



# (12) 发明专利申请

(10) 申请公布号 CN 114510561 A

(43) 申请公布日 2022.05.17

(21) 申请号 202210144860.7

G06N 3/04 (2006.01)

(22) 申请日 2022.02.17

G06N 3/08 (2006.01)

(71) 申请人 腾讯科技(深圳)有限公司

地址 518057 广东省深圳市南山区高新区  
科技中一路腾讯大厦35层

(72) 发明人 谢润泉 张恒 邵纪春

(74) 专利代理机构 北京三高永信知识产权代理  
有限责任公司 11138

专利代理师 李文静

(51) Int. Cl.

G06F 16/332 (2019.01)

G06F 16/33 (2019.01)

G06F 16/35 (2019.01)

G06F 40/30 (2020.01)

G06K 9/62 (2022.01)

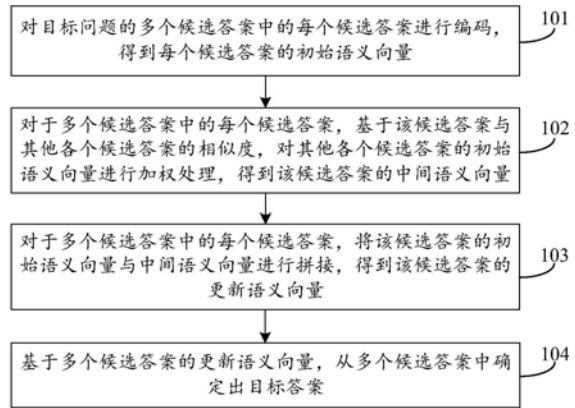
权利要求书2页 说明书17页 附图7页

## (54) 发明名称

答案选择方法、装置、设备及存储介质

## (57) 摘要

本申请公开了一种答案选择方法、装置、设备及存储介质,对于每个候选答案,本申请提供的方案能够基于该候选答案与其他各个候选答案的相似度,得到候选答案的中间语义向量,并将该中间语义向与初始语义向量进行拼接,得到候选答案的更新语义向量。由于该更新语义向量不仅包含了候选答案本身的信息,还包含了其他候选答案对其的支持性证据信息,因此可以确保基于多个候选答案的更新语义向量,从该多个候选答案中确定出的目标答案的准确性较高。



1. 一种答案选择方法,其特征在于,所述方法包括:

对目标问题的多个候选答案中的每个候选答案进行编码,得到每个候选答案的初始语义向量;

对于所述多个候选答案中的每个候选答案,基于所述候选答案与其他各个候选答案的相似度,对其他各个候选答案的初始语义向量进行加权处理,得到所述候选答案的中间语义向量;

对于所述多个候选答案中的每个候选答案,将所述候选答案的初始语义向量与中间语义向量进行拼接,得到所述候选答案的更新语义向量;

基于所述多个候选答案的更新语义向量,从所述多个候选答案中确定出目标答案。

2. 根据权利要求1所述的方法,其特征在于,所述对于所述多个候选答案中的每个候选答案,将所述候选答案的初始语义向量与中间语义向量进行拼接,得到所述候选答案的更新语义向量,包括:

对于所述多个候选答案中的每个候选答案,将所述候选答案的初始语义向量、所述候选答案的中间语义向量和所述候选答案的乘积向量进行拼接,得到所述候选答案的更新语义向量;

其中,所述乘积向量为所述候选答案的初始语义向量与所述候选答案的中间语义向量的向量积。

3. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

对于所述多个候选答案中的每个候选答案,基于所述候选答案的初始语义向量与其他各个候选答案的初始语义向量的点乘,确定所述候选答案与其他各个候选答案的相似度。

4. 根据权利要求1至3任一所述的方法,其特征在于,所述对目标问题的多个候选答案中的每个候选答案进行编码,得到每个候选答案的初始语义向量,包括:

对于所述目标问题的多个候选答案中的每个候选答案,将所述目标问题与所述候选答案进行拼接,得到拼接文本;

对所述拼接文本进行编码,得到所述候选答案的初始语义向量。

5. 根据权利要求1至3任一所述的方法,其特征在于,在对目标问题的多个候选答案中的每个候选答案进行编码之前,所述方法还包括:

基于所述目标问题,获取与所述目标问题关联的多个文档;

针对所述多个文档中的每个文档,获取所述目标问题的一个候选答案,得到多个候选答案。

6. 根据权利要求1至3任一所述的方法,其特征在于,所述基于所述多个候选答案的更新语义向量,从所述多个候选答案中确定出目标答案,包括:

将所述多个候选答案的更新语义向量输入至答案选择模型;

基于所述答案选择模型的输出结果,从所述多个候选答案中确定出目标答案。

7. 根据权利要求6所述的方法,其特征在于,所述方法还包括:

获取问题样本的多个答案样本,以及每个所述答案样本的标签,所述标签用于指示所述答案样本是否为所述问题样本的正确答案;

对所述多个答案样本中的每个答案样本进行编码,得到每个答案样本的初始语义向量;

对于所述多个答案样本中的每个答案样本,基于所述答案样本与其他各个答案样本的相似度,对其他各个答案样本的初始语义向量进行加权处理,得到所述答案样本的中间语义向量;

对于所述多个答案样本中的每个答案样本,将所述答案样本的初始语义向量和中间语义向量进行拼接,得到所述答案样本的更新语义向量;

基于所述多个答案样本的更新语义向量以及所述多个答案样本的标签,训练答案选择模型。

8. 一种答案选择装置,其特征在于,所述装置包括:

编码模块,用于对目标问题的多个候选答案中的每个候选答案进行编码,得到每个候选答案的初始语义向量;

加权模块,用于对于所述多个候选答案中的每个候选答案,基于所述候选答案与其他各个候选答案的相似度,对其他各个候选答案的初始语义向量进行加权处理,得到所述候选答案的中间语义向量;

拼接模块,用于对于所述多个候选答案中的每个候选答案,将所述候选答案的初始语义向量与中间语义向量进行拼接,得到所述候选答案的更新语义向量;

确定模块,用于基于所述多个候选答案的更新语义向量,从所述多个候选答案中确定出目标答案。

9. 一种计算机设备,其特征在于,所述计算机设备包括处理器和存储器,所述存储器中存储有至少一条指令、至少一段程序、代码集或指令集,所述至少一条指令、所述至少一段程序、所述代码集或指令集由所述处理器加载并执行以实现如权利要求1至7任一所述的答案选择方法。

10. 一种计算机可读存储介质,其特征在于,所述存储介质中存储有至少一条指令、至少一段程序、代码集或指令集,所述至少一条指令、所述至少一段程序、所述代码集或指令集由处理器加载并执行以实现如权利要求1至7任一所述的答案选择方法。

## 答案选择方法、装置、设备及存储介质

### 技术领域

[0001] 本申请涉及自然语言处理(Natural Language Processing, NLP)领域,特别涉及一种答案选择方法、装置、设备及存储介质。

### 背景技术

[0002] 近年来,基于文本的问答(Question Answering, QA)系统被广泛应用生活中的各个领域(例如医疗领域)。对于用户给出的文本形式的问题,该基于文本的问答系统能够通过检索语料库、知识图谱或问答知识库返回简洁且准确的匹配答案。

[0003] 相关技术中,基于文本的问答系统确定匹配答案的过程一般包括检索、抽取和答案选择三个阶段。其中,在检索阶段,问答系统能够获取多个包含有答案的文档。在抽取阶段,问答系统可以从获取到的多个文档中确定出能够回答问题的多个候选答案。之后,在答案选择阶段中,问答系统能够基于每个候选答案与问题在语法或者语义上的匹配程度,确定每个候选答案的正确性,并从该多个候选答案中选择出正确性最高的一个候选答案作为最佳答案。

[0004] 然而,仅基于候选答案与问题的匹配程度并不能准确地确定出的候选答案的正确性,上述答案选择方法所确定出的最佳答案的准确性较低。

### 发明内容

[0005] 本申请提供了一种答案选择方法、装置、设备及存储介质,能够有效提高答案选择方法所确定出的最佳答案的准确性。所述技术方案如下:

[0006] 一方面,提供了一种答案选择方法,所述方法包括:

[0007] 对目标问题的多个候选答案中的每个候选答案进行编码,得到每个候选答案的初始语义向量;

[0008] 对于所述多个候选答案中的每个候选答案,基于所述候选答案与其他各个候选答案的相似度,对其他各个候选答案的初始语义向量进行加权处理,得到所述候选答案的中间语义向量;

[0009] 对于所述多个候选答案中的每个候选答案,将所述候选答案的初始语义向量与中间语义向量进行拼接,得到所述候选答案的更新语义向量;

[0010] 基于所述多个候选答案的更新语义向量,从所述多个候选答案中确定出目标答案。

[0011] 另一方面,提供了一种答案选择模型的训练方法,所述方法包括:

[0012] 获取问题样本的多个答案样本,以及每个所述答案样本的标签,所述标签用于指示所述答案样本是否为所述问题样本的正确答案;

[0013] 对所述多个答案样本中的每个答案样本进行编码,得到每个答案样本的初始语义向量;

[0014] 对于所述多个答案样本中的每个答案样本,基于所述答案样本与其他各个答案样

本的相似度,对其他各个答案样本的初始语义向量进行加权处理,得到所述答案样本的中间语义向量;

[0015] 对于所述多个答案样本中的每个答案样本,将所述答案样本的初始语义向量和所述答案样本的中间语义向量进行拼接,得到所述答案样本的更新语义向量;

[0016] 基于所述多个答案样本的更新语义向量以及所述多个答案样本的标签,训练答案选择模型。

[0017] 又一方面,提供了一种答案选择装置,所述装置包括:

[0018] 编码模块,用于对目标问题的多个候选答案中的每个候选答案进行编码,得到每个候选答案的初始语义向量;

[0019] 加权模块,用于对于所述多个候选答案中的每个候选答案,基于所述候选答案与其他各个候选答案的相似度,对其他各个候选答案的初始语义向量进行加权处理,得到所述候选答案的中间语义向量;

[0020] 拼接模块,用于对于所述多个候选答案中的每个候选答案,将所述候选答案的初始语义向量与中间语义向量进行拼接,得到所述候选答案的更新语义向量;

[0021] 确定模块,用于基于所述多个候选答案的更新语义向量,从所述多个候选答案中确定出目标答案。

[0022] 再一方面,提供了一种答案选择模型的训练装置,所述装置包括:

[0023] 获取模块,用于获取问题样本的多个答案样本,以及每个所述答案样本的标签,所述标签用于指示所述答案样本是否为所述问题样本的正确答案;

[0024] 编码模块,用于对于所述多个答案样本中的每个答案样本进行编码,得到每个答案样本的初始语义向量;

