串在该信息类别上所具有的相似性,该信息类别上的相似性就属于该信息类别对应维度上的特征匹配度。

[0086] 可选的,该多个预设维度的数量与该第一编码向量或者说第二编码向量的维度相同,而每个预设维度可以代表第一编码向量以及第二编码向量中的一个维度。因此,在该步骤S105,可以是根据第一编码向量和第二编码向量在各个维度上的匹配情况,综合分析第一编码向量和第二编码向量在该多个预设维度上整体的特征匹配度。

[0087] 作为一种可选方式,本申请可以预先训练用于分析两个名称的编码向量之间特征匹配度的神经网络模型,如,预先训练二分类模型。其中,该神经网络模型为利用多份正样本和多份负样本训练得到的。

[0088] 其中,每份正样本包括正确标注的一对实体名称样本和标准实体名称样本,也就是说,正样本中的标准实体名称样本就是该正样本中实体名称样本对应的标准实体名称。

负样本包括错误标注的一对实体名称样本和标准实体名称样本,也就是说,负样本中,实体名称样本对应的标准实体名称并不是该负样本中的该标准实体名称样本。

[0089] 其中,按照正样本中实体名称样本的第一编码向量与该正样本中实体名称样本的第二编码向量所对应的特征匹配度相对较高;而负样本中实体名称样本的第一编码向量与该负样本中标准实体名称样本的第二编码向量之间的特征匹配度相对较低的原则,训练神经网络模型。具体训练方式可以比较多样,后续会结合一种情况为例进行说明。

[0090] 相应的,可以依据实体名称的第一编码向量和候选实体名称的第二编码向量,并利用该预先训练得到的神经网络模型,确定实体名称和候选实体名称之间在多个预设维度上的特征匹配度。

[0091] S106,将该特征匹配度符合预设要求的候选实体名称确定为该实体名称对应的标准实体名称。

[0092] 由于特征匹配度反映的是候选实体名称与该待匹配的实体名称在多个预设维度上的匹配程度,从而基于实体名称分别与各个候选实体名称的该特征匹配度,可以更为精准度的从候选实体名称中选取与实体名称匹配的标准实体名称。

[0093] 可见,本申请实施例中,在获取到待匹配的实体名称之后,会根据该实体名称包含的字符串,确定用于表征该实体名称所表达语义的第一编码向量,并依据该第一编码向量与标准名称库中各个标准实体名称的距离,从标准名称库中选取具有该距离较小的至少一个候选实体名称,实现对标准名称库的全面筛选,以筛选出与实体名称匹配度相对较高的至少一个候选实体名称,提高匹配的覆盖率。而且,为了提高匹配精准度,本申请还会根据实体名称的第一编码向量以及候选实体名称的第二编码向量,进一步确定候选实体名称与实体名称在多个预设维度上的特征匹配度,并将特征匹配度符合要求的候选实体名称确定为该实体名称的标准实体名称。

[0094] 由以上可知,本申请在不需要依靠人工通过经验观察得到的规则逻辑,便可以为待匹配的实体名称匹配到标准实体名称,而且,匹配过程的覆盖率较高,并有利于提高精准度。

[0095] 可以理解的是,在待匹配的实体名称确定的情况下,确定该实体名称对应的编码向量的方式可以有多种,为了便于理解,下面以一种情况为例进行详细说明。如,参见图2,其示出了本申请实施例中确定实体名称的第一编码向量的一种流程示意图,该流程可以包括:

[0096] S201,对该实体名称分词,得到由该实体名称分词出的至少一个词构成的第一词序列。

[0097] 其中,该第一词序列由该实体名称分词出的各个词组成。

[0098] 可选的,可以按照实体名称的字符串,将该实体名称分词为具有先后顺序的至少一个词,而该第一词序列中各个词的顺序与该各个词在实体名称中的顺序一致。

[0099] 其中,对实体名称的分词的具体实现可以有多种方式,本申请对此不加以限制。

[0100] S202,依次确定该第一词序列中各个词的编码向量。

[0101] 其中,对于第一词序列中各个词,确定词的编码向量的方式可以有多种,如,可以采用独热(one-hot)编码方式确定词的编码向量;或者采用预先训练好的word2vec编码该词的向量,以得到词的编码向量。

[0102] S203,根据该第一词序列中各个词的编码向量,构建出用于表征该实体名称所表达的语义的第一编码向量。

[0103] 在实体名称分词出的各个词的编码向量确定后,基于各个词的编码向量可以构建出用于表征该实体名称的第一编码向量。

[0104] 如,在一种可选的实现方式中,可以预先训练向量转换模型,并将该第一词序列中各个词的编码向量输入到该预先训练出的向量转换模型中,从而得到该向量转换模型输出的第一编码向量。

[0105] 其中,该向量转换模型可以为预先训练出的神经网络模型,如循环神经网络(RNN, Recurrent Neural Network)。

[0106] 其中,该向量转换模型为利用多份正样本和多份负样本训练得到,每份正样本包括一对正确标注的实体名称样本和实体标准名称样本,每份负样本包括一对错误标注的实体样本名称和实体标准名称样本。

[0107] 可以理解的是,在利用正样本和负样本训练向量转换模型时,可以先对正样本以及负样本中每个名称样本进行分词,得到每个名称样本的词序列中各个词的编码向量,然后,分别利用正样本中一对名称样本各自的词序列的编码向量以及负样本中一对名称样本各自的词序列的编码向量对向量转换模型进行训练。

