



(12) 发明专利申请

(10) 申请公布号 CN 115270746 A

(43) 申请公布日 2022. 11. 01

(21) 申请号 202210687384.3

G06N 3/08 (2006.01)

(22) 申请日 2022.06.17

G06N 5/02 (2006.01)

(71) 申请人 中国平安人寿保险股份有限公司
地址 518000 广东省深圳市福田区福田街
道益田路5033号平安金融中心14、15、
16、37、41、44、45、46、54、58、59层

(72) 发明人 刘丹

(74) 专利代理机构 广州嘉权专利商标事务所有
限公司 44205
专利代理师 廖慧贤

(51) Int. Cl.

G06F 40/186 (2020.01)

G06F 16/36 (2019.01)

G06F 40/30 (2020.01)

G06K 9/62 (2022.01)

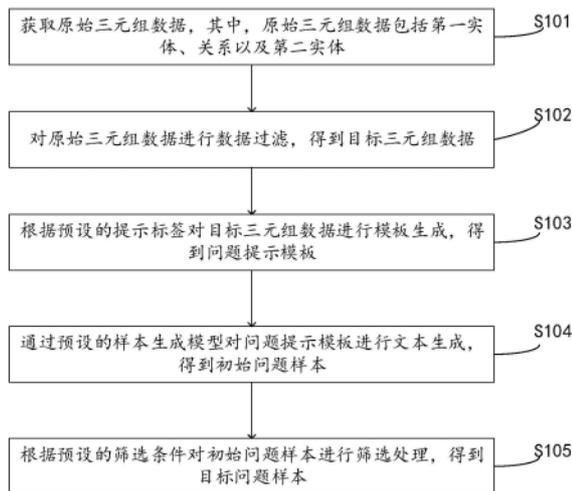
权利要求书2页 说明书15页 附图4页

(54) 发明名称

问题样本生成方法和装置、电子设备及存储介质

(57) 摘要

本申请实施例提供了一种问题样本生成方法和装置、电子设备及存储介质,属于人工智能技术领域。该方法包括:获取原始三元组数据,其中,原始三元组数据包括第一实体、关系以及第二实体;对原始三元组数据进行数据过滤,得到目标三元组数据;根据预设的提示标签对目标三元组数据进行模板生成,得到问题提示模板;通过预设的样本生成模型对问题提示模板进行文本生成,得到初始问题样本;根据预设的筛选条件对初始问题样本进行筛选处理,得到目标问题样本。本申请实施例能够提高问答样本的多样性。



1. 一种问题样本生成方法,其特征在于,所述方法包括:
获取原始三元组数据,其中,所述原始三元组数据包括第一实体、关系以及第二实体;
对所述原始三元组数据进行数据过滤,得到目标三元组数据;
根据预设的提示标签对所述目标三元组数据进行模板生成,得到问题提示模板;
通过预设的样本生成模型对所述问题提示模板进行文本生成,得到初始问题样本;
根据预设的筛选条件对所述初始问题样本进行筛选处理,得到目标问题样本。
2. 根据权利要求1所述的问题样本生成方法,其特征在于,所述获取原始三元组数据的步骤,包括:
获取预设的知识图谱库;
通过预设的知识抽取模型对所述知识图谱库进行知识抽取,得到所述原始三元组数据。
3. 根据权利要求2所述的问题样本生成方法,其特征在于,所述知识抽取模型包括实体抽取层和关系识别层,所述通过预设的知识抽取模型对所述知识图谱库进行知识抽取,得到所述原始三元组数据的步骤,包括:
通过所述实体抽取层的词性类别标签对所述知识图谱库的数据进行标签分类,得到带有词性标签的三元组特征;
通过所述关系识别层对所述三元组特征进行组合处理,得到所述原始三元组数据。
4. 根据权利要求1所述的问题样本生成方法,其特征在于,所述根据预设的提示标签对所述目标三元组数据进行模板生成,得到问题提示模板的步骤,包括:
根据所述提示标签对所述目标三元组数据进行特征定位,得到位置参数;
根据所述位置参数和预设的提示模板对所述三元组数据进行模板生成,得到所述问题提示模板。
5. 根据权利要求1至4任一项所述的问题样本生成方法,其特征在于,所述样本生成模型包括掩码层、编码层以及解码层,所述通过预设的样本生成模型对所述问题提示模板进行文本生成,得到初始问题样本的步骤,包括:
通过所述掩码层对所述问题提示模板进行掩码处理,得到掩码问题矩阵;
通过所述编码层对所述掩码问题矩阵进行编码处理,得到问题样本隐向量;
通过所述解码层对所述问题样本隐向量进行解码处理,得到所述初始问题样本。
6. 根据权利要求5所述的问题样本生成方法,其特征在于,所述通过所述编码层对所述掩码问题矩阵进行编码处理,得到问题样本隐向量的步骤,包括:
通过所述编码层对所述掩码问题矩阵进行编码处理,得到问题样本特征;
通过预设的第一注意力矩阵对所述问题样本特征进行注意力计算,得到所述问题样本隐向量。
7. 根据权利要求5所述的问题样本生成方法,其特征在于,所述通过所述解码层对所述问题样本隐向量进行解码处理,得到所述初始问题样本的步骤,包括:
通过所述解码层对所述问题样本隐向量进行解码处理,得到问题样本表征向量;
通过预设的第二注意力矩阵对所述问题样本表征向量进行注意力计算,得到所述初始问题样本。
8. 一种问题样本生成装置,其特征在于,所述装置包括:

数据获取模块,用于获取原始三元组数据,其中,所述原始三元组数据包括第一实体、关系以及第二实体;

数据过滤模块,用于对所述原始三元组数据进行数据过滤,得到目标三元组数据;

模板生成模块,用于根据预设的提示标签对所述目标三元组数据进行模板生成,得到问题提示模板;

样本生成模块,用于通过预设的样本生成模型对所述问题提示模板进行文本生成,得到初始问题样本;

筛选模块,用于根据预设的筛选条件对所述初始问题样本进行筛选处理,得到目标问题样本。

9. 一种电子设备,其特征在于,所述电子设备包括存储器、处理器、存储在所述存储器上并可在所述处理器上运行的程序以及用于实现所述处理器和所述存储器之间的连接通信的数据总线,所述程序被所述处理器执行时实现如权利要求1至7任一项所述的问题样本生成方法的步骤。

10. 一种存储介质,所述存储介质为计算机可读存储介质,用于计算机可读存储,其特征在于,所述存储介质存储有一个或者多个程序,所述一个或者多个程序可被一个或者多个处理器执行,以实现权利要求1至7中任一项所述的问题样本生成方法的步骤。

问题样本生成方法和装置、电子设备及存储介质

技术领域

[0001] 本申请涉及人工智能技术领域,尤其涉及一种问题样本生成方法和装置、电子设备及存储介质。

背景技术

[0002] 目前,高质量的问答模型的训练大多数需要依赖于高质量的问题样本,而相关技术中常常是通过人工收集和人工编写的方式来获取问题样本,这一方式往往会由于收集时间和收集范围有限而影响问题样本的多样性,因此,如何提高问题样本的多样性,成为了亟待解决的技术问题。

发明内容

[0003] 本申请实施例的主要目的在于提出一种问题样本生成方法和装置、电子设备及存储介质,旨在提高问题样本的多样性。

[0004] 为实现上述目的,本申请实施例的第一方面提出了一种问题样本生成方法,所述方法包括:

[0005] 获取原始三元组数据,其中,所述原始三元组数据包括第一实体、关系以及第二实体;

[0006] 对所述原始三元组数据进行数据过滤,得到目标三元组数据;

[0007] 根据预设的提示标签对所述目标三元组数据进行模板生成,得到问题提示模板;

[0008] 通过预设的样本生成模型对所述问题提示模板进行文本生成,得到初始问题样本;

[0009] 根据预设的筛选条件对所述初始问题样本进行筛选处理,得到目标问题样本。

[0010] 在一些实施例,所述获取原始三元组数据的步骤,包括:

[0011] 获取预设的知识图谱库;