[0025] 加权模块,用于对于所述多个答案样本中的每个答案样本,基于所述答案样本与其他各个答案样本的相似度,对其他各个答案样本的初始语义向量进行加权处理,得到所述答案样本的中间语义向量;

[0026] 拼接模块,用于对于所述多个候选答案中的每个候选答案,将所述答案样本的初始语义向量和中间语义向量进行拼接,得到所述答案样本的更新语义向量;

[0027] 训练模块,用于基于所述多个答案样本的更新语义向量以及所述多个答案样本的标签,训练答案选择模型。

[0028] 再一方面,提供了一种计算机设备,所述计算机设备包括处理器和存储器,所述存储器中存储有至少一条指令、至少一段程序、代码集或指令集,所述至少一条指令、所述至少一段程序、所述代码集或指令集由所述处理器加载并执行以实现如上述方面所提供的答案选择方法,或如上述方面所提供的答案选择模型的训练方法。

[0029] 再一方面,提供了一种计算机可读存储介质,所述存储介质中存储有至少一条指令、至少一段程序、代码集或指令集,所述至少一条指令、所述至少一段程序、所述代码集或指令集由处理器加载并执行以实现如上述方面所提供的答案选择方法,或如上述方面所提供的答案选择模型的训练方法。

[0030] 再一方面,提供了一种计算机程序产品或计算机程序,该计算机程序产品或计算机程序包括计算机指令,该计算机指令存储在计算机可读存储介质中。计算机设备的处理器从计算机可读存储介质读取该计算机指令,处理器执行该计算机指令,使得该计算机设

备执行上述方面所提供的答案选择方法,或上述方面提供的答案选择模型的训练方法。

[0031] 本申请提供的技术方案带来的有益效果至少包括:

[0032] 本申请提供了一种答案选择方法、装置、设备及存储介质。对于每个候选答案,本申请提供的方案能够基于该候选答案与其他各个候选答案的相似度,得到候选答案的中间语义向量,并将该中间语义向与初始语义向量进行拼接,得到候选答案的更新语义向量。由于该更新语义向量不仅包含了候选答案本身的信息,还包含了其他候选答案对其的支持性证据信息,因此可以确保基于多个候选答案的更新语义向量,从该多个候选答案中确定出的目标答案的准确性较高。

### 附图说明

[0033] 为了更清楚地说明本申请实施例中的技术方案,下面将对实施例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申请的一些实施例,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。

[0034] 图1是本申请实施例提供的一种问答系统的结构示意图;

[0035] 图2是本申请实施例提供的一种答案选择方法的流程图;

[0036] 图3是本申请实施例提供的另一种答案选择方法的流程图;

[0037] 图4是本申请实施例提供的一种服务器的结构框图;

[0038] 图5是本申请实施例提供的一种BERT模型的结构示意图;

[0039] 图6是本申请实施例提供的又一种答案选择方法的流程图;

[0040] 图7是本申请实施例提供的一种答案选择模型的训练方法的流程图;

[0041] 图8是本申请实施例提供的一种答案选择装置的结构框图;

[0042] 图9是本申请实施例提供的另一种答案选择装置的结构框图;

[0043] 图10是本申请实施例提供的一种答案选择模型的训练装置的结构框图;

[0044] 图11是本申请实施例提供的一种计算机设备的结构框图。

### 具体实施方式

[0045] 为使本申请的目的、技术方案和优点更加清楚,下面将结合附图对本申请实施方式作进一步地详细描述。

[0046] 本申请实施例提供的方案可以应用于云技术、人工智能或智慧交通等各种场景。首先,针对本申请实施例中涉及的名词进行简单介绍。

[0047] 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。换句话说,人工智能是计算机科学的一个综合技术,它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。人工智能也就是研究各种智能机器的设计原理与实现方法,使机器具有感知、推理与决策的功能。

[0048] 人工智能技术是一门综合学科,涉及领域广泛,既有硬件层面的技术也有软件层面的技术。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、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。

[0049] 随着人工智能技术研究和进步,人工智能技术在多个领域展开研究和应用,例如常见的智能家居、智能穿戴设备、虚拟助理、智能音箱、智能营销、无人驾驶、自动驾驶、无人机、机器人、智能医疗以及智能客服等,相信随着技术的发展,人工智能技术将在更多的领域得到应用,并发挥越来越重要的价值。

[0050] 机器学习(Machine Learning,ML):是一门多领域交叉学科,涉及概率论、统计学、逼近论、凸分析、算法复杂度理论等多门学科。专门研究计算机怎样模拟或实现人类的学习行为,以获取新的知识或技能,重新组织已有的知识结构使之不断改善自身的性能。机器学习是人工智能的核心,是使计算机具有智能的根本途径,其应用遍及人工智能的各个领域。机器学习和深度学习通常包括人工神经网络、置信网络、强化学习、迁移学习、归纳学习、示教学习等技术。

[0051] 自然语言处理:是计算机科学领域与人工智能领域中的一个重要方向。它研究能实现人与计算机之间用自然语言进行有效通信的各种理论和方法。自然语言处理是一门融语言学、计算机科学、数学于一体的科学。因此,这一领域的研究将涉及自然语言,即人们日常使用的语言,所以它与语言学的研究有着密切的联系。自然语言处理技术通常包括文本处理、语义理解、机器翻译、机器人问答、知识图谱等技术。

[0052] 问答系统:对于一个用户给出的自由文本形式的问题,问答系统通过检索语料库、知识图谱或问答知识库返回简洁、准确的匹配答案。问答系统可分为:基于结构化数据的问答系统,如基于知识图谱的问答(Knowledge Base Question Answering,KBQA),基于文本的问答(Document Question Answering,DocQA)系统,如机器阅读理解(Machine Reading Comprehension,MRC),基于问答对的问答系统如常见问题解答(Frequently Asked Questions,FAQ)系统。

[0053] 机器阅读理解:利用算法使计算机理解文章语义并回答相关问题的技术,常见问题类型有填空题、多项选择题以及问答题等。在DocQA中,通常可以采用片段抽取式的阅读理解技术,从问题相关的文档中抽取出连续的片段作为候选答案。机器阅读理解可以分为单文档MRC和多文档MRC。其中,单文档MRC能够对于单个问题,从单个召回文档中抽取能回答问题的连续片段作为答案。多文档MRC能够对个单个问题,从多个召回文档中抽取能回答问题的连续片段作为答案。区别于单文档MRC,多文档MRC的方法利用了同个问题下不同文档以及不同文档抽取的答案信息。

[0054] 答案选择:判断从检索和抽取步骤中得到的候选答案(通常这样的答案有多个)是否是对应问题的真正答案,并从这些真正的答案中选取一个答案作为最佳答案。

[0055] 对于答案选择,相关技术一般是基于问题和候选答案在语法或者语义上的匹配程度,确定每个候选答案的正确性,进而根据各个候选答案的正确性的高低,选择最佳答案。但是,在很多问答场景中,某个问题与其最佳答案在语法或者语义上的匹配程度并不高。例如,对于问题“为什么晚上不睡觉对身体危害很大”,其正确答案(或者称为最佳答案)中的关键词可以包括“长痘痘”,“毛孔粗大”和“熊猫眼”等,上述关键词与问题在语义和语法上并不匹配。此外,同一个问题可以有多个同义表示,但其对应的候选答案所包括的关键词一般是固定的。例如,对于一个关于某种病症的问题,该病症一般具有多个别称,但针对该病症问题的候选答案(例如症状、病因以及治疗方式)基本是固定的。由此,也会导致候选答案

与问题在语义和语法上的匹配程度较低。

[0056] 基于上述分析可知,问题和其对应的正确答案之间一般是因果关系。因此仅基于问题与候选答案在语义上的关系并不能准确地确定出候选答案的正确性。

[0057] 图1是本申请实施例提供的一种问答系统的结构示意图。参见图1,该系统包括:服务器110和终端设备120。该服务器110与终端设备120之间建立有线或无线通信连接。可选地,该服务器110可以是独立的物理服务器,也可以是多个物理服务器构成的服务器集群或者分布式系统,还可以是提供云计算服务的云服务器。该终端设备120可以是个人计算机(Personal Computer,PC)、车载终端、平板电脑、智能手机、可穿戴设备、智能机器人等具备数据计算、处理和存储能力的终端。

[0058] 作为第一种可能的示例,该问答系统中的终端设备120可以用于获取目标问题,并将该目标问题发送至服务器110。服务器110进而可以基于该目标问题,在相关数据库中检索和抽取该目标问题的候选答案。之后,该服务器110可以从该多个候选答案中选择出该目标问题的目标答案(即最佳答案),并将该最佳答案反馈至终端设备120。

[0059] 示例的,该问答系统可以是医疗领域的医疗知识问答系统。相应的,该目标问题是针对某个医疗知识的提问文本。该问答系统可以在医疗数据库(包含各种医学期刊或论文)中选择出能够回答该目标问题的最佳答案。

[0060] 作为第二种可能的示例,该问答系统可以只包括终端设备120。该终端设备120获取到目标问题后,可以直接在相关数据库中检索和抽取该目标问题的多个候选答案。终端设备120在该多个候选答案中选择出该目标问题的目标答案后,直接输出该目标答案。

[0061] 可选地,在上述两种示例中,问答系统可以通过答案选择模型,从多个候选答案中确定出最佳答案。下文对该答案选择模型的训练过程做简要介绍。

[0062] 在上述第一种可能的示例中,服务器110中预先存储有一个初始模型。该服务器110可以获取多个训练数据,每个训练数据包括一个问题样本,该问题样本对应的多个答案样本,以及每个答案样本的标签。服务器110可以对每个训练数据中的多个答案样本进行处理,并采用处理后的答案样本以及答案样本的标签对该初始模型进行训练,以得到能够执行答案选择任务的答案选择模型。其中,该多个训练数据可以是服务器110中预先存储的,也可以是终端设备120发送至服务器110的。

[0063] 在上述第二种可能的示例中,服务器110可以采用训练数据训练答案选择模型,并将训练得到的答案选择模型发送至终端设备120。该终端设备120进而可以通过该答案选择模型,从多个候选答案中确定出目标问题的最佳答案。

[0064] 结合上述名词简介和应用场景,对本申请实施例提供的答案选择方法进行说明。该方法可以应用于计算机设备,该计算机设备可以是图1所示场景中的服务器110或者终端设备120。下文以答案选择方法应用于服务器为例进行说明,如图2所示,该方法包括如下步骤。

[0065] 步骤101、对目标问题的多个候选答案中的每个候选答案进行编码,得到每个候选答案的初始语义向量。

[0066] 在本申请实施例中,服务器获取到目标问题后,可以基于该目标问题在数据库中进行检索,从而获取包含能够解答该目标问题的文本的多个文档。之后,服务器可以从每个文档中抽取出能够回答目标问题的片段作为一个候选答案,从而得到该目标问题的多个候