[0108] 为了便于理解,以一种训练向量转换模型的方式进行介绍,如参见图3,其示出了本申请一种训练向量转换模型的流程示意图,为了便于介绍,以向量转换模型为RNN模型为例。该流程可以包括:

[0109] S301,获取多份正样本和多份负样本。

[0110] 其中,每份正样本包括一对正确标注的实体名称样本和实体标准名称样本,每份负样本包括一对错误标注的实体样本名称和实体标准名称样本。

[0111] S302,针对每份正样本,分别对该正样本中的该实体名称样本以及标准实体名称样本的分词,得到正样本中该实体名称样本分词出的至少一个词构成的第一词序列,以及正样本中该标准实体名称样本分词出的至少一个词构成的第二词序列。

[0112] 其中,为了便于区分,将正样本以及负样本中表征实体别名的样本(即实体名称样本)分词得到的词序列称为第一词序列,而将标准实体名称的样本(即标准实体名称样本)分词得到的词序列称为第二词序列。

[0113] S303,针对每份正样本,依次确定该正样本中第一词序列中各个词的编码向量,并依次确定该正样本中第二词序列中各个词的编码向量。

[0114] 其中,对实体名称样本以及标准实体名称样本分词,以及确定分词得到的第一词序列以及第二词序列中各个词的编码向量的过程可以参见前面对实体名称的分词以及确定实体名称的第一词序列中各个词的编码向量的相关介绍,在此不再赘述。

[0115] S304,针对每份负样本,分别对该负样本中的该实体名称样本以及标准实体名称样本的分词,得到负样本中该实体名称样本分词出的至少一个词构成的第一词序列,以及负样本中该标准名称样本分词出的至少一个词构成的第二词序列。

[0116] S305,针对每份负样本,依次确定该负样本中第一词序列中各个词的编码向量,并依次确定该负样本中第二词序列中各个词的编码向量。

[0117] 其中,该步骤S302和S303也可以与步骤S304和步骤S305同时执行,或者是,先执行

该步骤S304和S305,再执行该步骤S302和S303。

[0118] S306,针对每份正样本,将该正样本的第一词序列中各个词的编码向量输入到待训练的RNN模型,并将该正样本的第二词序列中各个词的编码向量输入到该待训练的RNN模型,以通过该RNN模型分别输出该正样本中实体名称样本的第一编码向量以及该标准实体名称的第二编码向量。

[0119] 其中,该待训练的RNN模型可以采用单层或者双层的长短期记忆网络(Long Short-Term Memory,LSTM)。在该待训练的RNN模型中有可以输出实体名称样本(即实体别名样本)的第一编码向量的别名RNN层,以及用于输出标准实体名称样本的第二编码向量的标准RNN层,可以通过该待训练的RNN模型可以分别得到正样本中实体名称样本的第一编码向量以及正样本中标准实体名称样本的第二编码向量。

[0120] 其中,该第一编码向量和第二编码向量具有相同的维度。

[0121] S307,针对每份负样本,将该负样本的第一词序列中各个词的编码向量输入到待训练的RNN模型,并将该负样本的第二词序列中各个词的编码向量输入到该待训练的RNN模型,以通过该RNN模型分别输出该负样本中实体名称样本的第一编码向量以及该负样本中标准实体名称的第二编码向量。

[0122] 其中,对于负样本,得到该负样本对应的第一编码向量以及该第二编码向量的过程与前面确定正样本对应的第一编码向量和第二编码向量的过程相似,在此不再赘述。

[0123] S308,分别计算每个正样本中第一编码向量与第二编码向量之间的向量距离,并计算每个负样本中第一编码向量与第二编码向量之间的向量距离,得到每个正样本对应的向量距离以及每个负样本对应的向量距离。

[0124] 其中,计算第一编码向量与第二编码向量之间的向量距离的方式可以采用任意的向量距离计算方式,本申请对此不加限制。

[0125] S309,基于每个正样本对应的向量距离以及每个负样本对应的向量距离,判断该RNN模型的预测精准度是否达到预设要求,如果是,则训练完成;如果否,则调整该RNN模型的内部参数,并返回执行该步骤S306。

[0126] 可以理解的是,对于正样本而言,由于正样本中的实体名称样本与标准实体名称样本是正确标注的一对名称,因此,将正样本中实体名称样本的第一编码向量以及正样本中标准实体名称样本的第二编码向量投影到同一个向量空间,那么该第一编码向量和第二编码向量在空间距离上应该比较近。相应的,对于负样本而言,负样本中实体名称样本的第一编码向量与负样本中标准实体名称样本的第二编码向量之间的向量距离应该比较远。基于这个原理,可以不断训练RNN模型。

[0127] 可选的,在实际应用中,还可以基于如上原理,设置该RNN模型对应的损失函数。这样依据损失函数的输出值,则可以判断RNN模型的预测精准度满足预设要求。

[0128] 如,可以假设正样本对应的标签 $Y=1$ ;负样本对应的标签 $Y=-1$ ,其中,正样本以及负样本中的实体名称样本对应的第一编码向量均表示为 $V_a$ ,而正样本和正样本中标准实体名称样本对应的第二编码向量均表示为 $V_s$ ,则可以构建基于向量距离的损失函数 $L(V_a, V_s)$ 可以为如下公式一:

[0129]  $L(V_a, V_s) = \exp(Y * (V_a - V_s)^2)$ ; (公式一)