[0012] 通过预设的知识抽取模型对所述知识图谱库进行知识抽取,得到所述原始三元组数据。

[0013] 在一些实施例,所述知识抽取模型包括实体抽取层和关系识别层,所述通过预设的知识抽取模型对所述知识图谱库进行知识抽取,得到所述原始三元组数据的步骤,包括:

[0014] 通过所述实体抽取层的词性类别标签对所述知识图谱库的数据进行标签分类,得到带有词性标签的三元组特征;

[0015] 通过所述关系识别层对所述三元组特征进行组合处理,得到所述原始三元组数据。

[0016] 在一些实施例,所述根据预设的提示标签对所述目标三元组数据进行模板生成,得到问题提示模板的步骤,包括:

[0017] 根据所述提示标签对所述目标三元组数据进行特征定位,得到位置参数;

[0018] 根据所述位置参数和预设的提示模板对所述三元组数据进行模板生成,得到所述问题提示模板。

[0019] 在一些实施例,所述样本生成模型包括掩码层、编码层以及解码层,所述通过预设的样本生成模型对所述问题提示模板进行文本生成,得到初始问题样本的步骤,包括:

[0020] 通过所述掩码层对所述问题提示模板进行掩码处理,得到掩码问题矩阵;

[0021] 通过所述编码层对所述掩码问题矩阵进行编码处理,得到问题样本隐向量;

[0022] 通过所述解码层对所述问题样本隐向量进行解码处理,得到所述初始问题样本。

[0023] 在一些实施例,所述通过所述编码层对所述掩码问题矩阵进行编码处理,得到问题样本隐向量的步骤,包括:

[0024] 通过所述编码层对所述掩码问题矩阵进行编码处理,得到问题样本特征;

[0025] 通过预设的第一注意力矩阵对所述问题样本特征进行注意力计算,得到所述问题样本隐向量。

[0026] 在一些实施例,所述通过所述解码层对所述问题样本隐向量进行解码处理,得到所述初始问题样本的步骤,包括:

[0027] 通过所述解码层对所述问题样本隐向量进行解码处理,得到问题样本表征向量;

[0028] 通过预设的第二注意力矩阵对所述问题样本表征向量进行注意力计算,得到所述初始问题样本。

[0029] 为实现上述目的,本申请实施例的第二方面提出了一种问题样本生成装置,所述装置包括:

[0030] 数据获取模块,用于获取原始三元组数据,其中,所述原始三元组数据包括第一实体、关系以及第二实体;

[0031] 数据过滤模块,用于对所述原始三元组数据进行数据过滤,得到目标三元组数据;

[0032] 模板生成模块,用于根据预设的提示标签对所述目标三元组数据进行模板生成,得到问题提示模板;

[0033] 样本生成模块,用于通过预设的样本生成模型对所述问题提示模板进行文本生成,得到初始问题样本;

[0034] 筛选模块,用于根据预设的筛选条件对所述初始问题样本进行筛选处理,得到目标问题样本。

[0035] 为实现上述目的,本申请实施例的第三方面提出了一种电子设备,所述电子设备包括存储器、处理器、存储在所述存储器上并可在所述处理器上运行的程序以及用于实现所述处理器和所述存储器之间的连接通信的数据总线,所述程序被所述处理器执行时实现上述第一方面所述的方法。

[0036] 为实现上述目的,本申请实施例的第四方面提出了一种存储介质,所述存储介质为计算机可读存储介质,用于计算机可读存储,所述存储介质存储有一个或者多个程序,所述一个或者多个程序可被一个或者多个处理器执行,以实现上述第一方面所述的方法。

[0037] 本申请提出的问题样本生成方法、问题样本生成装置、电子设备及存储介质,其通过获取原始三元组数据,其中,原始三元组数据包括第一实体、关系以及第二实体,对原始三元组数据进行数据过滤,得到目标三元组数据,能够有效地剔除相关性较低的三元组

数据以及异常数据,提高数据准确性。进一步地,根据预设的提示标签对目标三元组数据进行模板生成,得到问题提示模板,能够较为方便地捕捉到实体和关系之间的知识,从而得到能够用于构建问题样本的问题提示模板,通过预设的样本生成模型对问题提示模板进行文本生成,得到初始问题样本,能够较为方便地对问题提示模板的语义信息进行分析,得到多个符合需求的初始问题样本。最后,根据预设的筛选条件对初始问题样本进行筛选处理,得到目标问题样本,能够提高目标问题样本的质量,也能够提高目标问题样本的生成效率和目标问题样本的内容丰富度,使得生成的目标问题样本具有较好的多样性。

附图说明

- [0038] 图1是本申请实施例提供的问题样本生成方法的流程图;
- [0039] 图2是图1中的步骤S101的流程图;
- [0040] 图3是图2中的步骤S202的流程图;
- [0041] 图4是图1中的步骤S103的流程图;
- [0042] 图5是图1中的步骤S104的流程图;
- [0043] 图6是图5中的步骤S502的流程图;
- [0044] 图7是图5中的步骤S503的流程图;
- [0045] 图8是本申请实施例提供的问题样本生成装置的结构示意图;
- [0046] 图9是本申请实施例提供的电子设备的硬件结构示意图。

具体实施方式

[0047] 为了使本申请的目的、技术方案及优点更加清楚明白,以下结合附图及实施例,对本申请进行进一步详细说明。应当理解,此处所描述的具体实施例仅用以解释本申请,并不用于限定本申请。

[0048] 需要说明的是,虽然在装置示意图中进行了功能模块划分,在流程图中示出了逻辑顺序,但是在某些情况下,可以以不同于装置中的模块划分,或流程图中的顺序执行所示出或描述的步骤。说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。

[0049] 除非另有定义,本文所使用的所有的技术和科学术语与属于本申请的技术领域的人员通常理解的含义相同。本文中所使用的术语只是为了描述本申请实施例的目的,不是旨在限制本申请。

[0050] 首先,对本申请中涉及的若干名词进行解析:

[0051] 人工智能(artificial intelligence, AI):是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学;人工智能是计算机科学的一个分支,人工智能企图了解智能的实质,并生产出一种新的能以人类智能相似的方式做出反应的智能机器,该领域的研究包括机器人、语言识别、图像识别、自然语言处理和专家系统等。人工智能可以对人的意识、思维的信息过程的模拟。人工智能还是利用数字计算机或者数字计算机控制的机器模拟、延伸和扩展人的智能,感知环境、获取知识并使用知识获得最佳结果的理论、方法、技术及应用系统。

[0052] 自然语言处理(natural language processing, NLP):NLP用计算机来处理、理解

以及运用人类语言(如中文、英文等),NLP属于人工智能的一个分支,是计算机科学与语言学的交叉学科,又常被称为计算语言学。自然语言处理包括语法分析、语义分析、篇章理解等。自然语言处理常用于机器翻译、手写体和印刷体字符识别、语音识别及文语转换、信息意图识别、信息抽取与过滤、文本分类与聚类、舆情分析和观点挖掘等技术领域,它涉及与语言处理相关的数据挖掘、机器学习、知识获取、知识工程、人工智能研究和与语言计算相关的语言学研究等。

[0053] 信息抽取(Information Extraction):从自然语言文本中抽取指定类型的实体、关系、事件等事实信息,并形成结构化数据输出的文本处理技术。信息抽取是从文本数据中抽取特定信息的一种技术。文本数据是由一些具体的单位构成的,例如句子、段落、篇章,文本信息正是由一些小的具体的单位构成的,例如字、词、词组、句子、段落或是这些具体的单位的组合。抽取文本数据中的名词短语、人名、地名等都是文本信息抽取,当然,文本信息抽取技术所抽取的信息可以是各种类型的信息。

[0054] 提示学习(Prompt Learning):提示学习主要针对few-shot场景下从文本中抽取关系(知识检索、对话、问答)的问题,其核心思想是通过插入文本,将分类任务转化为MASK语言模型。对于关系抽取问题,选择提示模板需要较多的领域知识和较大工作量,且在实体和边之间蕴藏的大量知识也不应该被忽略。提示学习通过学习模板词和答案词,将实体和关系的知识注入模型,并在知识约束下协同优化它们的表示。提示学习通过构造提示模板和标签映射,能够显著提升模型的少样本、跨领域、多任务等能力。