选答案。

[0067] 服务器获取到该目标问题的多个候选答案后,可以对该多个候选答案中的每个候选答案依次进行编码。通过编码,能够将文本形式的候选答案转换为向量形式,从而得到每个候选答案的初始语义向量。可选地,服务器可以将基于变换器的双向编码器表征量(Bidirectional Encoder Representations from Transformers,BERT)模型作为编码器,并将多个候选答案分别输入至该编码器,从而得到每个候选答案的初始语义向量。

[0068] 步骤102、对于多个候选答案中的每个候选答案,基于该候选答案与其他各个候选答案的相似度,对其他各个候选答案的初始语义向量进行加权处理,得到该候选答案的中间语义向量。

[0069] 在本申请实施例中,对于每两个候选答案,服务器可以基于该两个候选答案的初始语义向量的点乘,来确定该两个候选答案的相似度。并且,对于每个候选答案,服务器可以将该候选答案与其他各个候选答案的相似度作为权重,对其他各个候选答案的初始语义向量进行加权处理,从而得到该候选答案的中间语义向量。

[0070] 可选地,服务器可以采用注意力(Attention)机制,基于候选答案与其他各个候选答案的相似度,对其他各个候选答案的初始语义向量进行加权处理,以得到候选答案的更新语义向量。

[0071] 可以理解的是,对于每两个候选答案,服务器还可以采用余弦相似度、高斯距离或者相似度矩阵等方法来确定该两个候选答案的相似度。

[0072] 还可以理解的是,对于同一个目标问题,服务器从不同文档中获取到的正确答案之间是相似的(即不同的正确答案中包括的相似性的内容较多),而不正确的答案之间则是各不相同的。因此,若某个候选答案与其他候选答案之间的相似度越高,则表明该候选答案为正确答案的概率越高;若某个候选答案与其他候选答案之间的相似度越低,则表明该候选答案为正确答案的概率越低。

[0073] 基于此,服务器可以采用候选答案与其他各个候选答案的相似度,对其他各个候选答案的初始语义向量进行加权处理,得到中间语义向量。该中间语义向量可以作为候选答案的支持性证据信息。

[0074] 步骤103、对于多个候选答案中的每个候选答案,将该候选答案的初始语义向量与中间语义向量进行拼接,得到该候选答案的更新语义向量。

[0075] 对于多个候选答案中的每个候选答案,服务器可以将该候选答案的初始语义向量和中间语义向量进行拼接,以得到该候选答案的更新语义向量。可以理解的是,拼接后得到的每个候选答案的更新语义向量不仅包括该候选答案的本身信息,还包括其他候选答案对该候选答案的支持性证据信息。

[0076] 步骤104、基于多个候选答案的更新语义向量,从多个候选答案中确定出目标答案。

[0077] 服务器确定出多个候选答案的更新语义向量后,可以基于该多个候选答案的更新语义向量,来预测每个候选答案能够正确回答目标问题的概率,并将概率最高的候选答案作为该目标问题的目标答案(即最佳答案)。

[0078] 可选地,服务器可以通过分类器(也称为解码器)来预测每个候选答案能够正确回答目标问题的概率。

[0079] 在本申请实施例中,多个候选答案中的每个候选答案的更新语义向量,不仅包括该候选答案的本身信息,还包括其他候选答案对该候选答案的支持性证据信息。由于更新语义向量中的支持性的证据信息可以辅助分类器更为准确地判断每个候选答案的正确性,因此使得服务器从多个候选答案中确定出的目标答案的准确性较高。

[0080] 综上所述,本申请实施例提供了一种答案选择方法,对于每个候选答案,该方法能够基于该候选答案与其他各个候选答案的相似度,得到候选答案的中间语义向量,并将该中间语义向与初始语义向量进行拼接,得到候选答案的更新语义向量。由于该更新语义向量不仅包含了候选答案本身的信息,还包含了其他候选答案对其的支持性证据信息,因此可以确保基于多个候选答案的更新语义向量,从该多个候选答案中确定出的目标答案的准确性较高。

[0081] 图3是本申请实施例提供的另一种答案选择方法的流程图,该方法可以应用于计算机设备,该计算机设备可以是图1所示场景中的服务器110或者终端设备120。下文以答案选择方法应用于服务器为例进行说明,参见图3,该方法包括:

[0082] 步骤201、基于目标问题,获取与目标问题关联的多个文档。

[0083] 在本申请实施例中,服务器获取到目标问题后,可以基于该目标问题,在数据库中进行检索,以获取与该目标问题关联的多个文档,并将该多个文档作为该目标问题的召回文档。

[0084] 可选地,在获取与目标问题关联的多个文档时,服务器还可以对该目标问题进行同义改写,即将目标问题改写成意思相同但表达形式不同的文本。之后,服务器能够基于改写后的文本,在数据库中进行检索,以尽可能多地获取与该目标问题关联的文档。

[0085] 可选地,该数据库可以是搜索引擎。例如,在医疗问答系统中,该搜索引擎可以是涵盖各类医疗知识的搜索引擎。服务器从该搜索引擎中获取到的与目标问题关联的文档可以是期刊、论文或者是互联网上相关论坛和网站的文章。

[0086] 可以理解的是,通过检索可以有效缩小包含能够回答目标问题的答案的文档的范围,从而使得后续的答案选择过程更为高效。

[0087] 示例的,图4是本申请实施例提供的一种服务器的结构示意图,如图4所示,该服务器可以包括前序模块和答案选择模块。其中,前序模块包括检索模块11,该检索模块11能够基于目标问题在数据库10中检索,以获取与目标问题关联的多个文档。

[0088] 步骤202、针对多个文档中的每个文档,获取目标问题的一个候选答案,得到多个候选答案。

[0089] 对于获取到的多个文档中的每个文档,服务器可以从该文档中抽取出能够回答目标问题的精简片段作为一个候选答案。例如,例如,服务器可以采用多文档MRC技术从与目标问题关联的多个文档中获取多个候选答案。其中,对于该多个文档中的每个文档,服务器可以采用单文档MRC技术从该文档中获取一个候选答案。

[0090] 可以理解的是,在每个文档中选取能够与目标问题匹配的段落是一个文本匹配的过程。例如,服务器可以采用短文本-长文本匹配方法,在每个文档中选取与目标问题(即短文本)匹配且能够回答目标问题的段落(即长文本)作为目标问题的候选答案。在选取候选答案的过程中,服务器可以基于文档中各个段落与目标问题之间的语义信息及知识信息,从文档中选取多个片段作为候选答案。

[0091] 如图4所示,该前序模块还可以包括抽取模块12,对于检索模块11获取到的多个文档,该抽取模块12能够从每个文档中抽取出能够回答目标问题的精简片段作为一个候选答案。

[0092] 步骤203、对于目标问题的多个候选答案中的每个候选答案,将目标问题与该候选答案进行拼接,得到拼接文本。

[0093] 服务器可以将目标问题与该目标问题对应的多个候选答案分别进行拼接,以得到每个候选答案对应的拼接文本。在本申请实施例中,在后续对候选答案进行编码时,需要考虑目标问题和每个候选答案在语义上的匹配信息和交互信息。因此可以将目标问题和每个候选答案分别进行拼接,并将拼接得到的拼接文本输入至编码器进行编码。

[0094] 可选地,继续参考图4,为了提高后续答案选择的效率,服务器中的抽取模块113可以依据目标问题与候选答案的相关性,从上述步骤202中获取到的多个候选答案中,选择相关性最高的前n个(例如5个)候选答案进行拼接处理。例如,服务器可以基于候选答案与目标问题在语义和语法上的匹配程度,以及两者在知识层面的匹配程度来确定两者的相关性。

[0095] 若服务器采用BERT模型作为编码器,则如图5所示,在对候选答案进行编码前,服务器可以将目标问题和每个候选答案按照“[CLS]目标问题[SEP]候选答案”的方式进行拼接。其中,“[SEP]”是特殊的分隔符,BERT模型借此分隔符来区分目标问题和候选答案。“[CLS]”是一个特殊的标记,该标记可以将目标问题所包括的语义信息和候选答案所包括的语义信息汇集成整体的语义。

[0096] 步骤204、对拼接文本进行编码,得到候选答案的初始语义向量。

[0097] 服务器可以将得到的多个候选答案对应的拼接文本输入至编码器依次进行编码。通过编码,能够将文本形式的拼接文本转换为向量形式,从而得到每个候选答案的初始语义向量。

[0098] 可选地,服务器可以将BERT模型作为编码器(Encoder),并将多个拼接文本分别输入至该编码器,从而得到每个候选答案的初始语义向量。其中,编码器可以对目标问题的每个字块(Token, Tok)以及候选答案的每个字块分别进行编码。

[0099] 示例的,参考图5,假设目标问题的文本具有N个字块(也可以称为字符),则该目标问题的嵌入(Embedding)文本可以表示为:[Tok1, Tok2, ..., TokN]。该目标问题的某个候选答案的文本具有M个字块,则该候选答案的嵌入文本可以表示为:[Tok1, Tok2, ..., TokM]。其中,M和N均为大于1的正整数。

[0100] 可以理解的是,BERT模型包括多个变换器(Transformer)模型,该Transformer模型可以采用多头自我注意力(multi-head self-attention)的机制对拼接文本进行交互和编码,以获取拼接文本中目标问题和候选答案的匹配信息,并让拼接文本中的“[CLS]”注意到该拼接文本的全局语义信息。

[0101] 继续参考图4,该答案选择模块可以包括语义编码模块13。对于前序模块输出的多个候选答案中的每个候选答案,该语义编码模块13能够对每个候选答案对应的拼接文本进行编码,以得到候选答案的初始语义向量。

[0102] 步骤205、对于每个候选答案,基于该候选答案的初始语义向量与其他每个候选答案的初始语义向量的点乘,确定该候选答案与其他每个候选答案的相似度。

[0103] 在本申请实施例中,对于多个候选答案中的每个候选答案,服务器可以分别计算该候选答案的初始语义向量与其他每个候选答案的初始语义向量的点乘。其中,第*i*个候选答案的初始语义向量 $r^{Ai}$ 与第*j*个候选答案的语义向量 $r^{Aj}$ 的点乘 $s_{i,j}$ 的计算公式可以表示为:

$$[0104] \quad s_{i,j} = r^{Ai} \cdot r^{Aj};$$

[0105] 其中, $T$ 转置符号,且*i*≠*j*,*i*和*j*均为不大于*n*的正整数,*n*为目标问题的多个候选答案的总数,且*n*为大于1的整数。

[0106] 作为第一种可能的示例,对于多个候选答案中的每个候选答案,服务器可以直接将该候选答案的初始语义向量与其他每个候选答案的初始语义向量的点乘,作为该候选答案与其他每个候选答案的相似度。也即是,第*i*个候选答案与第*j*个候选答案的相似度为 $s_{i,j}$ 。