[0130] 在公式一的基础上,可以得到各个样本对应的损失函数的输出值,这样,结合所有

正样本和负样本对应的损失函数的输出值,可以分析待训练的RNN模型的精准度是否符合要求。

[0131] 可以理解的是,图3仅仅是本申请中向量转换模型的一种实现方式,在实际应用中,还可以有其他方式来训练该向量转换模型,本申请对此不加限制。

[0132] 可以理解的是,在本申请实施例中,得到待匹配的实体名称的第一编码向量以及候选实体名称的第二编码向量之后,确定候选实体名称与该待匹配的实体名称在多个维度上的特征匹配度的方式可以有多种可能。

[0133] 可选的,为了能够更加便捷高效的,确定该特征匹配度,在实际应用中,还可以训练二分类模型,如通过机器学习算法训练二分类模型。相应的,基于实体名称的第一编码向量和候选实体名称的第二编码向量,并利用该二分类模型可以得到候选实体名称与该实体名称之间在多个预设维度上的特征匹配度。

[0134] 在一种实现方式中,可以先利用该实体名称的第一编码向量和所述候选实体名称的第二编码向量,构建表征候选实体名称与实体名称之间特征关系的特征向量。然后,将构建的该特征向量输入到预先训练得到的该二分类模型,以通过该二分类模型预测出该候选实体名称与该实体名称之间在多个预设维度上的特征匹配度。

[0135] 其中,二分类模型的训练方式可以有多种,为了便于理解,下面以一种训练二分类模型的方式为例进行介绍。如,参见图4,其示出了本申请一种训练二分类模型的训练流程示意图,该流程可以包括:

[0136] S401,获取多份正样本和多份负样本。

[0137] 其中,每份正样本包括一对正确标注的实体名称样本和标准实体名称样本;每份负样本包括一对错误标注的实体名称样本和标准实体名称样本;

[0138] S402,确定该正样本中实体名称样本的第一编码向量和标准实体名称样本的第二编码向量,并利用正样本中实体名称样本的第一编码向量和标准实体名称样本的第二编码向量构建该正样本的特征向量。

[0139] S403,确定该负样本中实体名称样本的第一编码向量和标准实体名称样本的第二编码向量,并利用负样本中实体名称样本的第一编码向量和标准实体名称样本的第二编码向量构建该负样本的特征向量。

[0140] 其中,正样本以及负样本中实体名称样本的第一编码向量以及标准实体名称样本的第二编码向量可以分别按照前面提到的确定第一编码向量以及第二编码向量的方式确定。如,可以利用预先训练出的RNN模型,分别确定出正样本以及负样本中实体名称样本对应的第一编码向量以及标准实体名称对应的第二编码向量。

[0141] 对于正样本或者负样本而言,基于实体名称样本的第一编码向量和标准实体名称的第二编码向量,构建特征向量的过程可以有多种可能。如,可以将第一编码向量与第二特征编码向量进行拼接,并将拼接得到的向量确定为该特征向量。又如,还可以是将该第一特征向量与该第二特征向量相减,得到该特征向量。

[0142] S404,依次将该多份正样本的特征向量以及该多份负样本的特征向量输入到待训练的二分类模型中,并利用机器学习算法对该二分类模型训练,直至该二分类模型的预测精准度符合预设要求。

[0143] 可以理解的是,对于正样本而言,基于正样本对应的特征向量,利用二分类模型预

测出的特征匹配度应相对较高,如,特征匹配度应该大于表征该正样本中实体名称样本与标准实体名称样本为相互匹配的匹配度阈值;相应的,对于负样本而言,基于负样本对应的特征向量,利用该二分类模型预测出的特征匹配度应相对较低,如预测出的特征匹配度低于该匹配度阈值。基于此原理,可以判断该二分类模型对于一份样本的预测是否准确,从而统计对所有样本的预测情况,得到预测精准度。

[0144] 可以理解的是,如果判断出该二分类模型的预测精准度不符合要求,则可以调整该二分类模型的内部参数,并重新返回步骤S402进行训练,直至该二分类模型的预测精准度符合预设要求。

[0145] 其中,该机器学习算法可以有多种,如,可以为支持向量机、随机森林、逻辑回归、朴素贝叶斯等等算法。

[0146] 为了便于理解本申请的方案,下面以本申请的一种具体实现方式为例,对本申请方案进行介绍,如,参见图5,其示出了本申请一种实体匹配方法又一个实施例的流程示意图,本实施例的方法可以包括:

[0147] S501,获取待匹配的实体名称。

[0148] S502,对该实体名称分词,得到由该实体名称分词出的至少一个词构成的第一词序列。

[0149] S503,依次确定该第一词序列中各个词的编码向量。

[0150] S504,将该第一词序列中各个词的编码向量输入到该预先训练出的向量转换模型中,得到该向量转换模型输出的该实体名称的第一编码向量。

[0151] 该向量转换模型可以为前面提到的训练出的RNN模型。

[0152] 可以理解的是,得到该实体名称的第一编码向量的过程可以参见前面实施例的相关介绍,在此不再赘述。

[0153] S505,分别计算该实体名称的第一编码向量与标准名称库中各个标准实体名称的第二编码向量之间的距离,得到各个标准实体名称对应的该距离。

[0154] 其中,标准名称库中各个标准实体名称的第二编码向量可以与确定实体名称的第一编码向量的过程相似,具体可以参见前面的相关介绍。其中,该步骤S505可以是计算该实体名称的第一编码向量与预先存储的各个标准实体名称的第二编码向量之间的距离。