[0055] 知识图谱(Knowledge Graph):是通过将应用数学、图形学、信息可视化技术、信息科学等学科的理论与方法与计量学引文分析、共现分析等方法结合,并利用可视化的图谱形象地展示学科的核心结构、发展历史、前沿领域以及整体知识架构达到多学科融合目的的现代理论。知识图谱主要目标是用来描述真实世界中存在的各种实体和概念,以及他们之间的强关系,我们用关系去描述两个实体之间的关联。在Web视角下,知识图谱如同简单文本之间的超链接一样,通过建立数据之间的语义链接,支持语义搜索。在自然语言处理视角下,知识图谱就是从文本中抽取语义和结构化的数据。在人工智能视角下,知识图谱是利用知识库来辅助理解人类语言的工具。在数据库视角下,知识图谱是利用图的方式去存储知识的方法。知识图谱是比较通用的语义知识的形式化描述框架,用节点表示语义符号,用边表示语义之间的关系。知识图谱旨在描述真实世界中存在的各种实体或概念及其关系,其构成一张巨大的语义网络图,节点表示实体或概念,边则由属性或关系构成。现在的知识图谱已被用来泛指各种大规模的知识库。知识图谱又被称作语义网络(semantic network),从早期开始,语义网络就推动了基于图的知识表示,例如,推动RDF标准的过程中,在这样一个基于图的知识表示体系里面,实体作为图的节点,节点之间的连线作为关系。在构建知识图谱的过程中,往往需要将文本向量化,因此基于文本数据的Word2Vec应运而生,其通过浅神经网络语言模型将每个词用一个向量表示,通过构建输入层、映射层和输出层,利用神经网络学习可预测在该词上下文中出现概率最大的词语。通过对文本词库的训练将文本转化为n维向量空间中的向量,并通过在空间中的余弦相似度代表词语在语义上的接近程度。

[0056] 实体:指具有可区别性且独立存在的某种事物。如某一个人、某一个城市、某一种植物等、某一种商品等等。世界万物有具体事物组成,此指实体。实体是知识图谱中的最基

本元素,不同的实体间存在不同的关系。

[0057] 概念:某一类实体的集合。

[0058] 语义类(概念):具有同种特性的实体构成的集合,如书籍、电脑等。概念主要指集合、类别、对象类型、事物的种类,例如人物、地理等。

[0059] 关系:实体与实体之间、不同的概念与概念之间、概念与实体之间存在的某种相互关系。关系形式化为一个函数,它把 kk 个点映射到一个布尔值。在知识图谱上,关系则是一个把 kk 个图节点(实体、语义类、属性值)映射到布尔值的函数。

[0060] 三元组:三元组($\{E,R\}$)是知识图谱的一种通用表示方式;三元组的基本形式主要包括(实体1-关系-实体2)和(实体-属性-属性值)等。每个实体(概念的外延)可用一个全局唯一确定的ID来标识,每个属性-属性值对(attribute-value pair,AVP)用来刻画实体的内在特性,而关系用来连接两个实体,刻画它们之间的关联。例如,在一个知识图谱的例子中,张三是一个实体,教师是一个实体,张三-职业-教师是一个(实体-关系-实体)的三元组样例,甲市是一个实体,人口是一种属性,2069.3万是属性值。甲市-人口-2069.3万构成一个(实体-属性-属性值)的三元组样例。

[0061] 网络爬虫:又称为网页蜘蛛,网络机器人,在FOAF社区中间,更经常的称为网页追逐者,是一种按照一定的规则,自动地抓取万维网信息的程序或者脚本。另外一些不常用的名字还有蚂蚁、自动索引、模拟程序或者蠕虫。

[0062] 数据清洗(Data cleaning):是指发现并纠正数据文件中可识别的错误的最后一道程序,包括检查数据一致性,处理无效值和缺失值等。与问卷审核不同,录入后的数据清理一般是由计算机而不是人工完成。

[0063] 数据去重:是找到数据文件集合中重复的数据并将其删除,只保留唯一的数据单元,从而消除冗余数据。数据去重包括完全去重和不完全去重。完全去重指的是消除完全重复的数据,完全重复的数据指的是数据表记录字段值完全一样的数据。不完全去重指的是在数据清洗中,所有字段值都相等的重复值是一定要剔除的。

[0064] 编码(encoder):将输入序列转化成一个固定长度的向量。

[0065] 解码(decoder):将之前生成的固定向量再转化成输出序列;其中,输入序列可以是文字、语音、图像、视频;输出序列可以是文字、图像。

[0066] 注意力机制(Attention Mechanism):注意力机制可以使得神经网络具备专注于其输入(或特征)子集的能力,选择特定的输入,可以应用于任何类型的输入而不管其形状如何。在计算能力有限情况下,注意力机制是解决信息超载问题的主要手段的一种资源分配方案,将计算资源分配给更重要的任务。

[0067] 随着人工智能技术的不断发展,基于知识图谱的问答(Knowledge Base Question Answer,KBQA)可以用来实现多种复杂问题的问答需求。

[0068] 目前,高质量的问答模型的训练大多数需要依赖于高质量的问题样本,而相关技术中常常是通过人工收集和人工编写的方式来获取问题样本,这一方式往往会由于收集时间和收集范围有限而影响问题样本的多样性,因此,如何提高问题样本的多样性,成为了亟待解决的技术问题。

[0069] 基于此,本申请实施例提供了一种问题样本生成方法、问题样本生成装置、电子设备及存储介质,旨在提高问答样本的多样性。

[0070] 本申请实施例提供的问题样本生成方法、问题样本生成装置、电子设备及存储介质,具体通过如下实施例进行说明,首先描述本申请实施例中的问题样本生成方法。

[0071] 本申请实施例可以基于人工智能技术对相关的数据进行获取和处理。其中,人工智能(Artificial Intelligence, AI)是利用数字计算机或者数字计算机控制的机器模拟、延伸和扩展人的智能,感知环境、获取知识并使用知识获得最佳结果的理论、方法、技术及应用系统。

[0072] 人工智能基础技术一般包括如传感器、专用人工智能芯片、云计算、分布式存储、大数据处理技术、操作/交互系统、机电一体化等技术。人工智能软件技术主要包括计算机视觉技术、机器人技术、生物识别技术、语音处理技术、自然语言处理技术以及机器学习/深度学习等几大方向。

[0073] 本申请实施例提供的问题样本生成方法,涉及人工智能技术领域。本申请实施例提供的问题样本生成方法可应用于终端中,也可应用于服务器端中,还可以是运行于终端或服务器端中的软件。在一些实施例中,终端可以是智能手机、平板电脑、笔记本电脑、台式计算机等;服务器端可以配置成独立的物理服务器,也可以配置成多个物理服务器构成的服务器集群或者分布式系统,还可以配置成提供云服务、云数据库、云计算、云函数、云存储、网络服务、云通信、中间件服务、域名服务、安全服务、CDN以及大数据和人工智能平台等基础云计算服务的云服务器;软件可以是实现问题样本生成方法的应用等,但并不局限于以上形式。

[0074] 本申请可用于众多通用或专用的计算机系统环境或配置中。例如:个人计算机、服务器计算机、手持设备或便携式设备、平板型设备、多处理器系统、基于微处理器的系统、置顶盒、可编程的消费电子设备、网络PC、小型计算机、大型计算机、包括以上任何系统或设备的分布式计算环境等等。本申请可以在由计算机执行的计算机可执行指令的一般上下文中描述,例如程序模块。一般地,程序模块包括执行特定任务或实现特定抽象数据类型的例程、程序、对象、组件、数据结构等等。也可以在分布式计算环境中实践本申请,在这些分布式计算环境中,由通过通信网络而被连接的远程处理设备来执行任务。在分布式计算环境中,程序模块可以位于包括存储设备在内的本地和远程计算机存储介质中。