[0107] 作为第二种可能的示例,对于多个候选答案中的每个候选答案,服务器可以对该候选答案的初始语义向量与其他每个候选答案的初始语义向量的点乘进行归一化处理,并将归一化处理后的结果作为该候选答案与其他每个候选答案的相似度。其中,服务器对第*i*个候选答案的初始语义向量 $r^{Ai}$ 与第*j*个候选答案的语义向量 $r^{Aj}$ 的点乘 $s_{i,j}$ 进行归一化处理后的结果(即第*i*个候选答案与第*j*个候选答案的相似度 $\alpha_{i,j}$ )可以表示为:

$$[0108] \quad \alpha_{i,j} = \frac{e^{s_{i,j}}}{\sum_{k=1}^n e^{s_{i,k}}};$$

[0109] 其中, $s_{i,k}$ 为第*i*个候选答案的初始语义向量与第*k*个候选答案的初始语义向量的点乘, $e$ 为自然常数(即自然对数的底数),*k*为不大于*n*的正整数。

[0110] 可以理解的是,在上述两种示例中,服务器可以将每个候选答案与该候选答案自身的相似度设置为0,即服务器无需计算候选答案和它自身的相似度。因此,对于第一种示例,当*i*=*j*时, $s_{i,j}=0$ ,对于第二种示例,当*i*=*j*时, $\alpha_{i,j}=0$ 。

[0111] 还可以理解的是,通过对任意两个候选答案的初始语义向量的点乘做归一化处理,可以将点乘的结果(即相似度)映射到0~1的范围之内。也即是,通过归一化处理,能够使得任意两个候选答案的相似度位于0~1的范围之内,从而使得后续的处理更为高效和简便。

[0112] 步骤206、对于多个候选答案中的每个候选答案,基于该候选答案与其他各个候选答案的相似度,对其他各个候选答案的初始语义向量进行加权处理,得到该候选答案的中间语义向量。

[0113] 在本申请实施例中,对于多个候选答案中的第*i*个候选答案,服务器能够以该候选答案与其他各个候选答案的相似度为权重,对其他各个答案的初始语义向量进行加权求和,得到该第*i*个候选答案的中间语义向量 $\tilde{r}^{Ai}$ 。其中,该第*i*个候选答案的中间语义向量 $\tilde{r}^{Ai}$ 的计算公式可以表示为:

$$[0114] \quad \tilde{r}^{Ai} = \sum_{j=1}^n \alpha_{i,j} \cdot r^{Aj};$$

[0115] 基于上述步骤205可知,当*j*=*i*时, $\alpha_{i,j}=0$ 。因此,对于多个候选答案中的第*i*个候选答案,服务器不需要以该第*i*个候选答案与自身的相似度为权重,对第*i*个候选答案的初

始语义向量进行加权。或者可以理解为,在上述第*i*个候选答案的中间语义向量 $\tilde{r}^{A_i}$ 的计算公式中, $j \neq i$ 。

[0116] 可选地,如图6所示,服务器可以采用Attention机制实现每个候选答案的初始语义向量转换为中间语义向量的过程。参见图6,Attention机制可以基于*n*个候选答案的初始语义向量,确定出该*n*个候选答案中每个候选答案的中间语义向量。该中间语义向量可以作为候选答案的支持性证据信息。

[0117] 步骤207、对于该多个候选答案中的每个候选答案,将该候选答案的初始语义向量和中间语义向量进行拼接,得到该候选答案的更新语义向量。

[0118] 继续参考图6,对于*n*个候选答案中的每个候选答案,服务器可以将每个候选答案的初始语义向量和中间语义向量进行拼接(Fuse),以得到该候选答案的更新语义向量。其中,第*i*个候选答案的初始语义向量 $r^{A_i}$ 和中间语义向量 $\tilde{r}^{A_i}$ 拼接后得到的更新语义向量 $g_i^v$ 可以表示为:

$$[0119] \quad g_i^v = [r^{A_i}; \tilde{r}^{A_i}];$$

[0120] 可以理解的是,拼接后得到的每个候选答案的更新语义向量不仅包括该候选答案的本身信息,还包括其他候选答案对该候选答案的支持性证据信息。

[0121] 可选地,为了更好地实现候选答案之间的语义交互,对于每个候选答案,服务器可以将该候选答案的初始语义向量、该候选答案的中间语义向量和该候选答案的乘积向量进行拼接,以得到该候选答案的更新语义向量。其中,候选答案的乘积向量为该候选答案的初始语义向量与中间语义向量的向量积。相应的,该多个候选答案中的第*i*个候选答案的更新语义向量 $g_i^v$ 可以表示为:

$$[0122] \quad g_i^v = [r^{A_i}; \tilde{r}^{A_i}; r^{A_i} \times \tilde{r}^{A_i}];$$

[0123] 示例的,如图4所示,该服务器的答案选择模块还可以包括多答案比较模块14,该多答案比较模块14可以实现上述步骤205至步骤207所示的方法。

[0124] 步骤208、将多个候选答案的更新语义向量输入至答案选择模型。

[0125] 服务器确定出多个候选答案的更新语义向量后,可以将该多个候选答案的更新语义向量输入至答案选择模型。其中,该答案选择模型可以为二分类模型,且该答案选择模型也可以称为解码器。

[0126] 在本申请实施例中,该答案选择模型可以基于每个候选答案的更新语义向量,判断每个候选答案是否为目标答案的正确答案,并预测出每个候选答案为正确答案的概率。

[0127] 步骤209、基于答案选择模型的输出结果,从多个候选答案中确定出目标答案。

[0128] 在本申请实施例中,对于多个候选答案中的每个候选答案,答案选择模型的输出结果可以包括该候选答案的预测概率和预测标签。其中,对于预测概率大于或等于概率阈值的候选答案,答案选择模型输出的预测标签可以为第一标签,对于预测概率小于该概率阈值的候选答案,答案选择模型输出的预测标签可以为第二标签。

[0129] 其中,该概率阈值可以是服务器中预先配置的概率值,其可以为大于或等于0.5的数值。该第一标签和第二标签均可以为数值,例如第一标签可以为1,第二标签可以为0。

[0130] 服务器可以将预测标签为第一标签(例如1)的候选答案确定为目标问题的正确答

案,并将预测标签为第二标签(例如0)的候选答案确定为目标问题的错误答案。并且,服务器可以将预测标签为第一标签的至少一个正确答案中,预测概率最高的正确答案确定为目标问题的目标答案。

[0131] 参见图6,服务器可以将拼接得到的n个候选答案的更新语义向量分别输入至解码器进行解码(也可以称为预测),解码器进而可以输出每个候选答案的预测概率,或者输出每个候选答案的预测概率和预测标签。例如,假设概率阈值为0.5,第一标签为1,第二标签为0。对于某个候选答案,若答案选择模型输出的预测概率为0.6,则答案选择模型还可以输出该候选答案的预测标签:1。对于某个候选答案,若答案选择模型输出的预测概率为0.4,则答案选择模型还可以输出该候选答案的预测标签:0。

[0132] 示例的,如图4所示,该答案选择模块还可以包括结果分类模块15,该结果分类模块15可以实现上述步骤208至步骤209所示的方法。也即是,确定候选答案正确性的过程是一个二分类任务,通过二分类任务,可以将多个候选答案分为正确答案和错误答案。

[0133] 表1示出了本申请实施例提供的答案选择模型与基础参考(baseline)模型在判断候选答案为正确答案的过程中,在准确率(Precision),召回率(Recall)以及综合评价指标F值(F-Measure)方面的比较结果。如表1所示,对于候选答案中的正确答案,本申请实施例提供的答案选择模型的准确率,召回率以及综合评价指标F值均高于基础参考模型。

[0134] 表1

方法	正确答案 (%)		
	Precision	Recall	F
基础参考模型	74.7645	57.973	66.232
本申请实施例	84.128	58.052	68.577

[0136] 可以理解的是,在本申请实施例中,输入至答案选择模型的更新语义向量,不仅包括候选答案本身的信息,还包括其他各个候选答案对该候选答案的支持性证据信息。由此,答案选择模型可以基于该支持性证据信息,更为准确地判断每个候选答案的正确性。参见表1,采用本申请实施例提供的答案选择方法,可以在确保正确答案的召回率的基础上,提高答案选择模型判断候选答案为正确答案的准确率。由表1提供的数据可以看出,本申请实施例提供的答案选择模型判断候选答案为正确答案的准确率,相比于基础参考模型提高了10%。

[0137] 基于上述分析可知,本申请实施例提供的答案选择方法能够通过候选答案之间的支持性证据信息引入知识验证信息,从而有效提高从候选答案中选择出正确答案的准确率,以使得答案选择模型确定出的候选答案的正确性更为准确。

[0138] 可选地,本申请实施例提供的答案选择方法可以应用于医疗领域。例如,可以作为医典搜索中的离线数据收集模块的一部分。并且,通过本申请实施例提供的答案选择方法选择出的目标答案(即最佳答案),可以应用于医疗问答系统的线上直接问答模块,或者应用于医疗问答系统的问答式摘要等模块。本申请实施例提供的答案选择方法还可以应用于其他领域,例如金融领域的金融知识问答,农业领域的农业百科知识问答等,本申请在此并不做限定。

[0139] 可以理解的是,本申请实施例提供的答案选择方法的步骤的先后顺序可以进行适

当调整,步骤也可以根据情况进行相应增减。例如,上述步骤203可以根据情况删除,也即是,服务器可以只对候选答案进行编码。任何熟悉本技术领域的技术人员在本申请揭露的技术范围内,可轻易想到变化的方法,都应涵盖在本申请的保护范围之内,因此不再赘述。

[0140] 综上所述,本申请实施例提供了一种答案选择方法,对于每个候选答案,能够基于该候选答案与其他各个候选答案的相似度,得到候选答案的中间语义向量,并将该中间语义向量与初始语义向量进行拼接,得到候选答案的更新语义向量。由于该更新语义向量不仅包含了候选答案本身的信息,还包含了其他候选答案对其的支持性证据信息,因此可以确保基于多个候选答案的更新语义向量,从该多个候选答案中确定出的目标答案的准确性较高。

[0141] 本申请实施例还提供了一种答案选择模型的训练方法,该方法可以用于训练上述实施例中所述的答案选择模型。并且,该答案选择模型的训练方法可以应用于计算机设备,该计算机设备可以是图1所示场景中的服务器110。参见图7,该答案选择模型的训练过程包括如下步骤。

[0142] 步骤301、获取问题样本的多个答案样本,以及每个答案样本的标签。

[0143] 服务器可以根据答案选择模型的具体应用场景,来获取与场景相对应的多个问题样本,以及该多个问题样本中每个问题样本的多个答案样本。其中,每个答案样本具有人工标注的标签,该标签可以用于指示答案样本是否为问题样本的正确答案。若某个答案样本为问题样本的正确答案,则该答案样本的标签为第一标签(例如1),若某个答案样本为问题样本的错误答案,则该答案样本的标签为第二标签(例如0)。