[0155] 其中,该第一编码向量的维度与该第二编码向量的维度相同。

[0156] S506,从该标准名称库的多个标准实体名称中,选取对应的该距离较小的至少一个候选实体名称。

[0157] S507,针对每个候选实体名称,利用该实体名称的第一编码向量和该候选实体名称的第二编码向量,构建表征该候选实体名称与实体名称之间特征关系的特征向量。

[0158] 如,将第一编码向量与第二编码向量拼接为该特征向量。具体的,针对每个候选实体名称,将该实体名称的第一编码向量和该候选实体名称的第二编码向量合并为一个多维度的向量,并将合并出的向量确定为表征候选实体名称与实体名称之间特征关系的特征向量。可见,合并得到的该特征向量的维度为第一编码向量和所述第二编码向量的维度之和。

[0159] S508,依次将各个候选实体名称对应的特征向量输入到预先通过机器学习算法训练得到的二分类模型中,获得该二分类模型针对各个候选实体名称对应的特征向量输出的特征匹配度。

[0160] 其中,该二分类模型可以根据特征向量所包含的各个维度,确定该特征向量所对应的特征匹配度,其中,该特征匹配度可以从多个预设维度上反映出候选实体名称与待匹配的实体名称之间的匹配程度。该多个预设维度可以认为是该特征向量对应的多个维度。

[0161] S509,将该至少一个候选实体名称中对应的特征匹配度最高的候选实体名称确定为该实体名称对应的标准实体名称。

[0162] 在本实施例中,是以将特征匹配度最高的候选实体名称作为与该实体名称匹配的标准实体名称为例说明,对于选取特征匹配度排名处于前指定位中的任意一个候选实体名称作为该标准实体名称,或者基于该特征匹配度,采取其他方式选取标准实体名称的方式也同样适用于本实施例。

[0163] 对应本申请的一种实体匹配方法,本申请还提供了一种实体匹配装置。如参见图6,其示出了本申请一种实体匹配装置一个实施例的组成示意图,该装置可以应用于电子设备,该装置可以包括:

[0164] 别名获取单元601,用于获取待匹配的实体名称;

[0165] 名称编码单元602,用于根据所述实体名称所包含的字符串,确定用于表征所述实体名称所表达的语义的第一编码向量;

[0166] 距离计算单元603,用于分别计算所述实体名称的第一编码向量与标准名称库中各个标准实体名称的第二编码向量之间的距离,得到各个标准实体名称对应的所述距离;所述标准实体名称的第二编码向量为用于表征所述标准实体名称所表达的语义的编码向量;

[0167] 候选选取单元604,用于从所述标准名称库的多个标准实体名称中,选取出对应的所述距离较小的至少一个候选实体名称;

[0168] 匹配度确定单元605,用于根据所述实体名称的第一编码向量和所述候选实体名称的第二编码向量,确定所述候选实体名称与所述实体名称在多个预设维度上的特征匹配度;

[0169] 标准匹配单元606,用于将所述特征匹配度符合预设要求的候选实体名称确定为所述实体名称对应的标准实体名称。

[0170] 在一种可能的实现方式中,所述名称编码单元,包括:

[0171] 分词子单元,用于对所述实体名称分词,得到由所述实体名称分词出的至少一个词构成的第一词序列;

[0172] 词编码子单元,用于依次确定所述第一词序列中各个词的编码向量;

[0173] 名称编码子单元,用于根据所述第一词序列中各个词的编码向量,构建出用于表征所述实体名称所表达的语义的第一编码向量。

[0174] 可选的,所述名称编码子单元,包括:

[0175] 模型转换子单元,用于将所述第一词序列中各个词的编码向量输入到预先训练出的向量转换模型中,并获取所述向量转换模型输出的第一编码向量,所述向量转换模型为利用多份正样本和多份负样本训练得到,其中,每份正样本包括一对正确标注的实体名称样本和实体标准名称样本,每份负样本包括一对错误标注的实体样本名称和实体标准名称样本。

[0176] 在一种可能的实现方式中,以上装置的任意一个实施例中,所述匹配度确定单元



可以包括：

[0177] 特征构建单元，用于利用所述实体名称的第一编码向量和所述候选实体名称的第二编码向量，构建表征候选实体名称与实体名称之间特征关系的特征向量；

[0178] 模型匹配单元，用于根据所述特征向量，并利用预先通过机器学习算法训练出的二分类模型，确定所述候选实体名称与所述实体名称之间的特征匹配度。

[0179] 可选的，所述特征构建单元，包括：

[0180] 向量拼接子单元，用于将所述实体名称的第一编码向量和所述候选实体名称的第二编码向量合并为一个多维度的向量，并将合并出的向量确定为表征候选实体名称与实体名称之间特征关系的特征向量，所述特征向量的维度为所述第一编码向量和所述第二编码向量的维度之和。

[0181] 可选的，该装置还包括二分类模型训练单元，用于通过如下方式训练所述二分类模型：

[0182] 获取多份正样本和多份负样本，每份正样本包括一对正确标注的实体名称样本和标准实体名称样本；每份负样本包括一对错误标注的实体名称样本和标准实体名称样本；