[0075] 图1是本申请实施例提供的问题样本生成方法的一个可选的流程图,图1中的方法可以包括但不限于包括步骤S101至步骤S105。

[0076] 步骤S101,获取原始三元组数据,其中,原始三元组数据包括第一实体、关系以及第二实体;

[0077] 步骤S102,对原始三元组数据进行数据过滤,得到目标三元组数据;

[0078] 步骤S103,根据预设的提示标签对目标三元组数据进行模板生成,得到问题提示模板;

[0079] 步骤S104,通过预设的样本生成模型对问题提示模板进行文本生成,得到初始问题样本;

[0080] 步骤S105,根据预设的筛选条件对初始问题样本进行筛选处理,得到目标问题样本。

[0081] 本申请实施例所示意的步骤S101至步骤S105,通过获取原始三元组数据,其中,原始三元组数据包括第一实体、关系以及第二实体,对原始三元组数据进行数据过滤,得

到目标三元组数据,能够有效地剔除相关性较低的三元组数据以及异常数据,提高数据准确性。进一步地,根据预设的提示标签对目标三元组数据进行模板生成,得到问题提示模板,能够较为方便地捕捉到实体和关系之间的知识,从而得到能够用于构建问题样本的问题提示模板,通过预设的样本生成模型对问题提示模板进行文本生成,得到初始问题样本,能够较为方便地对问题提示模板的语义信息进行分析,得到多个符合需求的初始问题样本。最后,根据预设的筛选条件对初始问题样本进行筛选处理,得到目标问题样本,能够提高目标问题样本的质量,也能够提高目标问题样本的生成效率和目标问题样本的内容丰富度,使得生成的目标问题样本具有较好的多样性。

[0082] 需要说明的是,在本申请的各个具体实施方式中,当涉及到需要根据用户信息、用户行为数据,用户历史数据以及用户位置信息等与用户身份或特性相关的数据进行相关处理时,都会先获得用户的许可或者同意,而且,对这些数据的收集、使用和处理等,都会遵守相关国家和地区的相关法律法规和标准。此外,当本申请实施例需要获取用户的目标个人信息时,会通过弹窗或者跳转到确认页面等方式获得用户的单独许可或者单独同意,在明确获得用户的单独许可或者单独同意之后,再获取用于使本申请实施例能够正常运行的必要的用户相关数据。

[0083] 在一些实施例的步骤S101中,可以通过编写网络爬虫,设置好数据源之后对数据源的数据进行有目标性地爬取,得到原始三元组数据。也可以通过其他方式获取原始三元组数据,不限于此,其中,数据源可以是各种类型的网络平台,原始三元组数据包括第一实体、关系以及第二实体。该原始三元组形式具体表示为(第一实体->关系->第二实体)的形式,将结构化数据(例如,人物与地名的关系等)直接转换成三元组的表示形式,得到原始三元组数据。例如,人物与地名的三元组形式可以表示为(人物-出生地-地名)。

[0084] 请参阅图2,在一些实施例中,步骤S101可以包括但不限于包括步骤S201至步骤S202:

[0085] 步骤S201,获取预设的知识图谱库;

[0086] 步骤S202,通过预设的知识抽取模型对知识图谱库进行知识抽取,得到原始三元组数据。

[0087] 在一些实施例的步骤S201中,预设的知识图谱库可以根据预先获取到的结构化数据和结构化数据构建而成。具体地,可以通过预设的API接口链接到基于已有的实体之间的关系构建的关系库,从关系库直接提取出相应的结构化数据,其中,结构化数据主要包括不同领域的实体特征以及实体特征之间的关系。以医疗领域为例,结构化数据包括常见的疾病实体、疾病与症状的关系等等。对于非结构化数据,则可以通过编写网络爬虫,设置好数据源之后对数据源的数据进行有目标性地爬取,得到非结构化数据。非结构化数据包括不同领域的百科知识、新闻、咨询以及文章等等,以医疗领域为例,非结构化数据可以是健康话题、疾病描述、症状描述以及病因描述等方面的科普文章、新闻咨询等等。将获取到的结构化数据和非结构化数据统一存储在目标数据库中,从而得到该知识图谱库。

[0088] 在一些实施例的步骤S202中,知识抽取模型可以基于pipeline方法构建而成,其中,知识抽取模型包括实体抽取层和关系识别层,通过实体抽取层的词性类别对知识图谱库的非结构化数据进行词性标注,根据标注的词性标签对非结构化数据进行分类处理,得到每一词性标签对应的三元组特征,该词性标签可以包括主语标签、谓语标签、宾语标签

等。进而通过关系识别层对带有不同词性标签的三元组特征进行组合处理,得到原始三元组数据。

[0089] 在一些具体实施例中,原始三元组数据中的每一三元组包括第一实体、第二实体以及第一实体和第二实体之间的关系。例如,历史名人的背景分析、歌曲和歌手的对应关系等特征信息可以表示为三元组,更具体地,请参阅表1所示,根据对应的实体“张三”所提取到的二个三元组表示为(第一实体->关系->第二实体)的形式:(张三-职业-教师)、(张三-性别-男性),在第一三元组中,第一实体为“张三”,关系为“职业”,第二实体为“教师”;在第二三元组中,第一实体为“张三”,关系为“性别”,第二实体为“男性”。

[0090]

目标三元组	第一实体	关系	第二实体
张三-职业-教师	张三	职业	教师
张三-性别-男性	张三	性别	男性

[0091] 表1

[0092] 请参阅图3,在一些实施例中,知识抽取模型包括实体抽取层和关系识别层,步骤S202可以包括但不限于包括步骤S301至步骤S302:

[0093] 步骤S301,通过实体抽取层的词性类别标签对知识图谱库的数据进行标签分类,得到带有词性标签的三元组特征;

[0094] 步骤S302,通过关系识别层对三元组特征进行组合处理,得到原始三元组数据。

[0095] 在一些实施例的步骤S301中,通过词性类别标签和分类函数(如softmax函数、sigmioid函数等等)对知识图谱库的结构化数据和非结构化数据进行标签分类,以softmax函数为例,通过softmax函数创建结构化数据/非结构化数据在每一词性类别标签的概率分布,通过概率分布情况表征出结构化数据/非结构化数据属于每一词性类别标签的可能性,并将结构化数据/非结构化数据归类到概率值最大的词性类别标签对应的数据集合中,并对这些结构化数据/非结构化数据进行标注处理,从而得到带有词性标签的三元组特征,其中,该词性标签可以包括主语标签、谓语标签、宾语标签等。

[0096] 在一些实施例的步骤S302中,通过关系识别层对带有不同词性标签的三元组特征进行组合处理,得到原始三元组数据,其中,组合形式可以依据常规的语法规则,以主谓宾的形式提取出相应的三元组特征,并将这一系列的三元组特征进行组合处理,得到原始三元组数据,例如,原始三元组数据可以为(湖北-省会-武汉)、(湖北-简称-鄂)等等。

[0097] 在一些实施例的步骤S102中,对原始三元组数据进行数据过滤可以包括对原始三元组数据的数据清洗和数据去重等操作,具体地,对原始三元组数据进行数据清洗时,首先对原始三元组数据中的缺失值进行清洗,具体地,对原始三元组数据中的每个数据(包括三元组中的实体和关系)进行计算,得到缺失值比例,并按照缺失比例和数据重要性对原始三元组数据中的数据进行清洗。进一步地,对不重要的数据进行删除,并且根据业务知识或者经验推断等方式对缺失数据进行填充处理,得到填充数据。对于重要性较高且缺失率较高的数据,可以向业务人员咨询或者通过其他渠道取得对应的完整数据,从而得到初始三元组数据。进一步地,统计初始三元组数据中每一数据的字段值,将字段值相同的数据划分在同一个集合,对每个集合进行数据剔除,每个集合只保留其中的一个数据,从而实现对初始三元组数据的完全去重,使得每个数据为唯一数据,从而消除冗余数据,得到目标三元组数据。

[0098] 请参阅图4,在一些实施例中,步骤S103可以包括但不限于包括步骤S401 至步骤S402:

[0099] 步骤S401,根据提示标签对目标三元组数据进行特征定位,得到位置参数;