[0144] 可以理解的是,该多个问题样本以及多个问题样本中的每个问题样本的答案样本可以是服务器中预先存储的,也可以是终端设备发送给服务器的。

[0145] 步骤302、对多个答案样本中的每个答案样本进行编码,得到每个答案样本的初始语义向量。

[0146] 步骤303、对于多个答案样本中的每个答案样本,基于该答案样本与其他各个答案样本的相似度,对其他各个答案样本的初始语义向量进行加权处理,得到该答案样本的中间语义向量。

[0147] 步骤304、对于多个答案样本中的每个答案样本,将该答案样本的初始语义向量和语义向量进行拼接,得到该答案样本的更新语义向量。

[0148] 上述步骤302至步骤304的实现过程可以参考步骤101至步骤103的实现过程,以及步骤203至步骤207的实现过程,本申请实施例在此不再赘述。

[0149] 步骤305、基于多个答案样本的更新语义向量以及多个答案样本的标签,训练答案选择模型。

[0150] 在本申请实施例中,服务器中预先存储有初始模型。服务器可以采用多个答案样本的更新语义向量以及多个答案样本的标签,对该初始模型进行训练,以得到答案选择模型。或者,服务器可以采用多个答案样本的更新语义向量以及多个答案样本的标签,对已训练好的答案选择模型进行更新。

[0151] 可选地,服务器可以采用二分类交叉熵函数作为该答案选择模型的损失函数,并可以采用端到端的训练方式对答案选择模型进行训练。其中,该损失函数L可以表示为:

[0152]  $L = -(y \cdot \log p + (1-y) \cdot \log(1-p))$ ;

[0153] 其中, $y$ 表示答案样本的标签,例如, $y=1$ 表示该答案样本是问题样本的正确答案, $y=0$ 表示该答案样本是问题样本的错误答案。 $p$ 表示答案选择模型输出的答案样本的预测概率,即答案样本为正确答案的概率。

[0154] 在训练答案选择模型的过程中,服务器可以基于答案选择模型输出的答案样本的预测概率与该答案样本的标签之间的误差,对该答案选择模型的相关参数进行调节。当损失函数 $L$ 收敛,或者该答案选择模型的训练效果达到预期效果时,服务器可以停止对该答案选择模型的训练。

[0155] 可以理解的是,由于用于训练的答案样本的更新语义向量不仅包含答案样本本身的信息,还包括其他各个答案样本对该答案样本的支持性证据信息。因此在训练过程中,答案选择模型能够较好地学习到每个候选答案的语义信息,以及该答案样本与问题样本的相关性。由此,可以确保答案选择模型的训练效果较好。

[0156] 示例的,假设该答案选择模型应用于医疗场景,则可以采用相关医疗问题作为问题样本,并采用针对该问题样本的相关医疗知识作为答案样本。参考表2,对于问题样本“切除子宫一侧的卵巢对身体有影响吗”,该问题样本对应的4个真实标签为1的答案样本在语义和语法上,与该问题样本的匹配程度并不高。

[0157] 相应的,若采用基础参考模型对该4个答案样本的正确性进行预测,则如表2所示,该基础参考模型输出的该4个答案样本的标签均为0。也即是,基础参考模型将这4个答案样本都判断成了错误答案,基础参考模型输出的标签与该4个答案样本的真实标签不一致。

[0158] 而采用本申请实施例提供的答案选择模型后,输入至答案选择模型的答案样本的更新语义向量不仅包括该答案样本自身的信息,还包括其他各个答案样本的支持性证据信息。参见表2,该支持性证据信息可以为各个答案样本中加粗的文本。由于答案选择模型基于每个答案样本的支持性证据信息,可以更为准确地判断答案样本的正确性,因此对于表2示出的4个答案样本,答案选择模型输出的标签均为1。也即是,该答案选择模型输出的标签与该4个答案样本的真实标签一致。

[0159] 表2

问题样本	答案样本	基础参考模型	答案选择模型	人工标注
[0160] 切除子宫一侧的卵巢对身体有影响吗	<b>切除一侧卵巢, 只要另一侧卵巢形态、功能是正常的, 则对身体影响基本非常小</b>	0	1	1
	<b>切除卵巢对绝经后的女性影响不大, 但是绝经前的女性切除双侧卵巢, 就会出现相应的更年期症状。</b>	0	1	1
	<b>切除一侧卵巢, 女性激素分泌基</b>	0	1	1



	<p><b>本保持正常的水平。</b>如果两侧都切除的话，雌性激素和孕激素会分泌停止，患者会出现头疼、全身性的神经紧张、分泌物减少、阴道干涩等情况</p>			
<p>[0161]</p>	<p>正常的<b>卵巢</b>是一对，可以分泌性激素，能排卵。<b>切除一侧</b>后，可能容易引起卵巢功能早衰。加入不需要再生育，影响会相对小些。</p>	<p>0</p>	<p>1</p>	<p>1</p>

[0162] 综上所述，本申请实施例提供了一种答案选择模型的训练方法，对于每个答案样本，能够基于该答案样本与其他各个答案样本的相似度，得到答案样本的中间语义向量，并将该中间语义向与初始语义向量进行拼接，得到答案样本的更新语义向量。由于该更新语义向量不仅包含了答案样本本身的信息，还包含了其他答案样本对其的支持性证据信息。因此在训练过程中，答案选择模型能够较好地学习到每个候选答案的语义信息，以及该答案样本与问题样本的相关性。由此，可以确保答案选择模型能够较为准确地对答案样本准确性进行判断，从而使得该答案选择模型的训练效果较好。

[0163] 图8是本申请实施例提供的一种答案选择装置的结构示意图，该答案选择装置可以应用于计算机设备中，该计算机设备可以是图1所示场景中的服务器110。如图8所示，该装置包括：

[0164] 编码模块401，用于对目标问题的多个候选答案中的每个候选答案进行编码，得到每个候选答案的初始语义向量。

[0165] 其中，该编码模块401可以为图4所示的语义编码模块13。

[0166] 加权模块402，用于对于多个候选答案中的每个候选答案，基于候选答案与其他各个候选答案的相似度，对其他各个候选答案的初始语义向量进行加权处理，得到候选答案的中间语义向量；

[0167] 拼接模块403，用于对于多个候选答案中的每个候选答案，将候选答案的初始语义向量与中间语义向量拼接进行拼接，得到候选答案的更新语义向量。

[0168] 其中，上述加权模块402和拼接模块403可以均属于图4所示的多答案比较模块14。

[0169] 确定模块404，用于基于多个候选答案的更新语义向量，从多个候选答案中确定出目标答案。

[0170] 其中，该确定模块404可以为图4所示的结果分类模块15。

[0171] 可选地，该拼接模块403用于对于多个候选答案中的每个候选答案，将候选答案的初始语义向量、候选答案的中间语义向量和候选答案的乘积向量进行拼接，得到候选答案的更新语义向量；其中，乘积向量为候选答案的初始语义向量与候选答案的中间语义向量的向量积。

[0172] 可选地，该加权模块402用于对于多个候选答案中的每个候选答案，基于候选答案

的初始语义向量与其他每个候选答案的初始语义向量的点乘,确定候选答案与其他每个候选答案的相似度。

[0173] 可选地,该多个候选答案中,第*i*个候选答案与第*j*个候选答案的相似度 $\alpha_{i,j}$ 满足:

$$[0174] \quad \alpha_{i,j} = \frac{e^{s_{i,j}}}{\sum_{k=1}^n e^{s_{i,k}}};$$

[0175] 其中, $s_{i,j}$ 为第*i*个候选答案的初始语义向量与第*j*个候选答案的初始语义向量的点乘, $s_{i,k}$ 为第*i*个候选答案的初始语义向量与第*k*个候选答案的初始语义向量的点乘, $e$ 为自然常数, $n$ 为多个候选答案的总数, $i,j$ 和*k*均为不大于*n*的正整数。

[0176] 可选地,该编码模块401用于对目标问题的多个候选答案中的每个候选答案,将目标问题与候选答案进行拼接,得到拼接文本;对拼接文本进行编码,得到候选答案的初始语义向量。

[0177] 可选地,如图9所示,该答案选择装置还可以包括第一获取模块405,该第一获取模块405用于基于目标问题,获取与目标问题关联的多个文档;针对多个文档中的每个文档,获取目标问题的一个候选答案,得到多个候选答案。

[0178] 可选地,该确定模块404用于将多个候选答案的更新语义向量输入至答案选择模型;基于答案选择模型的输出结果,从多个候选答案中确定出目标答案。

[0179] 可选地,继续参考图9,该答案选择装置还可以包括第二获取模块406,该第二获取模块406,用于获取问题样本的多个答案样本,以及每个答案样本的标签,该标签用于指示答案样本是否为问题样本的正确答案。

[0180] 该编码模块401,还用于对多个答案样本中的每个答案样本进行编码,得到每个答案样本的初始语义向量。

[0181] 该加权模块402,还用于对于多个答案样本中的每个答案样本,基于答案样本与其他各个答案样本的相似度,对其他各个答案样本的初始语义向量进行加权处理,得到答案样本的中间语义向量。

[0182] 该拼接模块403,用于对于多个答案样本中的每个答案样本,将答案样本的初始语义向量和中间语义向量进行拼接,得到答案样本的更新语义向量。

[0183] 继续参考图9,该答案选择装置还可以包括训练模块407,该训练模块407用于基于多个答案样本的更新语义向量以及多个答案样本的标签,训练答案选择模型。

[0184] 综上所述,本申请实施例提供了一种答案选择装置,对于每个候选答案,该装置能够基于该候选答案与其他各个候选答案的相似度,得到候选答案的中间语义向量,并将该中间语义向与初始语义向量进行拼接,得到候选答案的更新语义向量。由于该更新语义向量不仅包含了候选答案本身的信息,还包含了其他候选答案对其的支持性证据信息,因此可以确保基于多个候选答案的更新语义向量,从该多个候选答案中确定出的目标答案的准确性较高。

[0185] 图10是本申请实施例提供的一种答案选择模型的训练装置的结构框图,该答案选择装置可以应用于计算机设备中,该计算机设备可以是图1所示场景中的服务器110。如图10所示,该装置包括:

[0186] 获取模块501,用于获取问题样本的多个答案样本,以及每个答案样本的标签,标签用于指示答案样本是否为问题样本的正确答案。

[0187] 编码模块502,用于对多个答案样本中的每个答案样本进行编码,得到每个答案样本的初始语义向量。

[0188] 加权模块503,用于对于多个答案样本中的每个答案样本,基于答案样本与其他各个答案样本的相似度,对其他各个答案样本的初始语义向量进行加权处理,得到答案样本的中间语义向量。