[0183] 确定所述正样本中实体名称样本的第一编码向量和标准实体名称样本的第二编码向量，并利用正样本中实体名称样本的第一编码向量和标准实体名称样本的第二编码向量构建所述正样本的特征向量；

[0184] 确定所述负样本中实体名称样本的第一编码向量和标准实体名称样本的第二编码向量，并利用负样本中实体名称样本的第一编码向量和标准实体名称样本的第二编码向量构建所述负样本的特征向量；

[0185] 依次将所述多份正样本的特征向量以及所述多份负样本的特征向量输入到待训练的二分类模型中，并利用机器学习算法对所述二分类模型训练，直至所述二分类模型的预测精准度符合预设要求。

[0186] 又一方面，本申请还提供了一种电子设备，该电子设备可以为个人计算机、服务器等具有数据处理能力的设备。如，参见图7，其示出了本申请一种电子设备的组成结构示意图，本实施例的电子设备可以包括：

[0187] 处理器701和存储器702；

[0188] 其中，所述处理器701用于，获取待匹配的实体名称；根据所述实体名称所包含的字符串，确定用于表征所述实体名称所表达的语义的第一编码向量；分别计算所述实体名称的第一编码向量与标准名称库中各个标准实体名称的第二编码向量之间的距离，得到各个标准实体名称对应的所述距离；所述标准实体名称的第二编码向量为用于表征所述标准实体名称所表达的语义的编码向量；从所述标准名称库的多个标准实体名称中，选取出对应的所述距离较小的至少一个候选实体名称；根据所述实体名称的第一编码向量和所述候选实体名称的第二编码向量，确定所述候选实体名称与所述实体名称在多个预设维度上的特征匹配度；将所述特征匹配度符合预设要求的候选实体名称确定为所述实体名称对应的标准实体名称。

[0189] 存储器702，用于存储所述处理器执行以上操作所需的程序。

[0190] 可以理解的是，图7仅仅是本申请的电子设备的简单构成的示意图，在实际应用中，该电子设备还可以包括显示器，用于显示确定出的标准实体名称；通信模块，用于接收

待匹配的实体名称或者发送匹配出的标准实体名称等等,当然,该电子设备还可以包括输入装置等等,在此不再赘述。

[0191] 可以理解的是,该处理器所执行的具体操作可以参见前面图1到图5实施例中电子设备侧所执行的操作,在此不再重复赘述。

[0192] 对所公开的实施例的上述说明,使本领域专业技术人员能够实现或使用本发明。对这些实施例的多种修改对本领域的专业技术人员来说将是显而易见的,本文中所定义的一般原理可以在不脱离本发明的精神或范围的情况下,在其它实施例中实现。因此,本发明将不会被限制于本文所示的这些实施例,而是要符合与本文所公开的原理和新颖特点相一致的最宽的范围。

