



(12) 发明专利申请

(10) 申请公布号 CN 121980025 A

(43) 申请公布日 2026. 05. 05

(21) 申请号 202512019087.8

(22) 申请日 2025.12.30

(71) 申请人 中煤科工开采研究院有限公司

地址 101399 北京市顺义区中关村科技园  
区顺义园临空二路1号

(72) 发明人 吕依濛

(74) 专利代理机构 北京清亦华知识产权代理事

务所(普通合伙) 11201

专利代理师 汤宝平

(51) Int. Cl.

G06F 16/34 (2025.01)

G06F 40/216 (2020.01)

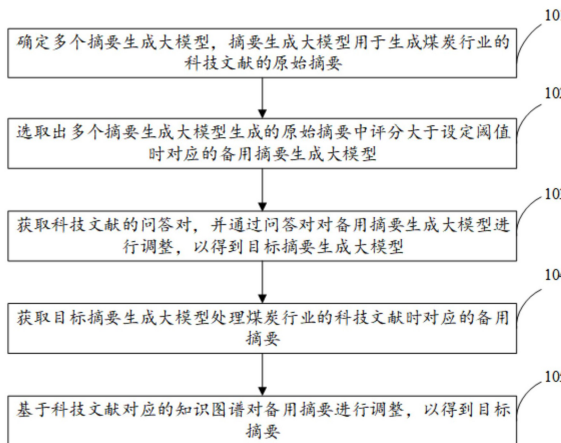
权利要求书2页 说明书8页 附图2页

(54) 发明名称

一种结合大模型技术的摘要生成方法及装置

(57) 摘要

本公开提出一种结合大模型技术的摘要生成方法及装置,方法包括:确定多个摘要生成大模型,摘要生成大模型用于生成煤炭行业的科技文献的原始摘要;选取多个摘要生成大模型生成的原始摘要中评分大于设定阈值时对应的备用摘要生成大模型;获取科技文献的问答对,并通过问答对对备用摘要生成大模型进行调整,以得到目标摘要生成大模型;获取目标摘要生成大模型处理煤炭行业的科技文献时对应的备用摘要;基于科技文献对应的知识图谱对备用摘要进行调整,以得到目标摘要。由此,通过问答对对备用摘要生成大模型的调整以及知识图谱对备用摘要的调整,使摘要不断优化,贴近煤炭开采垂直领域,显著提高科技文献摘要的质量和效率。



1. 一种结合大模型技术的摘要生成方法,其特征在于,所述方法包括:
  - 确定多个摘要生成大模型,所述摘要生成大模型用于生成煤炭行业的科技文献的原始摘要;
  - 选取多个摘要生成大模型生成的原始摘要中评分大于设定阈值时对应的备用摘要生成大模型;
  - 获取所述科技文献的问答对,并通过所述问答对对所述备用摘要生成大模型进行调整,以得到目标摘要生成大模型;
  - 获取所述目标摘要生成大模型处理所述煤炭行业的科技文献时对应的备用摘要;
  - 基于所述科技文献对应的知识图谱对所述备用摘要进行调整,以得到目标摘要。
2. 根据权利要求1所述的方法,其特征在于,所述选取多个摘要生成大模型生成的原始摘要中评分大于设定阈值时对应的备用摘要生成大模型,包括:
  - 获取各所述摘要生成大模型生成的原始摘要的第一评分;其中,所述第一评分是通过所述原始摘要与备用摘要之间的文本质量评估指标计算得到;
  - 获取各所述摘要生成大模型生成的原始摘要的第二评分;其中,所述第二评分是通过预设的用户评分机制、所述问答对的浏览数、采用数、讨论数计算得到;
  - 获取各所述摘要生成大模型的评分平均值和权重,所述评分平均值为各所述摘要生成大模型生成的多个原始摘要的评分的均值;
  - 基于所述第一评分、第二评分、问答对对应知识服务平台的活跃度参数、评分平均值和权重,计算出各所述摘要生成大模型生成的原始摘要的评分;
  - 选取所述原始摘要的评分大于设定阈值时对应的备用摘要生成大模型。
3. 根据权利要求2所述的方法,其特征在于,所述问答对对应知识服务平台的活跃度参数是通过所述知识服务平台的月活跃数、所述问答对的浏览数、采用数、讨论数计算得到。
4. 根据权利要求2所述的方法,其特征在于,所述获取所述科技文献的问答对,包括:
  - 从所述知识服务平台中获取所述科技文献对应的多个初始问答对;
  - 根据各初始问答对的浏览数、采用数、讨论数,确定出各初始问答对的选用概率;
  - 将选用概率大于设定概率阈值的问答对作为科技文献的问答对。
5. 根据权利要求1所述的方法,其特征在于,所述基于所述科技文献对应的知识图谱对所述备用摘要进行调整,以得到目标摘要,包括:
  - 获取所述科技文献对应的知识图谱;
  - 确定所述备用摘要中煤炭行业的文本;
  - 将所述知识图谱作为向量对所述煤炭行业的文本进行梳理调整,以得到目标摘要。
6. 一种结合大模型技术的摘要生成装置,其特征在于,所述装置包括:
  - 确定模块,用于确定多个摘要生成大模型,所述摘要生成大模型用于生成煤炭行业的科技文献的原始摘要;
  - 选取模块,用于选取多个摘要生成大模型生成的原始摘要中评分大于设定阈值时对应的备用摘要生成大模型;
  - 第一调整模块,用于获取所述科技文献的问答对,并通过所述问答对对所述备用摘要生成大模型进行调整,以得到目标摘要生成大模型;
  - 获取模块,用于获取所述目标摘要生成大模型处理所述煤炭行业的科技文献时对应的