[0100] 步骤S402,根据位置参数和预设的提示模板对三元组数据进行模板生成,得到问题提示模板。

[0101] 在一些实施例的步骤S401中,提示标签可以是BIO标签,其中,在BIO标签中,B是指该部分所在的片段属于句子且处于句子的开头位置,I指该部分所在的片段属于句子且在处于句子的中间位置,O指该部分所在的片段属于句子且在处于句子的结尾位置。在特征定位时主要需要构建提示标签和目标三元组数据的标签映射,即根据目标三元组数据的实体以及实体与实体之间的关系,确定实体和关系在整个句子中所处的位置,得到位置参数,位置参数用于表征哪些实体以及关系应该放在起始位置,哪些实体以及关系应该放在中间位置,以及哪些实体以及关系应该放在结尾位置,从而得到每一实体/每一关系与提示标签之间的映射关系。

[0102] 在一些实施例的步骤S402中,预设的提示模板一般是由人工设计,比如,提示模板可以是:“请用三元组()生成句子:”,提示模板也可以构建为:“三元组()生成的句子是:”。具体选择哪个模板通过人工核查的方式来确定,即通过观察哪个提示模板生成的问题提示模板在问题样本生成模型中生成的句子效果更好来确定。在大多数实施例中,由于prefix模板更适用于NLG生成任务,因此,预设的提示模板一般为prefix模板。通过提示模板和位置参数对三元组数据进行模板生成,限定生成句子的位置为句子末尾,将位置参数为结束标签的句子放置在整体文本句子的末尾,从而得到问题提示模板。

[0103] 通过上述步骤S401至步骤S402能够较为方便地捕捉到实体和关系之间的知识,从而得到能够用于构建问题样本的问题提示模板,提高问题样本的生成多样性和生成效率。

[0104] 在一些实施例的步骤S104之前,问题样本生成方法包括预先训练样本生成模型,其中,样本生成模型可以基于BART模型、T5模型或者GPT模型等构建而成。样本生成模型包括掩码层、编码层以及解码层,其中,掩码层主要用于对问题提示模板的一些特点的元素进行选择或者屏蔽处理;编码层主要用于实现对经过掩码处理的问题提示模板从文本空间到隐向量空间的映射,从而提取问题提示模板的语义信息,解码层主要用于对提取到的语义信息进行重构处理,得到符合需求的问题样本。在对问题样本生成模型进行训练时,将样本数据输入至样本生成模型,并通过样本生成模型的损失函数来计算模型损失,其中,损失函数可以是常用的交叉熵损失函数,同时,可以采用梯度下降法来对模型损失进行反向传播至输入层(即掩码层),在反向传播的过程中,根据模型损失调整各种参数的值,不断迭代上述过程,直至收敛,以训练样本生成模型。

[0105] 请参阅图5,在一些实施例中,样本生成模型包括掩码层、编码层以及解码层,步骤S104可以包括但不限于包括步骤S501至步骤S503:

[0106] 步骤S501,通过掩码层对问题提示模板进行掩码处理,得到掩码问题矩阵;

[0107] 步骤S502,通过编码层对掩码问题矩阵进行编码处理,得到问题样本隐向量;

[0108] 步骤S503,通过解码层对问题样本隐向量进行解码处理,得到初始问题样本。

[0109] 在一些实施例的步骤S501中,通过掩码层的掩码向量对问题提示模板的一些特

定元素进行选择或者屏蔽处理,以实现问题提示模板的元素过滤,得到掩码问题矩阵,其中,掩码向量可以取值为0或者1,在一些其他实施例中,也可以采用其他向量作为掩码向量,不限于此。

[0110] 在一些实施例的步骤S502中,通过编码层对掩码问题矩阵进行编码处理,实现掩码问题矩阵从文本空间到隐向量空间的映射,从而提取问题提示模板的重要语义信息,得到问题样本隐向量。

[0111] 在一些实施例的步骤S503中,通过解码层对问题样本隐向量进行语义重构处理,得到符合需求的问题样本,即初始问题样本。

[0112] 通过上述步骤S501至步骤S503能够较为方便地对问题提示模板的语义信息进行分析,得到多个符合需求的初始问题样本。

[0113] 请参阅图6,在一些实施例,步骤S502包括但不限于包括步骤S601至步骤S602:

[0114] 步骤S601,通过编码层对掩码问题矩阵进行编码处理,得到问题样本特征;

[0115] 步骤S602,通过预设的第一注意力矩阵对问题样本特征进行注意力计算,得到问题样本隐向量。

[0116] 在一些实施例的步骤S601中,通过编码层对掩码问题矩阵进行编码处理,实现掩码问题矩阵从文本空间到隐向量空间的映射,从而提取问题提示模板的重要语义信息,得到问题样本特征。

[0117] 在一些实施例的步骤S602中,通过预设的第一注意力矩阵对问题样本特征进行注意力计算,获取问题样本特征的语义表征信息,根据每一问题样本特征的语义表征信息的重要程度,对每一问题样本特征赋予不同的权重值,并对所有的问题样本特征进行加权计算,得到问题样本隐向量。该加权计算的过程主要包括将每一向量形式的问题样本特征与对应的权重值相乘,得到加权问题向量,将所有的加权问题向量进行求和处理,得到问题样本隐向量。

[0118] 通过上述步骤S601至步骤S602能够有针对性地对提取出掩码问题矩阵中更为重要的语义信息,使得生成的问题样本隐向量更接近问题提示模板的真实语义内容,从而提高问题样本生成的准确性。

[0119] 请参阅图7,在一些实施例中,步骤S503可以包括但不限于包括步骤S701至步骤S702:

[0120] 步骤S701,通过解码层对问题样本隐向量进行解码处理,得到问题样本表征向量;

[0121] 步骤S702,通过预设的第二注意力矩阵对问题样本表征向量进行注意力计算,得到初始问题样本。

[0122] 在一些实施例的步骤S701中,通过解码层对问题样本隐向量进行解码处理,将问题样本隐向量由向量形式转换为文本序列形式,实现对问题样本隐向量的语义重构,得到问题样本表征向量。

[0123] 在一些实施例的步骤S702中,通过预设的第二注意力矩阵对问题样本表征向量进行注意力计算,获取问题样本表征向量的关键表征信息,根据每一问题样本表征向量的关键表征信息的重要程度,对每一问题样本表征向量赋予不同的权重值,并对所有的问题样本表征向量进行加权计算,得到初始问题样本。该加权计算的过程主要包括将每一向量

形式的问题样本表征向量与对应的权重值相乘,得到加权表征向量,将所有的加权表征向量进行求和处理,得到初始问题样本。例如,基于原始三元组数据<某某-出生地-湖南>,可以构成的初始问题样本包括“某某是哪里人?”、“某某的出生地是哪里?”以及“某某的出生地在哪里”等等。

[0124] 通过上述步骤S701至步骤S702能够有针对性地对提取出问题样本隐向量中的语义信息,通过不同的语义偏向来尽可能地构建初始问题样本,提高初始问题样本的样本数量,提高初始问题样本的多样性。

[0125] 在一些实施例的步骤S105中,根据预设的筛选条件对初始问题样本进行筛选处理,将不符合三元组信息的初始问题样本进行剔除,该过程可以通过将初始问题样本与预设的参考问题样本进行比对或者人工复核等方式来实现,从而得到目标问题样本。例如,初始问题样本包括“某某是哪里人?”、“某某的出生地是哪里?”、“某某的出生日期是多少”,由于当前的问答场景是询问地点,通过人工复核的方式发现“某某的出生日期是多少”这一初始问题样本不符合当前的问答场景,则将这一初始问题样本进行剔除。