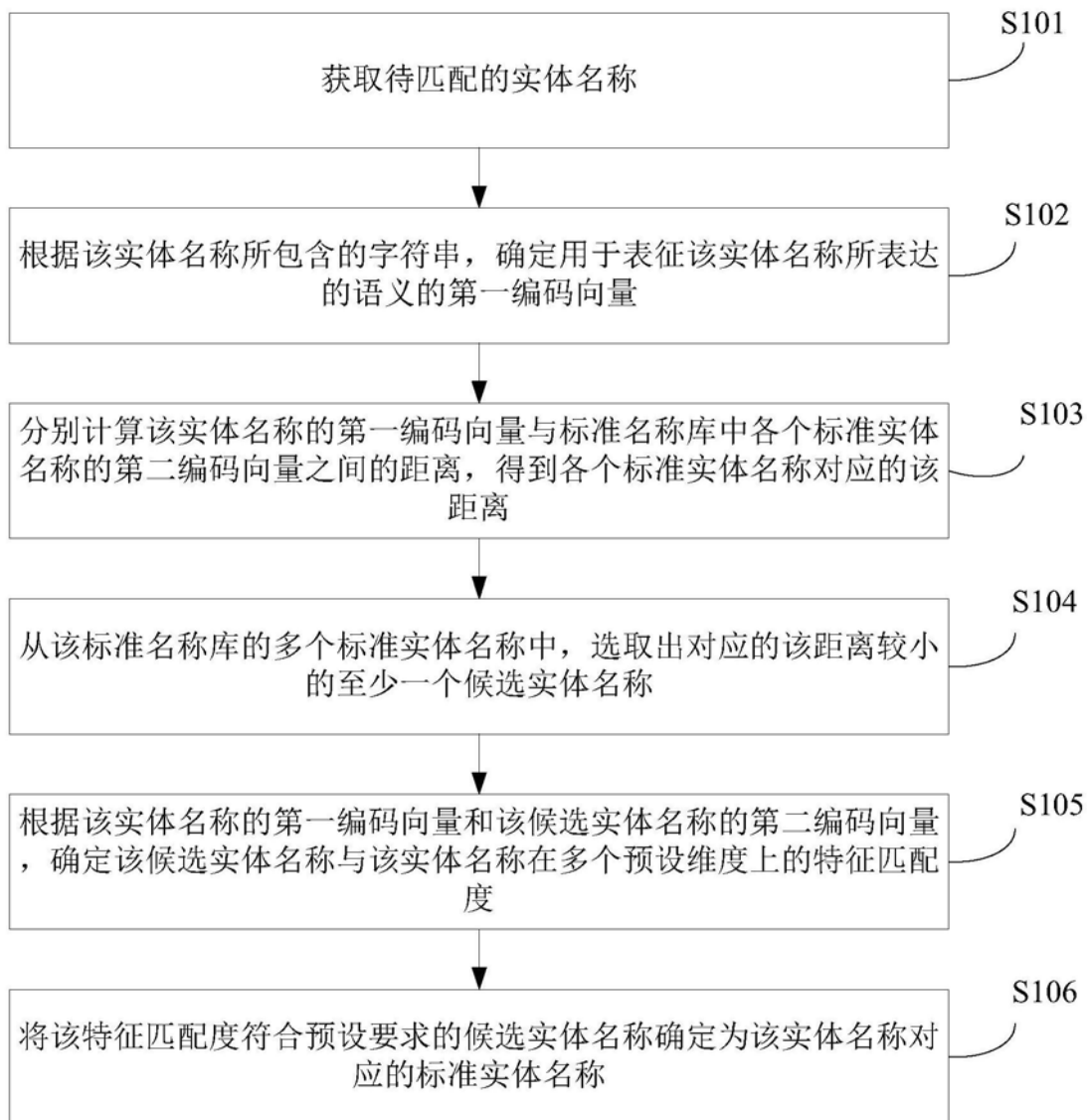


图1

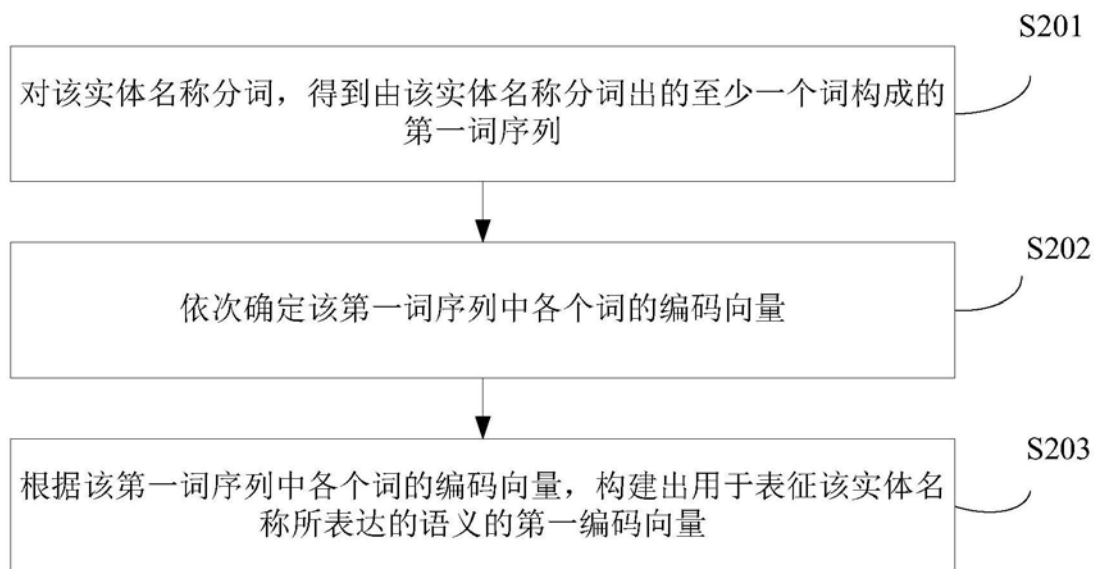


图2