备用摘要；

第二调整模块,用于基于所述科技文献对应的知识图谱对所述备用摘要进行调整,以得到目标摘要。

7. 根据权利要求6所述的装置,其特征在于,所述选取模块,具体用于:

获取各所述摘要生成大模型生成的原始摘要的第一评分;其中,所述第一评分是通过所述原始摘要与备用摘要之间的文本质量评估指标计算得到;

获取各所述摘要生成大模型生成的原始摘要的第二评分;其中,所述第二评分是通过预设的用户评分机制、所述问答对的浏览数、采用数、讨论数计算得到;

获取各所述摘要生成大模型的评分平均值和权重,所述评分平均值为各所述摘要生成大模型生成的多个原始摘要的评分的均值;

基于所述第一评分、第二评分、问答对对应知识服务平台的活跃度参数、评分平均值和权重,计算出各所述摘要生成大模型生成的原始摘要的评分;

选取出所述原始摘要的评分大于设定阈值时对应的备用摘要生成大模型。

8. 根据权利要求7所述的装置,其特征在于,所述问答对对应知识服务平台的活跃度参数是通过所述知识服务平台的月活跃数、所述问答对的浏览数、采用数、讨论数计算得到。

9. 根据权利要求7所述的装置,其特征在于,所述第一调整模块,具体用于:

从所述知识服务平台中获取所述科技文献对应的多个初始问答对;

根据各初始问答对的浏览数、采用数、讨论数,确定出各初始问答对的选用概率;

将选用概率大于设定概率阈值的问答对作为科技文献的问答对。

10. 根据权利要求6所述的装置,其特征在于,所述第二调整模块,具体用于:

获取所述科技文献对应的知识图谱;

确定所述备用摘要中煤炭行业的文本;

将所述知识图谱作为向量对所述煤炭行业的文本进行梳理调整,以得到目标摘要。

## 一种结合大模型技术的摘要生成方法及装置

### 技术领域

[0001] 本公开涉及摘要生成技术领域,尤其涉及一种结合大模型技术的摘要生成方法和装置。

### 背景技术

[0002] 在数字化时代,知识库的自动生成摘要技术是一种成熟的自然语言处理技术,传统自动生成摘要技术分为抽取式和生成式文本摘要两种,识别文本中的核心句子和短语,自动生成摘要文本,但在煤炭开采领域,由于对于煤炭开采垂直领域,过于专业,传统的自动摘要技术处理技术报告生成的摘要逻辑性极差,无法深度挖掘科技文献的专业词汇及逻辑,导致摘要冗余多,错误多。

### 发明内容

[0003] 本公开提供一种结合大模型技术的摘要生成方法和装置,以至少在一定程度上解决相关技术中的技术问题之一。本公开的技术方案如下:

根据本公开实施例的第一方面,提供一种结合大模型技术的摘要生成方法,所述方法包括:确定多个摘要生成大模型,所述摘要生成大模型用于生成煤炭行业的科技文献的原始摘要;选取出多个摘要生成大模型生成的原始摘要中评分大于设定阈值时对应的备用摘要生成大模型;获取所述科技文献的问答对,并通过所述问答对对所述备用摘要生成大模型进行调整,以得到目标摘要生成大模型;获取所述目标摘要生成大模型处理所述煤炭行业的科技文献时对应的备用摘要;基于所述科技文献对应的知识图谱对所述备用摘要进行调整,以得到目标摘要。