[0126] 本申请实施例的问题样本生成方法,其通过获取原始三元组数据,其中,原始三元组数据包括第一实体、关系以及第二实体,对原始三元组数据进行数据过滤,得到目标三元组数据,能够有效地剔除相关性较低的三元组数据以及异常数据,提高数据准确性。进一步地,根据预设的提示标签对目标三元组数据进行模板生成,得到问题提示模板,能够较为方便地捕捉到实体和关系之间的知识,从而得到能够用于构建问题样本的问题提示模板,通过预设的样本生成模型对问题提示模板进行文本生成,得到初始问题样本,能够较为方便地对问题提示模板的语义信息进行分析,得到多个符合需求的初始问题样本。最后,根据预设的筛选条件对初始问题样本进行筛选处理,得到目标问题样本,能够提高目标问题样本的质量,也能够提高目标问题样本的生成效率和目标问题样本的内容丰富度,使得生成的目标问题样本具有较好的多样性。由于提示学习能够借助简单的数据信息就可以学习到大规模知识,并根据这些知识数据生成多样的问题样本,本申请通过引入提示学习这一思路来生成知识图谱的问题样本,能够获取到较高质量的问题样本,能够有效地解决基于知识图谱的问答过程的冷启动的问题,适用于各种问答场景下的问题生成,普适性较好。

[0127] 请参阅图8,本申请实施例还提供一种问题样本生成装置,可以实现上述问题样本生成方法,该装置包括:

[0128] 数据获取模块801,用于获取原始三元组数据,其中,原始三元组数据包括第一实体、关系以及第二实体;

[0129] 数据过滤模块802,用于对原始三元组数据进行数据过滤,得到目标三元组数据;

[0130] 模板生成模块803,用于根据预设的提示标签对目标三元组数据进行模板生成,得到问题提示模板;

[0131] 样本生成模块804,用于通过预设的样本生成模型对问题提示模板进行文本生成,得到初始问题样本;

[0132] 筛选模块805,用于根据预设的筛选条件对初始问题样本进行筛选处理,得到目标问题样本。

[0133] 在一些实施例中,数据获取模块801包括:

- [0134] 获取单元,用于获取预设的知识图谱库;
- [0135] 知识抽取单元,用于通过预设的知识抽取模型对知识图谱库进行知识抽取,得到原始三元组数据。
- [0136] 在一些实施例中,知识抽取模型包括实体抽取层和关系识别层,知识抽取单元包括:
- [0137] 标签分类子单元,用于通过实体抽取层的词性类别标签对知识图谱库的数据进行标签分类,得到带有词性标签的三元组特征;
- [0138] 组合子单元,用于通过关系识别层对三元组特征进行组合处理,得到原始三元组数据。
- [0139] 在一些实施例中,模板生成模块803包括:
- [0140] 定位单元,用于根据提示标签对目标三元组数据进行特征定位,得到位置参数;
- [0141] 生成单元,用于根据位置参数和预设的提示模板对三元组数据进行模板生成,得到问题提示模板。
- [0142] 在一些实施例中,样本生成模型包括掩码层、编码层以及解码层,样本生成模块804包括:
- [0143] 掩码单元,用于通过掩码层对问题提示模板进行掩码处理,得到掩码问题矩阵;
- [0144] 编码单元,用于通过编码层对掩码问题矩阵进行编码处理,得到问题样本隐向量;
- [0145] 解码单元,用于通过解码层对问题样本隐向量进行解码处理,得到初始问题样本。
- [0146] 在一些实施例中,编码单元包括:
- [0147] 编码子单元,用于通过编码层对掩码问题矩阵进行编码处理,得到问题样本特征;
- [0148] 第一计算子单元,用于通过预设的第一注意力矩阵对问题样本特征进行注意力计算,得到问题样本隐向量。
- [0149] 在一些实施例中,解码单元包括:
- [0150] 解码子单元,用于通过解码层对问题样本隐向量进行解码处理,得到问题样本表征向量;
- [0151] 第二计算子单元,用于通过预设的第二注意力矩阵对问题样本表征向量进行注意力计算,得到初始问题样本。
- [0152] 该问题样本生成装置的具体实施方式与上述问题样本生成方法的具体实施例基本相同,在此不再赘述。
- [0153] 本申请实施例还提供了一种电子设备,电子设备包括:存储器、处理器、存储在存储器上并可在处理器上运行的程序以及用于实现处理器和存储器之间的连接通信的数据总线,程序被处理器执行时实现上述问题样本生成方法。该电子设备可以为包括平板电脑、车载电脑等任意智能终端。
- [0154] 请参阅图9,图9示意了另一实施例的电子设备的硬件结构,电子设备包括:
- [0155] 处理器901,可以采用通用的CPU(CentralProcessingUnit,中央处理器)、微处理器、应用专用集成电路(ApplicationSpecificIntegratedCircuit,ASIC)、或者一个或多个

个集成电路等方式实现,用于执行相关程序,以实现本申请实施例所提供的技术方案;

[0156] 存储器902,可以采用只读存储器(ReadOnlyMemory,ROM)、静态存储设备、动态存储设备或者随机存取存储器(RandomAccessMemory,RAM)等形式实现。存储器902可以存储操作系统和其他应用程序,在通过软件或者固件来实现本说明书实施例所提供的技术方案时,相关的程序代码保存在存储器902中,并由处理器901来调用执行本申请实施例的问题样本生成方法;

[0157] 输入/输出接口903,用于实现信息输入及输出;

[0158] 通信接口904,用于实现本设备与其他设备的通信交互,可以通过有线方式(例如USB、网线等)实现通信,也可以通过无线方式(例如移动网络、WIFI、蓝牙等)实现通信;

[0159] 总线905,在设备的各个组件(例如处理器901、存储器902、输入/输出接口903和通信接口904)之间传输信息;

[0160] 其中处理器901、存储器902、输入/输出接口903和通信接口904通过总线905实现彼此之间在设备内部的通信连接。

[0161] 本申请实施例还提供了一种存储介质,存储介质为计算机可读存储介质,用于计算机可读存储,存储介质存储有一个或者多个程序,一个或者多个程序可被一个或者多个处理器执行,以实现上述问题样本生成方法。

[0162] 存储器作为一种非暂态计算机可读存储介质,可用于存储非暂态软件程序以及非暂态性计算机可执行程序。此外,存储器可以包括高速随机存取存储器,还可以包括非暂态存储器,例如至少一个磁盘存储器件、闪存器件、或其他非暂态固态存储器件。在一些实施方式中,存储器可选包括相对于处理器远程设置的存储器,这些远程存储器可以通过网络连接至该处理器。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

[0163] 本申请实施例提供的问题样本生成方法、问题样本生成装置、电子设备及存储介质,其通过获取原始三元组数据,其中,原始三元组数据包括第一实体、关系以及第二实体,对原始三元组数据进行数据过滤,得到目标三元组数据,能够有效地剔除相关性较低的三元组数据以及异常数据,提高数据准确性。进一步地,根据预设的提示标签对目标三元组数据进行模板生成,得到问题提示模板,能够较为方便地捕捉到实体和关系之间的知识,从而得到能够用于构建问题样本的问题提示模板,通过预设的样本生成模型对问题提示模板进行文本生成,得到初始问题样本,能够较为方便地对问题提示模板的语义信息进行分析,得到多个符合需求的初始问题样本。最后,根据预设的筛选条件对初始问题样本进行筛选处理,得到目标问题样本,能够提高目标问题样本的质量,也能够提高目标问题样本的生成效率和目标问题样本的内容丰富度,使得生成的目标问题样本具有较好的多样性。由于提示学习能够借助简单的数据信息就可以学习到大规模知识,并根据这些知识数据生成多样的问题样本,本申请通过引入提示学习这一思路来生成知识图谱的问题样本,能够获取到较高质量的问题样本,能够有效地解决基于知识图谱的问答过程的冷启动的问题,适用于各种问答场景下的问题生成,普适性较好。

[0164] 本申请实施例描述的实施例是为了更加清楚的说明本申请实施例的技术方案,并不构成对于本申请实施例提供的技术方案的限定,本领域技术人员可知,随着技术的演变和新应用场景的出现,本申请实施例提供的技术方案对于类似的技术问题,同样适用。

[0165] 本领域技术人员可以理解的是,图1-7中示出的技术方案并不构成对本申请实施例的限定,可以包括比图示更多或更少的步骤,或者组合某些步骤,或者不同的步骤。

[0166] 以上所描述的装置实施例仅仅是示意性的,其中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,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模块来实现本实施例方案的目的。