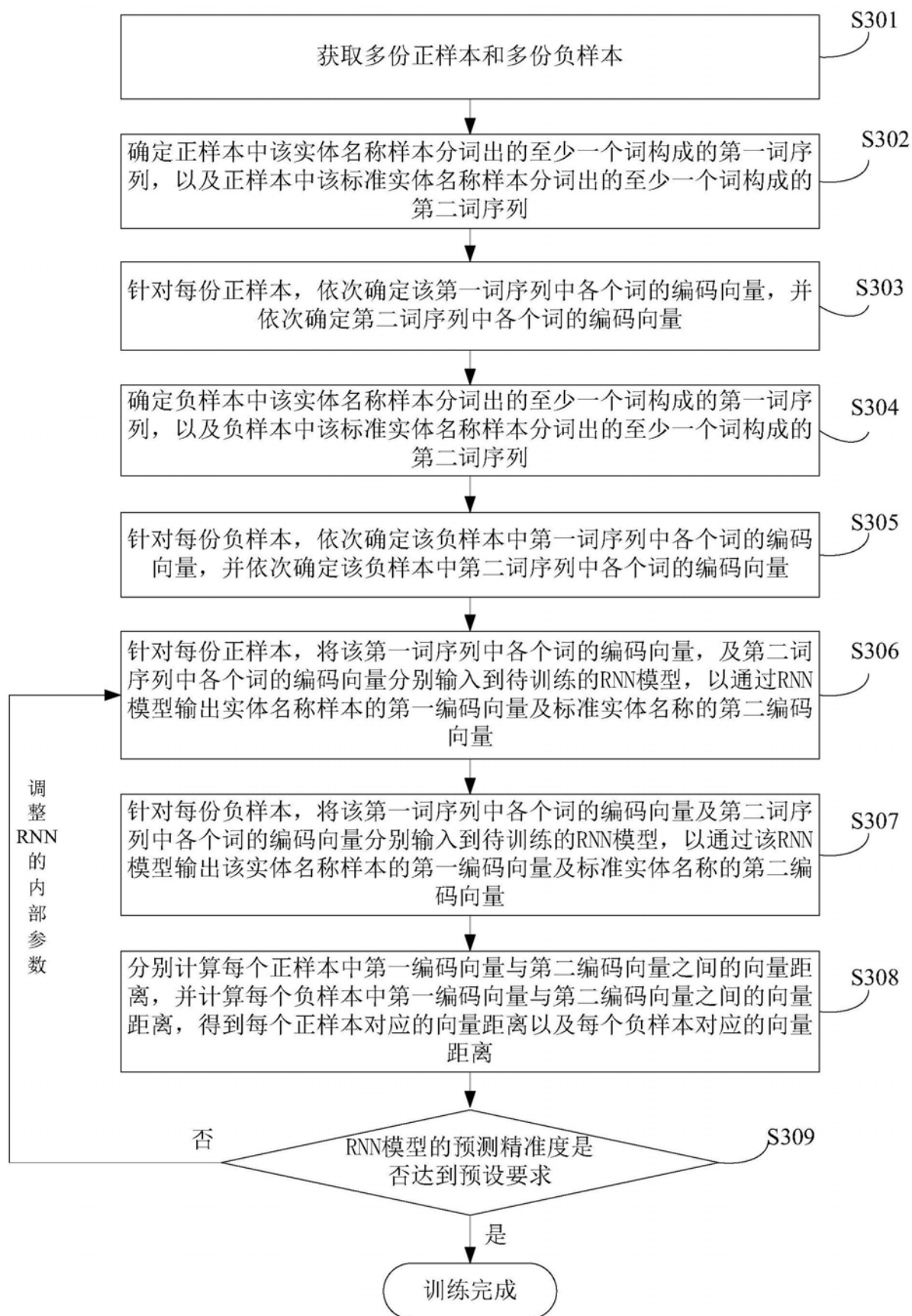


图3

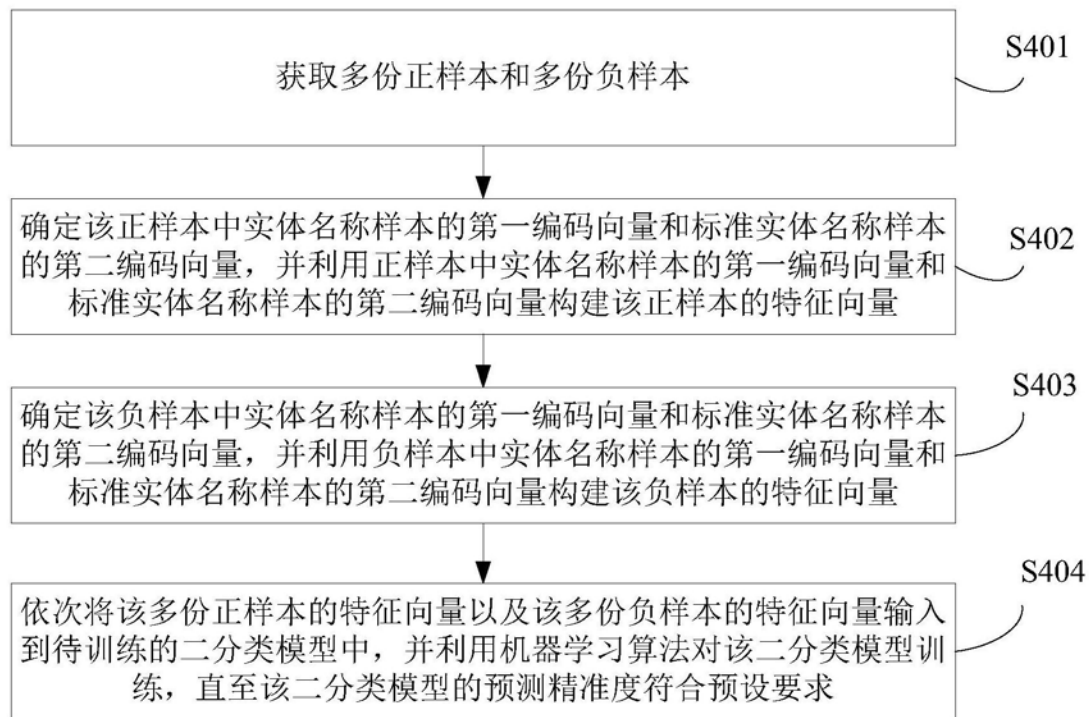


图4

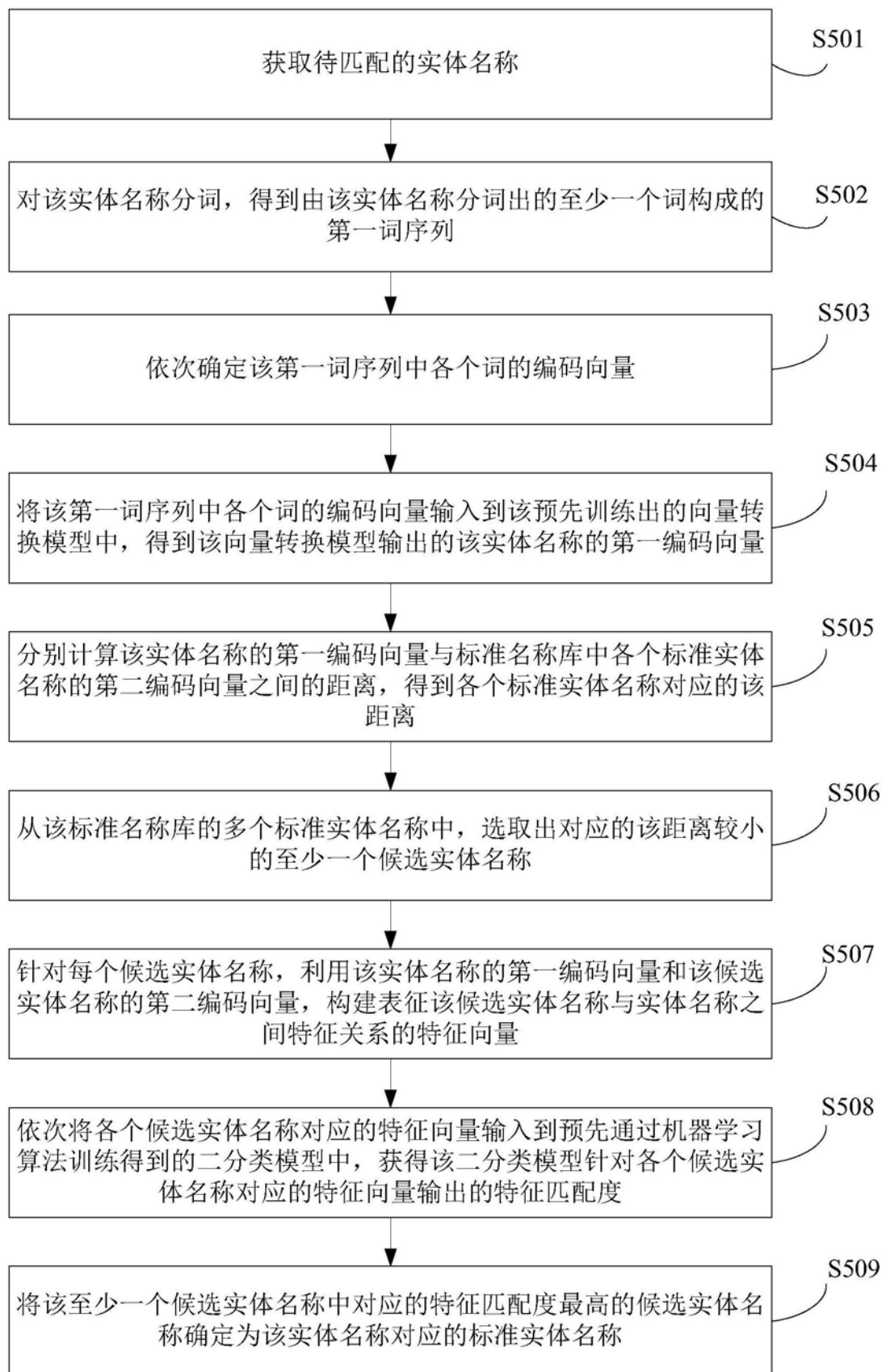