[0189] 拼接模块504,用于对答案样本的初始语义向量和答案样本的中间语义向量进行拼接,得到答案样本的更新语义向量。

[0190] 训练模块505,用于基于多个答案样本的更新语义向量以及多个答案样本的标签,训练答案选择模型。

[0191] 综上所述,本申请实施例提供了一种答案选择模型的训练装置,对于每个候选答案,该装置能够基于该答案样本与其他各个答案样本的相似度,得到答案样本的中间语义向量,并将该中间语义向与初始语义向量进行拼接,得到答案样本的更新语义向量。由于该更新语义向量不仅包含了答案样本本身的信息,还包含了其他答案样本对其的支持性证据信息,因此在训练过程中,答案选择模型能够较好地学习到每个候选答案的语义信息,以及该答案样本与问题样本的相关性。由此,可以确保答案选择模型能够较为准确地对答案样本准确性进行判断,从而使得该答案选择模型的训练效果较好。

[0192] 可以理解的是:上述实施例提供的答案选择装置和答案选择模型的训练装置,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,即将设备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。

[0193] 另外,上述实施例提供的答案选择装置和答案选择方法实施例属于同一构思,答案选择模型的训练装置与答案选择模型的训练方法实施例属于同一构思,其具体实现过程详见方法实施例,这里不再赘述。

[0194] 图11是本申请实施例提供的一计算机设备的结构框图,该计算机设备可以是图1所示场景中的服务器110,下文以该计算机设备为服务器110进行说明。

[0195] 具体来讲:服务器110包括中央处理单元(Central Processing Unit,CPU)111、包括随机存取存储器(Random Access Memory,RAM)1121和只读存储器(Read Only Memory,ROM)1122的系统存储器112,以及连接系统存储器112和中央处理单元111的系统总线113。服务器110还包括用于存储操作系统1141、应用程序1142和其他程序模块1143的大容量存储设备114。

[0196] 大容量存储设备114通过连接到系统总线113的大容量存储控制器(未示出)连接到中央处理单元111。大容量存储设备114及其相关联的计算机可读介质为服务器110提供非易失性存储。也就是说,大容量存储设备114可以包括诸如硬盘或者紧凑型光盘只读存储器(Compact Disc Read Only Memory,CD-ROM)驱动器之类的计算机可读介质(未示出)。

[0197] 不失一般性,计算机可读介质可以包括计算机存储介质和通信介质。计算机存储介质包括以用于存储诸如计算机可读指令、数据结构、程序模块或其他数据等信息的任何方法或技术实现的易失性和非易失性、可移动和不可移动介质。计算机存储介质包括RAM、ROM、可擦除可编程只读存储器(Erasable Programmable Read Only Memory,EPR0M)、带电可擦可编程只读存储器(Electrically Erasable Programmable Read Only Memory,

EEPROM)、闪存或其他固态存储其技术,CD-ROM、数字通用光盘(Digital Versatile Disc, DVD)或其他光学存储、磁带盒、磁带、磁盘存储或其他磁性存储设备。当然,本领域技术人员可知计算机存储介质不局限于上述几种。上述的系统存储器112和大容量存储设备114可以统称为存储器。

[0198] 根据本申请的各种实施例,服务器110还可以通过诸如因特网等网络连接到网络上的远程计算机运行。也即服务器110可以通过连接在系统总线113上的网络接口单元115连接到网络116,或者说,也可以使用网络接口单元115来连接到其他类型的网络或远程计算机系统(未示出)。

[0199] 上述存储器还包括一个或者一个以上的程序,一个或者一个以上程序存储于存储器中,被配置由CPU执行。

[0200] 本申请实施例还提供了一种计算机设备,该计算机设备包括处理器和存储器,该存储器中存储有至少一条指令、至少一段程序、代码集或指令集,至少一条指令、至少一段程序、代码集或指令集由处理器加载并执行以实现上述各方法实施例提供的答案选择方法,或答案选择模型的训练方法。

[0201] 本申请实施例还提供了一种计算机可读存储介质,该计算机可读存储介质上存储有至少一条指令、至少一段程序、代码集或指令集,至少一条指令、至少一段程序、代码集或指令集由处理器加载并执行,以实现上述各方法实施例提供的答案选择方法,或答案选择模型的训练方法。

[0202] 本申请实施例还提供了一种计算机程序产品或计算机程序,该计算机程序产品或计算机程序包括计算机指令,该计算机指令存储在计算机可读存储介质中。计算机设备的处理器从计算机可读存储介质读取该计算机指令,处理器执行该计算机指令,使得该计算机设备执行上述实施例中任一所述的答案选择方法,或答案选择模型的训练方法。

[0203] 可选地,该计算机可读存储介质可以包括:只读存储器(ROM,Read Only Memory)、随机存取记忆体(RAM,Random Access Memory)、固态硬盘(SSD,Solid State Drives)或光盘等。其中,随机存取记忆体可以包括电阻式随机存取记忆体(ReRAM,Resistance Random Access Memory)和动态随机存取存储器(DRAM,Dynamic Random Access Memory)。上述本申请实施例序号仅仅为了描述,不代表实施例的优劣。

[0204] 可以理解的是,本申请中术语“至少一个”是指一个或多个,“多个”的含义是指两个或两个以上。

[0205] 还可以理解的是,本申请所涉及的信息和数据(包括但不限于用于训练模型的数据,以及用于提供候选答案的数据),均为经用户授权或者经过各方充分授权的,且相关数据的收集、使用和处理需要遵守相关国家和地区的相关法律法规和标准。例如,本申请中涉及到的与目标问题关联的文档,以及答案样本都是在充分授权的情况下获取的。

[0206] 本领域普通技术人员可以理解实现上述实施例的全部或部分步骤可以通过硬件来完成,也可以通过程序来指令相关的硬件完成,所述的程序可以存储于一种计算机可读存储介质中,上述提到的存储介质可以是只读存储器,磁盘或光盘等。本申请提供的上述实施例可应用于云技术、人工智能、智慧交通等各种场景,在此不做限定。

[0207] 以上所述仅为本申请的可选实施例,并不用以限制本申请,凡在本申请的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本申请的保护范围之内。