[0004] 根据本公开实施例的第二方面,提供了一种结合大模型技术的摘要生成装置,所述装置包括:确定模块,用于确定多个摘要生成大模型,所述摘要生成大模型用于生成煤炭行业的科技文献的原始摘要;选取模块,用于选取出多个摘要生成大模型生成的原始摘要中评分大于设定阈值时对应的备用摘要生成大模型;第一调整模块,用于获取所述科技文献的问答对,并通过所述问答对对所述备用摘要生成大模型进行调整,以得到目标摘要生成大模型;获取模块,用于获取所述目标摘要生成大模型处理所述煤炭行业的科技文献时对应的备用摘要;第二调整模块,用于基于所述科技文献对应的知识图谱对所述备用摘要进行调整,以得到目标摘要。

[0005] 根据本公开实施例的第三方面,提供一种电子设备,包括:处理器;用于存储所述处理器可执行指令的存储器;其中,所述处理器被配置为执行所述指令,以实现如本公开第一方面实施例所述的结合大模型技术的摘要生成方法。

[0006] 根据本公开实施例的第四方面,提供一种计算机可读存储介质,当所述计算机可读存储介质中的指令由电子设备的处理器执行时,使得电子设备能够执行如本公开第一方面实施例所述的结合大模型技术的摘要生成方法。

[0007] 本公开的实施例提供的技术方案至少带来以下有益效果:

在本技术方案中,确定多个摘要生成大模型,摘要生成大模型用于生成煤炭行业的科技文献的原始摘要;选取出多个摘要生成大模型生成的原始摘要中评分大于设定阈值时对应的备用摘要生成大模型;获取科技文献的问答对,并通过问答对对备用摘要生成大模型进行调整,以得到目标摘要生成大模型;获取目标摘要生成大模型处理煤炭行业的科技文献时对应的备用摘要;基于科技文献对应的知识图谱对备用摘要进行调整,以得到目标摘要。由此,通过问答对对备用摘要生成大模型的调整以及知识图谱对备用摘要的调整,使摘要不断优化,贴近煤炭开采垂直领域,显著提高科技文献摘要的质量和效率。

[0008] 应当理解的是,以上的一般描述和后文的细节描述仅是示例性和解释性的,并不能限制本公开。

### 附图说明

[0009] 此处的附图被并入说明书中并构成本说明书的一部分,示出了符合本公开的实施例,并与说明书一起用于解释本公开的原理,并不构成对本公开的不当限定。

[0010] 图1是本公开第一实施例所示出的结合大模型技术的摘要生成方法的流程示意图;

图2是本公开第二实施例所示出的结合大模型技术的摘要生成方法的原理示意图;

图3是本公开第三实施例所示出的结合大模型技术的摘要生成装置的结构示意图;

图4是本公开一示例性实施例所示出的电子设备的结构示意图。

### 具体实施方式

[0011] 为了使本领域普通人员更好地理解本公开的技术方案,下面将结合附图,对本公开实施例中的技术方案进行清楚、完整地描述。

[0012] 需要说明的是,本公开的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。应该理解这样使用的数据在适当情况下可以互换,以便这里描述的本公开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。以下示例性实施例中所描述的实施方式并不代表与本公开相一致的所有实施方式。相反,它们仅是与如所附权利要求书中所详述的、本公开的一些方面相一致的装置和方法的例子。

[0013] 需要说明的是,本公开的技术方案中,所涉及的用户个人信息的收集、存储、使用、加工、传输、提供和公开等处理,均在征得用户同意的前提下进行,并且均符合相关法律法规的规定,且不违背公序良俗。

[0014] 下面参考附图描述本公开实施例的结合大模型技术的摘要生成方法和装置。

[0015] 图1是本公开第一实施例所示出的结合大模型技术的摘要生成方法的流程示意图。

[0016] 如图1所示,该结合大模型技术的摘要生成方法包括如下步骤:

步骤101,确定多个摘要生成大模型,摘要生成大模型用于生成煤炭行业的科技文献的原始摘要。

[0017] 需要说明的是,科技文献包括但不限于煤炭开采领域的技术报告、论文、专利等。

[0018] 在本公开实施例中,可通过外接应用程序编程接口(Application Programming Interface,API)调用多个通用大模型(如GPT4,Deepseek等)作为摘要生成大模型处理煤炭行业的科技文献,以得到对应的原始摘要 $\{a_1, a_2, \dots, a_n\}$ ,该部分的原始摘要侧重结果增强,一个科技文献生成多份原始摘要,以增强摘要的多样性。其中,通过调用各煤炭行业主流通用大模型,调用生成的多份原始摘要的数量分别为 $s_1, s_2, \dots, s_n$ ,几个通用大模型的评分总和为 $p_1, p_2, \dots, p_n$ 。其中, $\sum_1^n S_n = n, S_n = \lfloor \text{softmax}(P_n) * n \rfloor$ ,softmax为将一组数值转换为概率分布的函数。

[0019] 作为一种示例,以煤炭行的知识服务平台为例,可通过知识服务平台的API调用多个通用大模型,以快速准确生成煤炭行业的科技文献的原始摘要,例如,外接了GPT3.5, Deepseek R1, Q-wen 7B三个通用大模型。

[0020] 步骤102,选取出多个摘要生成大模型生成的原始摘要中评分大于设定阈值时对应的备用摘要生成大模型。

[0021] 在本公开实施例中,为了优化摘要,提升摘要质量,作为一种示例,获取各摘要生成大模型生成的原始摘要的第一评分;其中,第一评分是通过原始摘要与备用摘要之间的文本质量评估指标计算得到;获取各摘要生成大模型生成的原始摘要的第二评分;其中,第二评分是通过预设的用户评分机制、问答对的浏览数、采用数、讨论数计算得到;获取各摘要生成大模型的评分平均值和权重,评分平均值为各摘要生成大模型生成的多个原始摘要的评分的均值;基于第一评分、第二评分、问答对对应知识服务平台的活跃度参数、评分平均值和权重,计算出各摘要生成大模型生成的原始摘要的评分;选取出原始摘要的评分大于设定阈值时对应的备用摘要生成大模型。备用摘要为目标摘要生成大模型生成的摘要,目标摘要生成大模型是通过对摘要生成大模型进行评分选取后通过问答对进行调整得到的,此外,还可以根据实际需要调整API调用摘要生成大模型的数量和种类,生成摘要数量和比例,根据原始摘要的评分进行调整。

[0022] 其中,需要说明的是,问答对对应知识服务平台的活跃度参数是通过知识服务平台的月活跃数、问答对的浏览数c、采用数d、讨论数e计算得到。

[0023] 为了清楚原始摘要中评分的计算方式,例如,这里使用原始摘要对比备用摘要,原始摘要为处理知识服务平台科技文献数据生成的摘要,默认选择评分最高的前n个摘要,如果摘要尚未评分,则随机选择n个摘要,记为abstract(n),生成摘要记为text。然后通过ROUGE计算第一评分,ROUGE(Recall-Oriented Understudy for Gisting Evaluation)是自然语言生成领域中的一类自动化文本质量评估指标,尤其常用于摘要、生成式问答、对话等任务的结果评估。其中,第一评分 $P_1$ 的计算方式如下:

$$P_1 = \frac{A * \lambda_1 + B * \lambda_2 + C * \lambda_3 + D * \lambda_4 + E * \lambda_5 + F * \lambda_6}{\sum \lambda}$$

[0024] 其中,ROUGE-1,用于统计原始摘要与备用摘要之间一元词组的重合数,记为ROUGE-1(text, abstract(n))作为上述公式中的A,调节参数设为 $\lambda_1$ 。ROUGE-2,用于统计原始摘要与备用摘要之间二元词组的重合数,记为ROUGE-2(text, abstract(n))作为上述公式中的B,调节参数设为 $\lambda_2$ 。ROUGE-L,用于统计原始摘要与备用摘要之间最长公共子序列

(LCS),记为ROUGE-L(text,abstract(n))作为上述公式中的C,调节参数设为 $\lambda_3$ 。同时计算原始摘要与备用摘要的TF-IDF。TF-IDF(Term Frequency-Inverse Document Frequency),是文本表示和信息检索中的特征提取方法,记为TF-IDF(text,abstract(n))作为上述公式中的D,调节参数设为 $\lambda_4$ 。并计算两个摘要的余弦相似度(Cosine Similarity),记为cosine\_similarity(text,abstract(n))作为上述公式中的E,调节参数设为 $\lambda_5$ 。BLEU(双语评估替代方法)用于衡量机器原始摘要与备用摘要之间的相似程度,记为BLEU(text,abstract(n))作为上述公式中的F,调节参数设为 $\lambda_6$ 。 $\lambda_1, \dots, \lambda_6$ 为根据实际情况设定的调节参数,视实际运行情况而定,值越大,相应技术参数越重要。