[0167] 本领域普通技术人员可以理解,上文中所公开方法中的全部或某些步骤、系统、设备中的功能模块/单元可以被实施为软件、固件、硬件及其适当的组合。

[0168] 本申请的说明书及上述附图中的术语“第一”、“第二”、“第三”、“第四”等(如果存在)是用于区别类似的对象,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或描述的那些以外的顺序实施。此外,术语“包括”和“具有”以及他们的任何变形,意图在于覆盖不排他的包含,例如,包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元,而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0169] 应当理解,在本申请中,“至少一个(项)”是指一个或者多个,“多个”是指两个或两个以上。“和/或”,用于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:只存在A,只存在B以及同时存在A和B三种情况,其中A,B可以是单数或者复数。字符“/”一般表示前后关联对象是一种“或”的关系。“以下至少一项(个)”或其类似表达,是指这些项中的任意组合,包括单项(个)或复数项(个)的任意组合。例如,a,b或c中的至少一项(个),可以表示:a,b,c,“a和b”,“a和c”,“b和c”,或“a和b和c”,其中a,b,c可以是单个,也可以是多个。

[0170] 在本申请所提供的几个实施例中,应该理解到,所揭露的装置和方法,可以通过其它的方式实现。例如,以上所描述的装置实施例仅仅是示意性的,例如,上述单元的划分,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。

[0171] 上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。

[0172] 另外,在本申请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。

[0173] 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,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,本申请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,包括多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)执行本申请各个实施例的方法的全部或部

分步骤。而前述的存储介质包括：U盘、移动硬盘、只读存储器 (Read-Only Memory, 简称 ROM)、随机存取存储器 (Random Access Memory, 简称RAM)、磁碟或者光盘等各种可以存储程序的介质。

[0174] 以上参照附图说明了本申请实施例的优选实施例,并非因此局限本申请实施例的权利范围。本领域技术人员不脱离本申请实施例的范围和实质内所作的任何修改、等同替换和改进,均应在本申请实施例的权利范围之内。