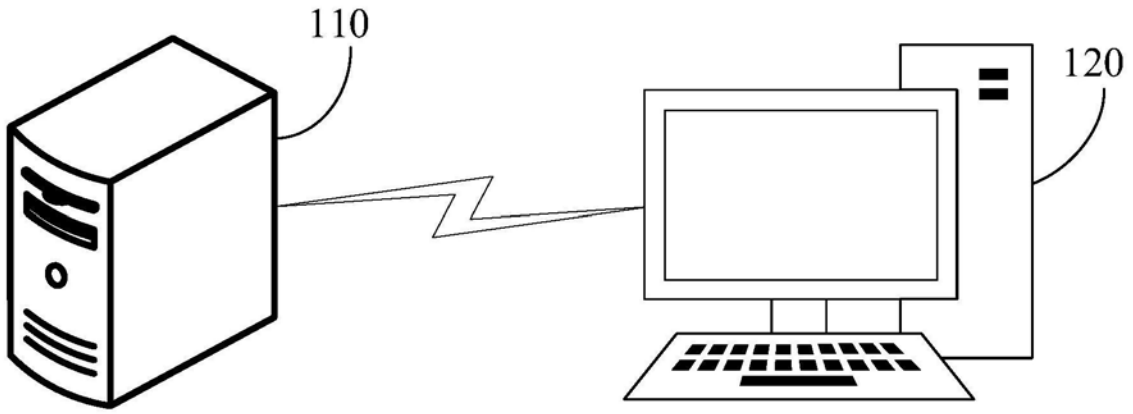


图1

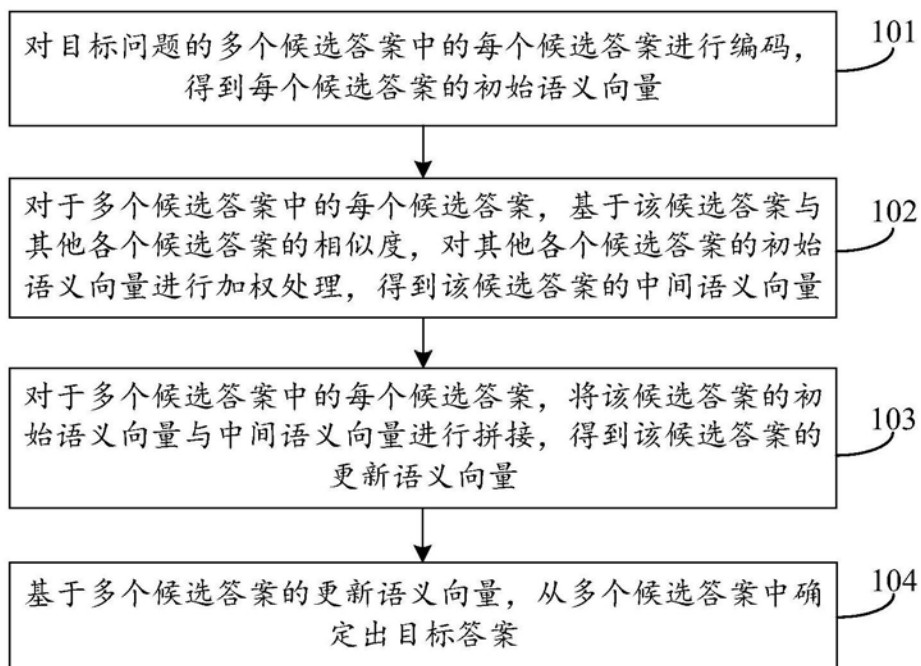


图2

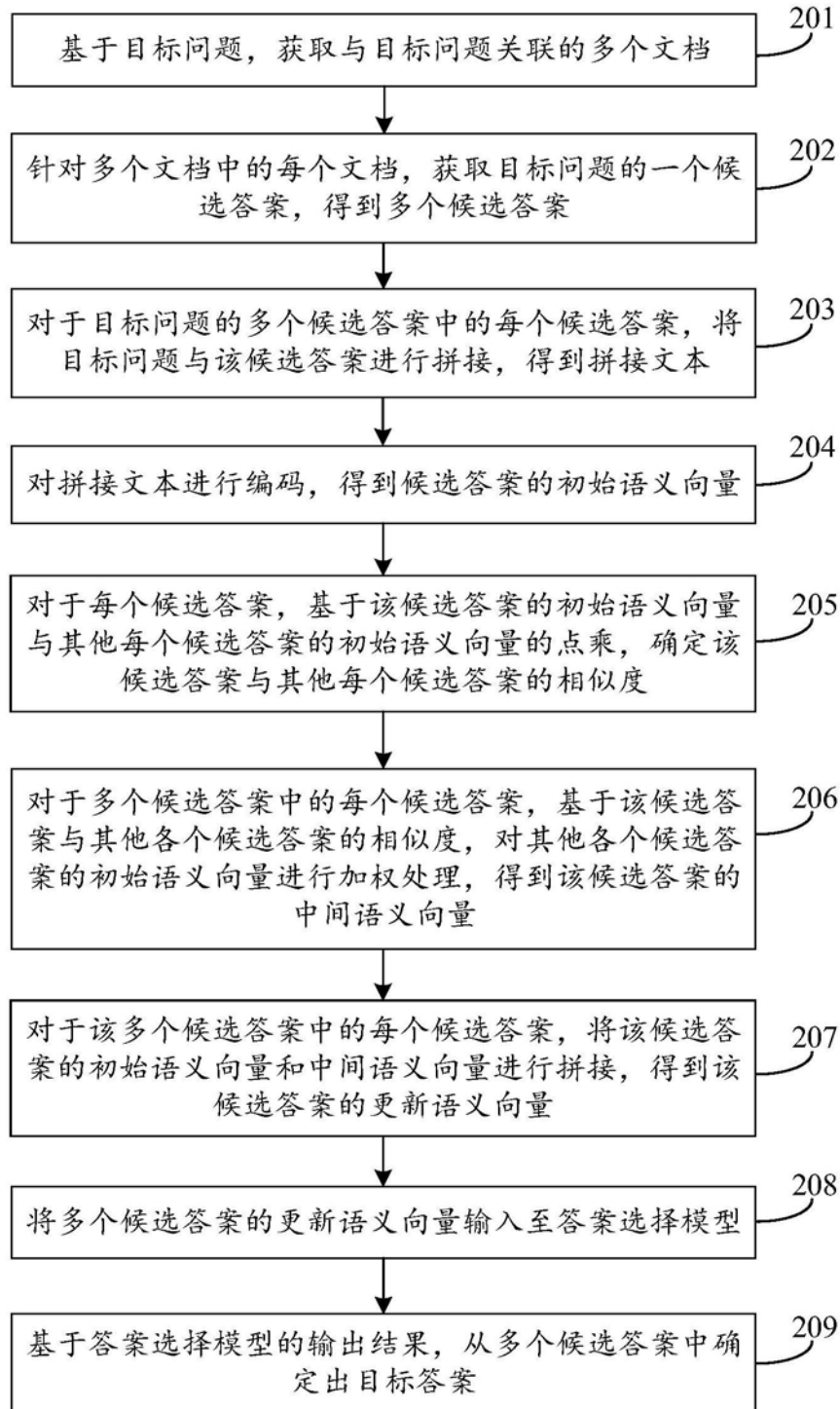


图3

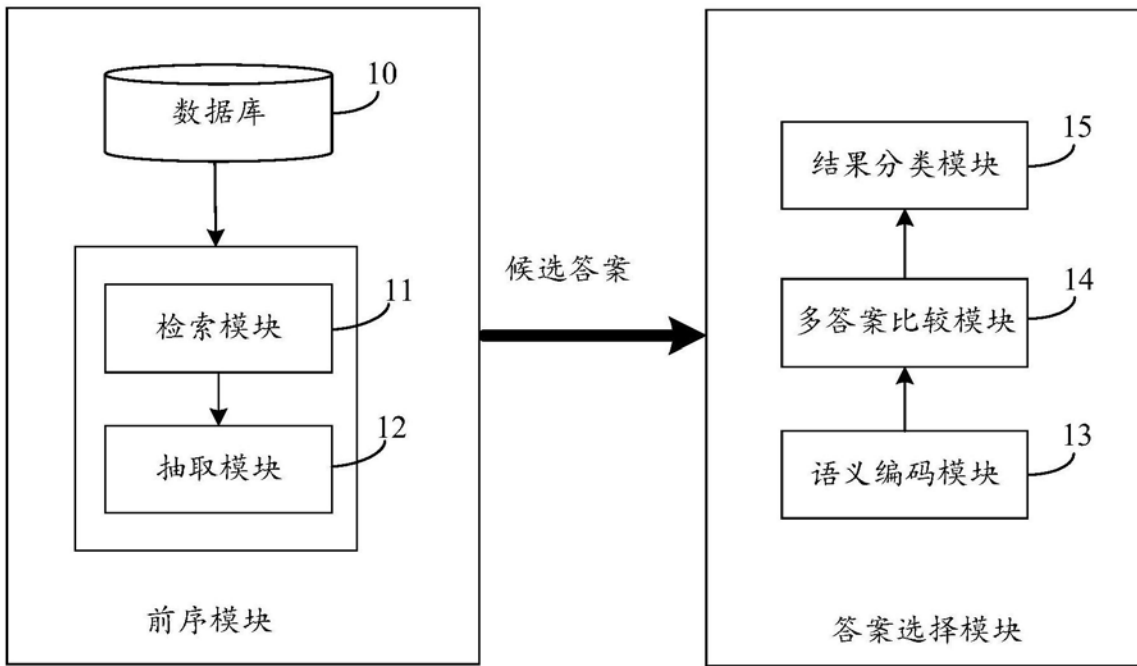


图4

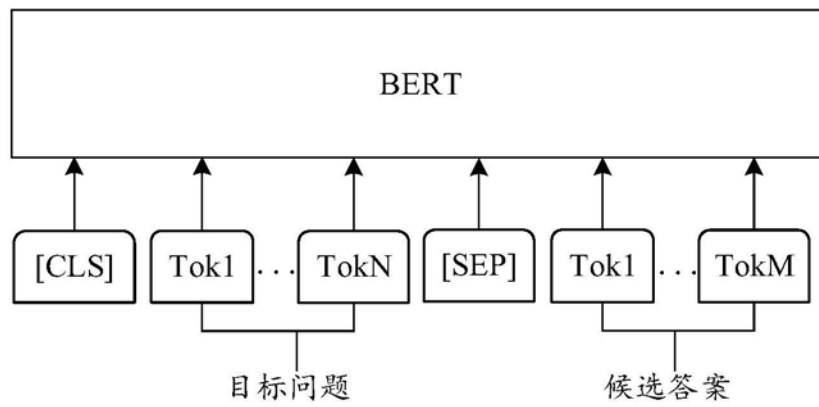


图5

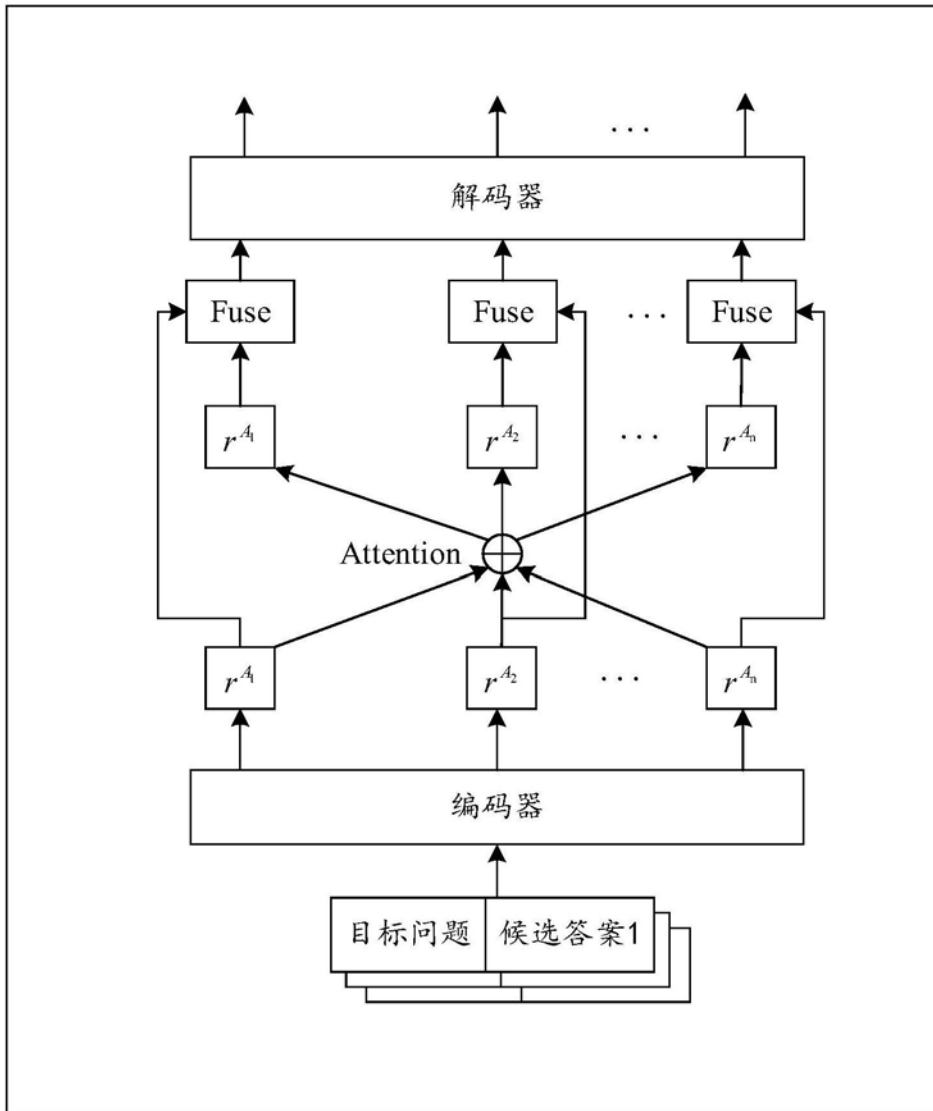