图5

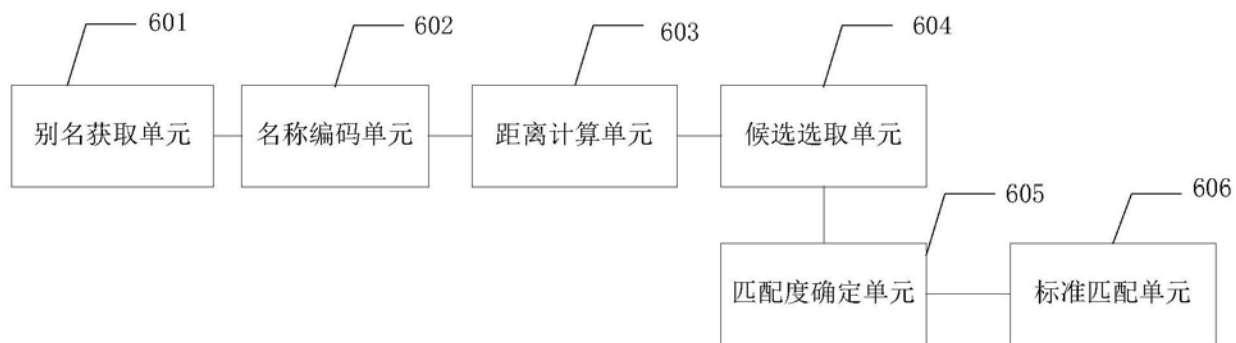


图6

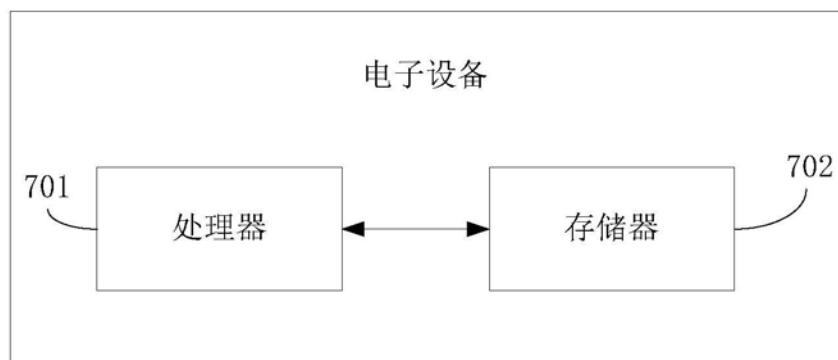


图7