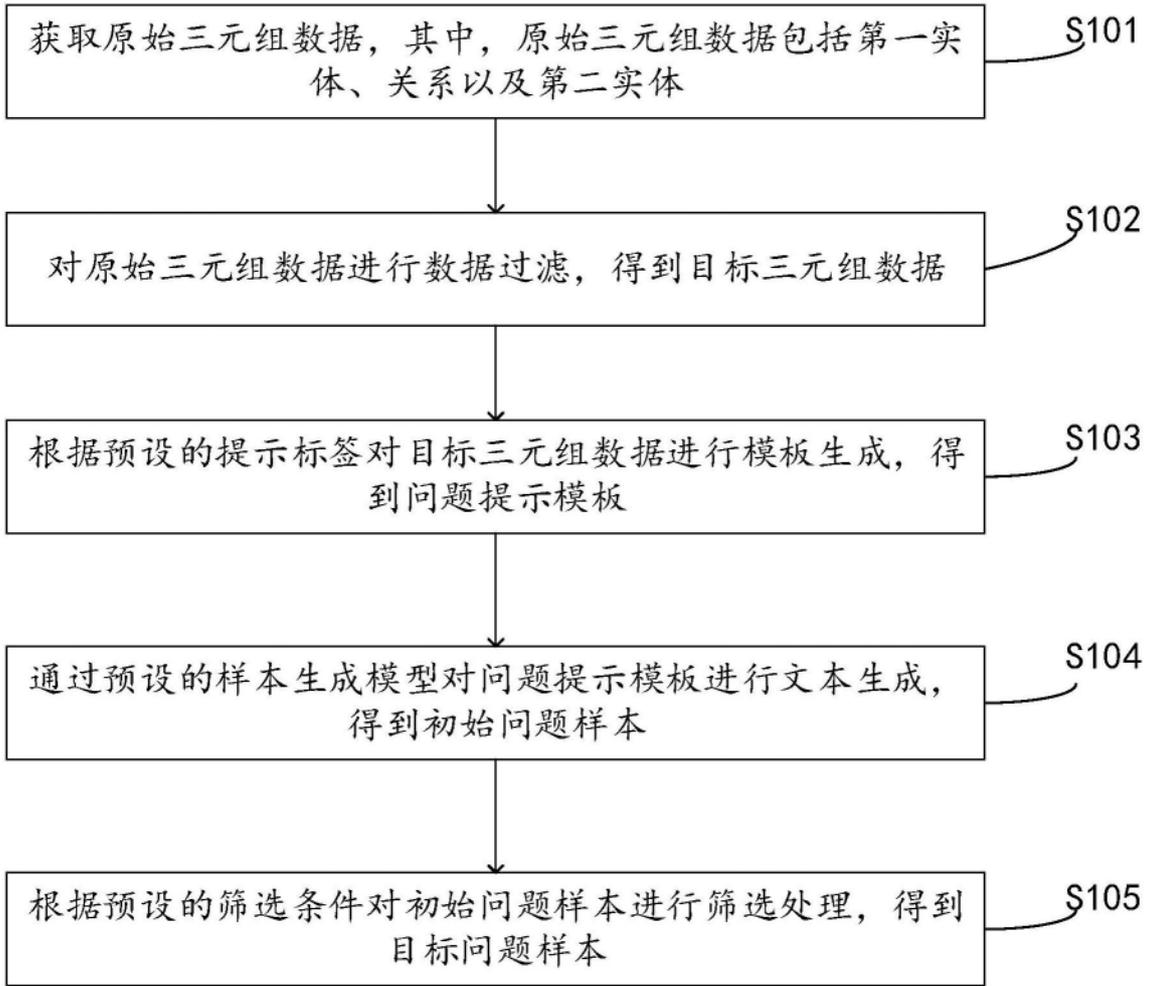


图1

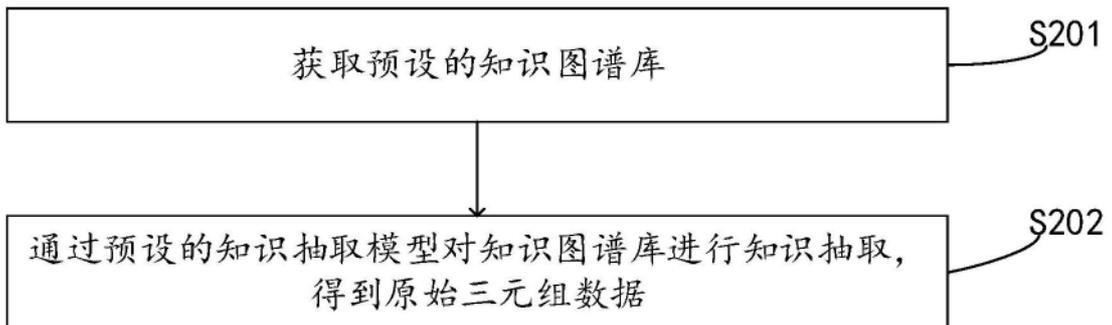


图2

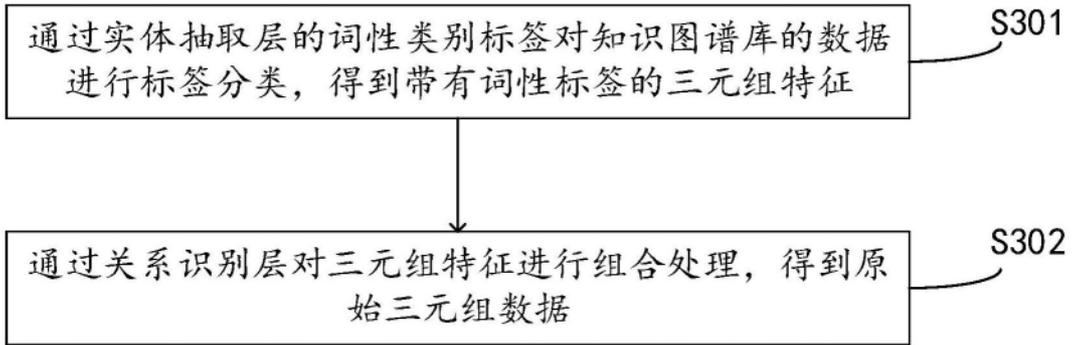


图3

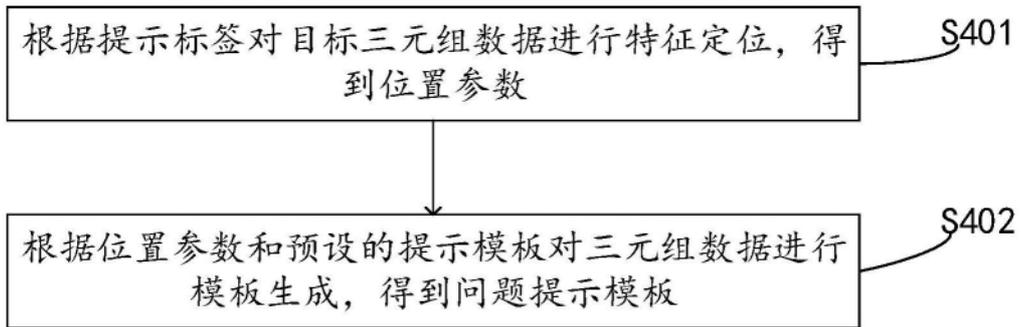


图4

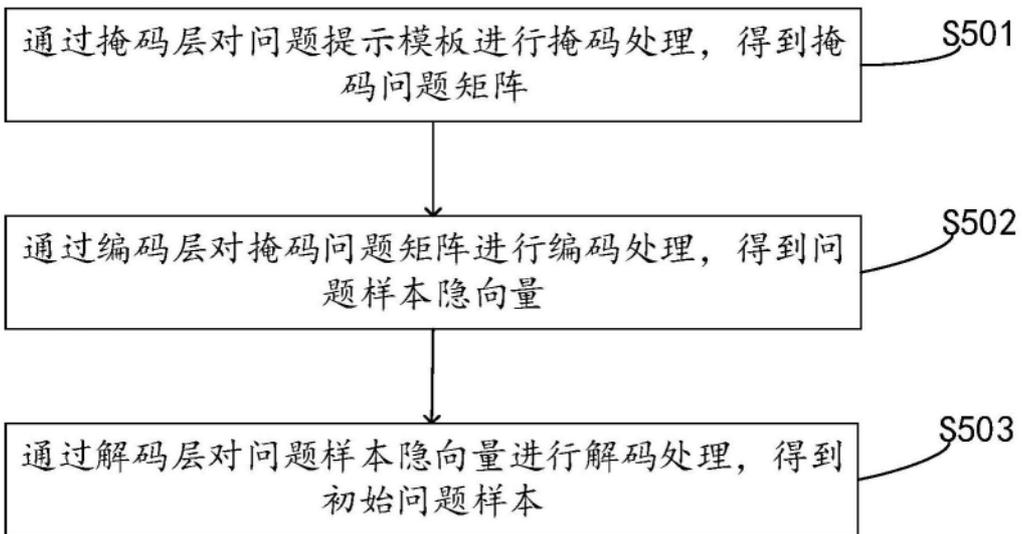


图5

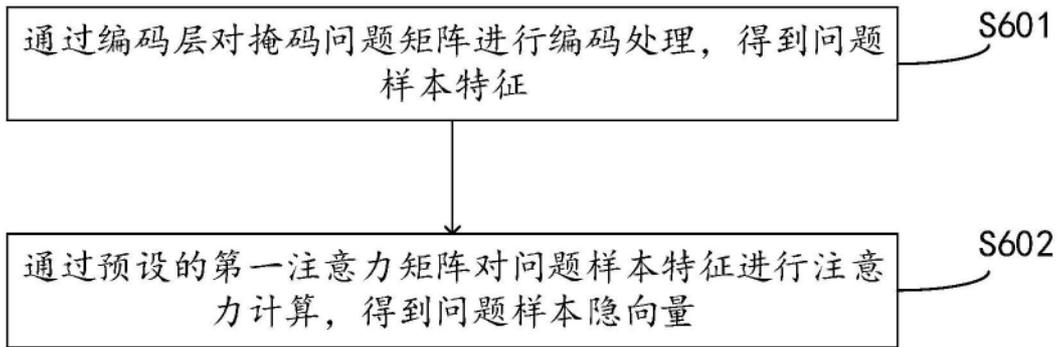


图6

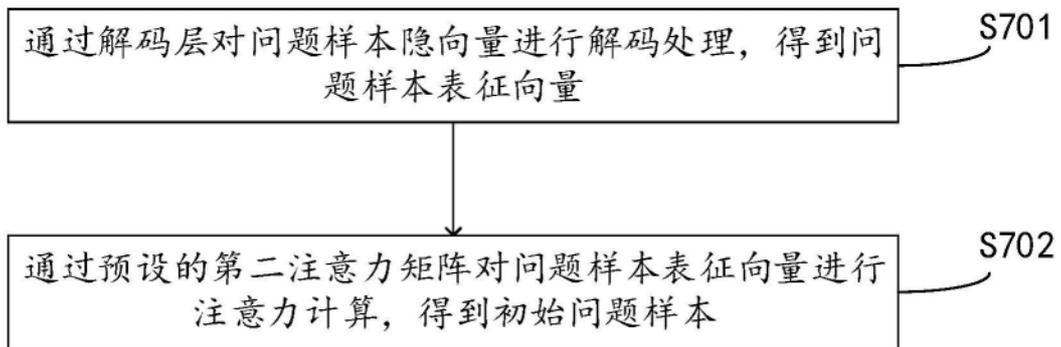


图7

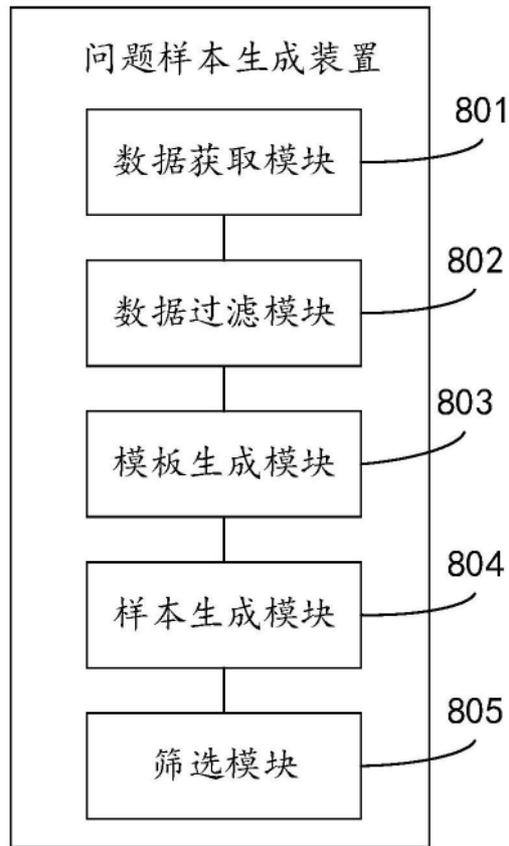


图8

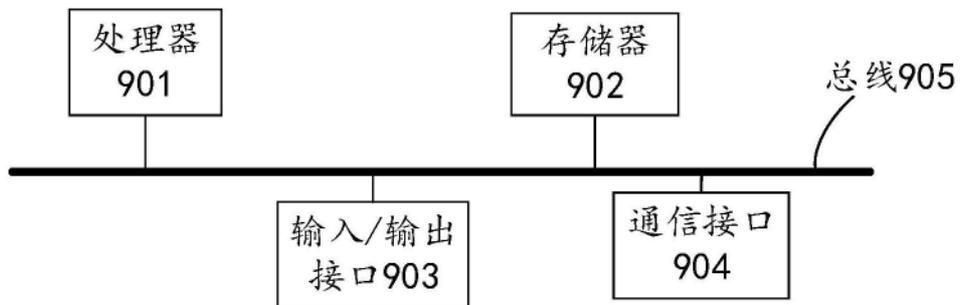


图9