图6



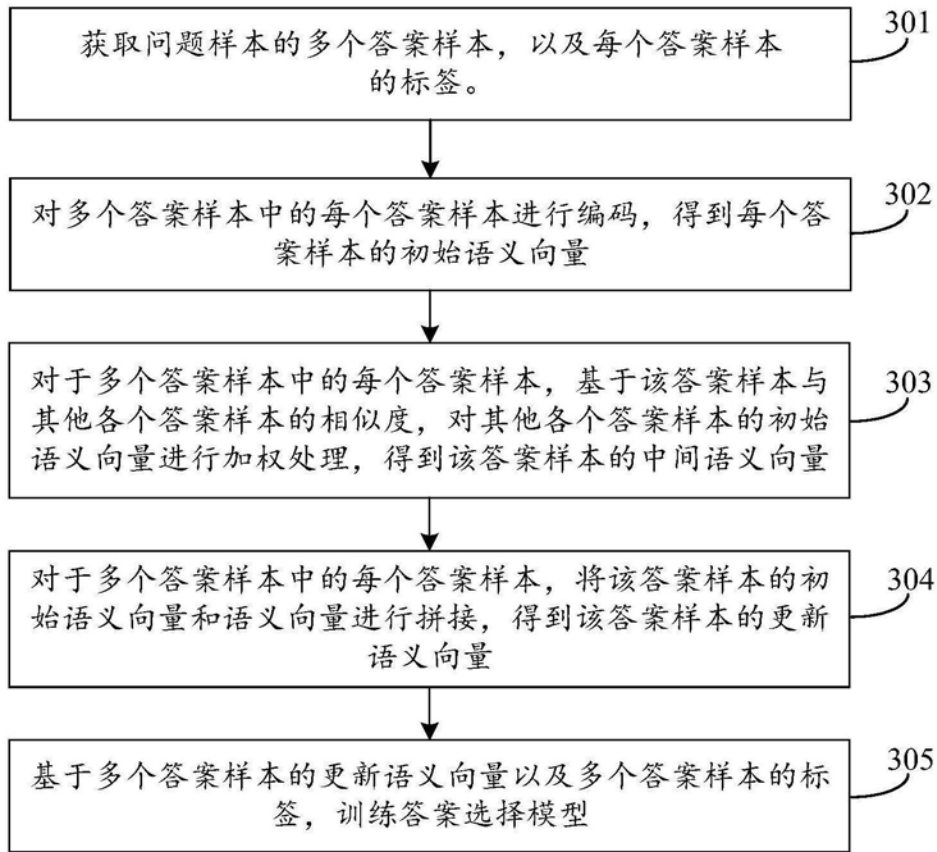


图7

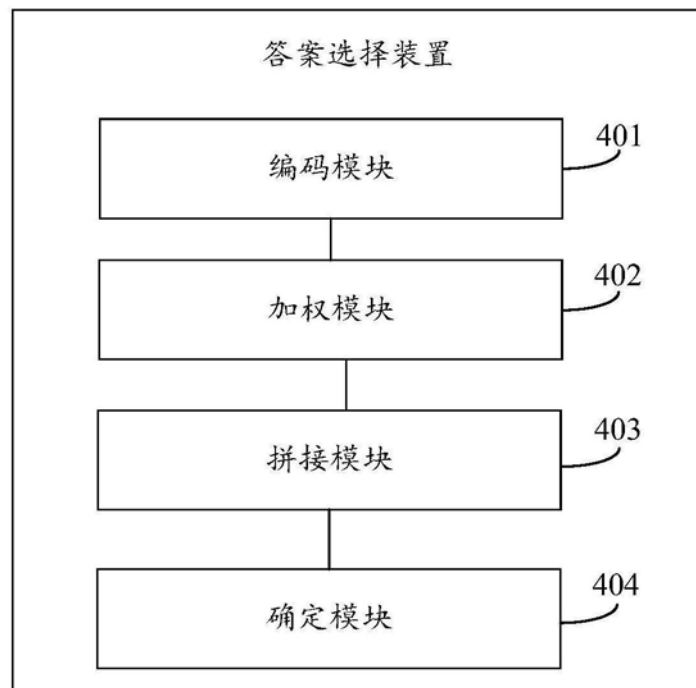


图8

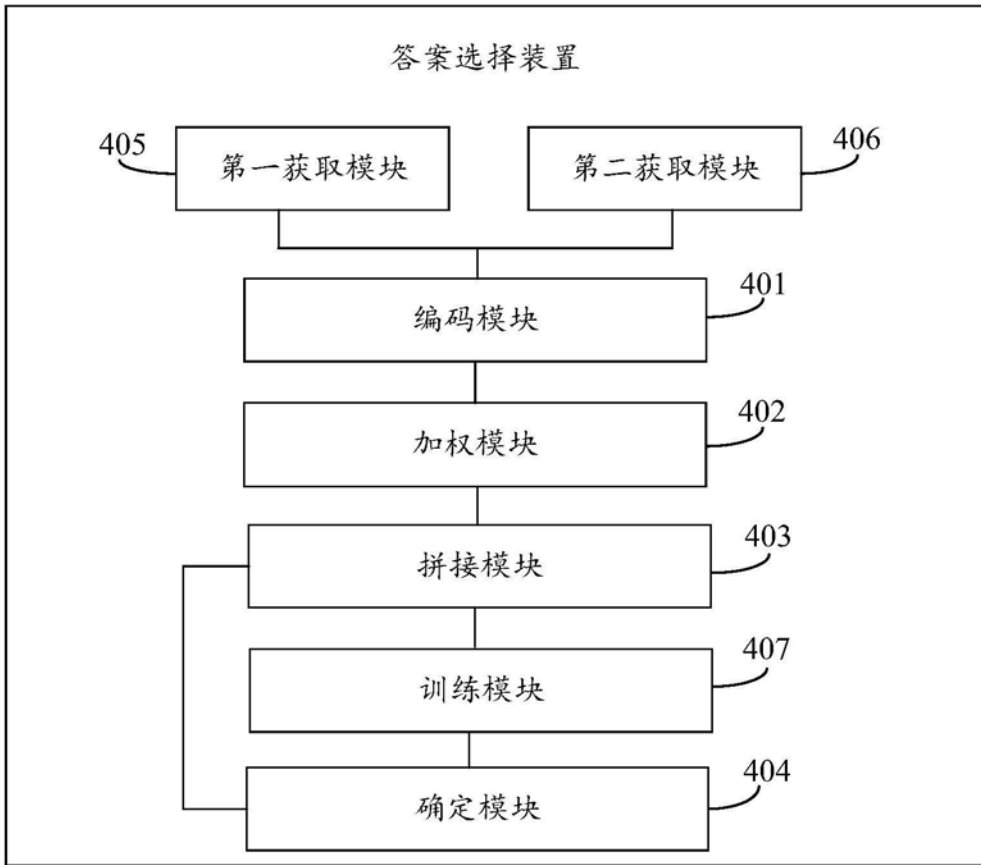


图9

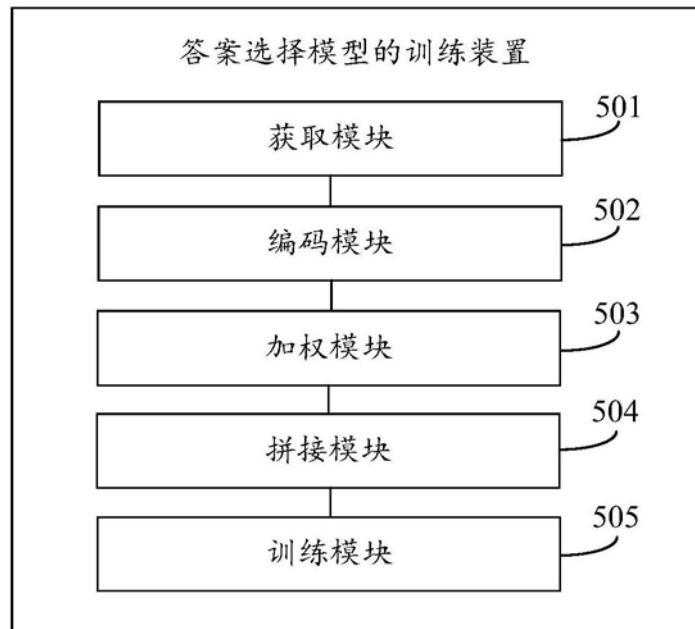


图10

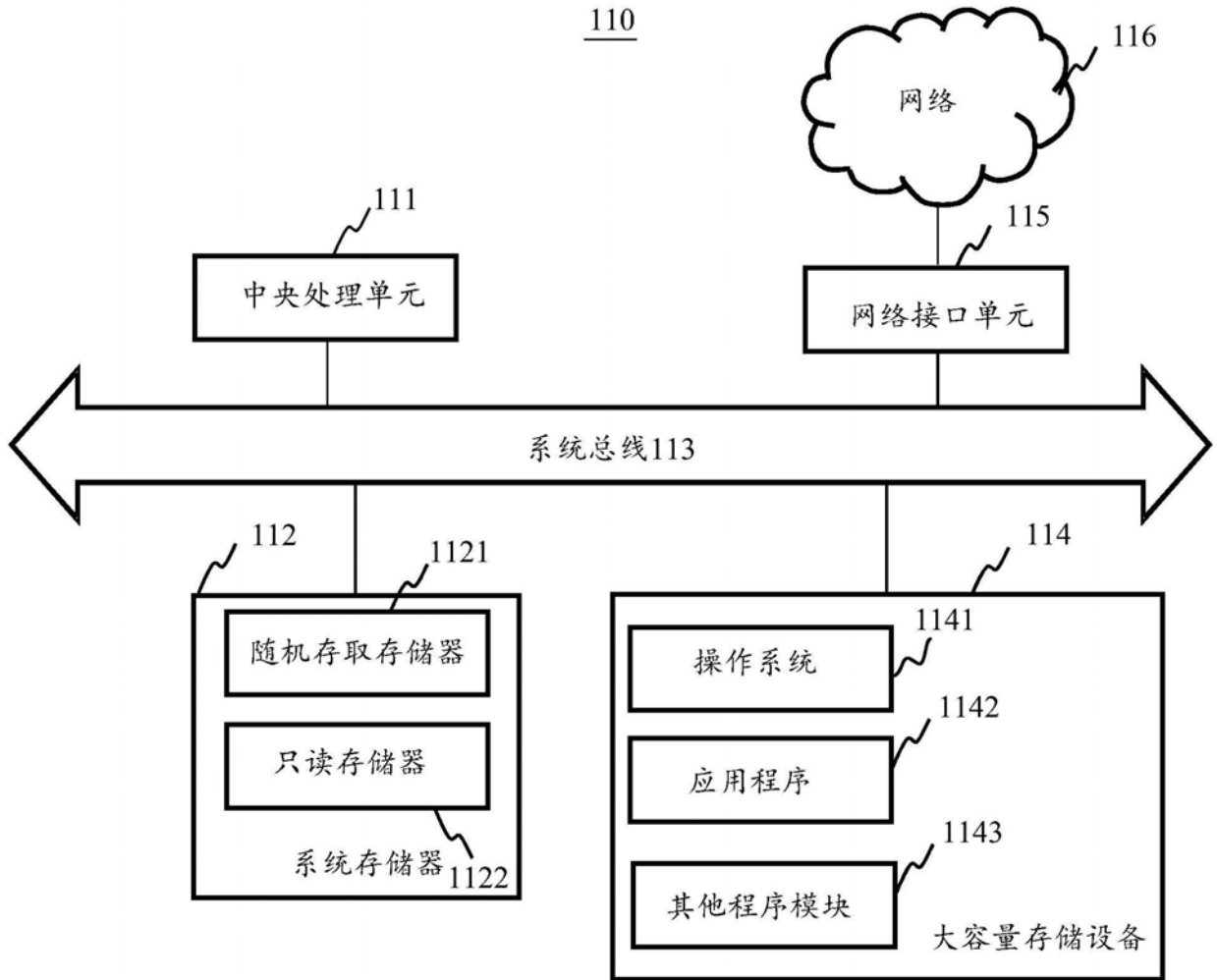


图11