[0025] 第二评分 $P_{hu}$ 的计算方式如下:例如,将人工评估作为第二评分 $P_{hu}$ ,具体地,采用煤炭开采专家评分机制作为用户评分机制,取所有专家用户评分的平均分 $P_{ave}$ 作为加权,知识服务平台用户越活跃,该评估价值越高,这里统计知识服务平台的月活跃数f,活跃度参数记为 $\alpha$ ,其中,

$$P_{hu} > \frac{d}{c} \ln \frac{1 + \frac{e}{c}}{1 - \frac{e}{c}} + \cos \frac{d}{c} + P_{ave}$$

[0026]  $\alpha = \log_f(d+e) + \log_f(c)$ 。

[0027] 其中,各摘要生成大模型生成的多个原始摘要的评分的均值记为 $P_n$ ,相应摘要生成大模型的权重记为 $\beta_n$ , $\sum \beta_n = 1$ 。

[0028] 综上,原始摘要的评分的计算公式可以为: $P = (1-\alpha) * P_t + P_{hu} * \alpha + \sum_1^n P_n \beta_n$ ,评分反应了原始摘要的质量,评分越高,质量越好,将评分P反馈给目标摘要生成大模型及备用摘要生成大模型进行模型微调。

[0029] 步骤103,获取科技文献的问答对,并通过问答对对备用摘要生成大模型进行调整,以得到目标摘要生成大模型。

[0030] 在本公开实施例中,为了筛选出高质量问答对,以用于优化备用摘要生成大模型,作为一种示例,从知识服务平台中获取科技文献对应的多个初始问答对;根据各初始问答对的浏览数、采用数、讨论数,确定出各初始问答对的选用概率;将选用概率大于设定概率阈值的问答对作为科技文献的问答对。

[0031] 例如,知识服务平台生成的关于某个科技文献的问答对 $\{b_1, b_2, \dots, b_n\}$ ,每个问答对的浏览数为c,采用数为d,讨论数为e,设定一个阈值 $P_1$ 作为选用概率,当 $P_1$ 大于设定概率阈值即满足下式时:

$$P_1 > \frac{d}{c} \ln \frac{1 + \frac{e}{c}}{1 - \frac{e}{c}} + \cos \frac{d}{c}$$

[0032] 选取出高质量的所需的科技文献的问答对。

[0033] 步骤104,获取目标摘要生成大模型处理煤炭行业的科技文献时对应的备用摘要。

[0034] 需要说明的是,科技文献还可以是煤炭开采、安全、机电设备维护、碳排放治理等领域的期刊论文、专利、会议报告等。

[0035] 步骤105,基于科技文献对应的知识图谱对备用摘要进行调整,以得到目标摘要。

[0036] 在本公开实施例中,为了增强摘要的准确性、完整性与逻辑性,使其更符合煤炭行业场景下的信息需求,作为一种示例,获取科技文献对应的知识图谱;确定备用摘要中煤炭行业的文本;将知识图谱作为向量对煤炭行业的文本进行梳理调整,以得到目标摘要。

[0037] 其中,科技文献对应的知识图谱是以科技文献为核心数据源,通过可视化技术揭示学科领域知识结构、研究热点及发展趋势的图谱化工具。始数据为自动。

[0038] 综上,确定多个摘要生成大模型,摘要生成大模型用于生成煤炭行业的科技文献的原始摘要;选取出多个摘要生成大模型生成的原始摘要中评分大于设定阈值时对应的备用摘要生成大模型;获取科技文献的问答对,并通过问答对对备用摘要生成大模型进行调整,以得到目标摘要生成大模型;获取目标摘要生成大模型处理煤炭行业的科技文献时对应的备用摘要;基于科技文献对应的知识图谱对备用摘要进行调整,以得到目标摘要。由此,通过问答对对备用摘要生成大模型的调整以及知识图谱对备用摘要的调整,使摘要不断优化,贴近煤炭开采垂直领域,显著提高科技文献摘要的质量和效率。

[0039] 为了清楚的描述结合大模型技术的摘要生成方法,如图2所示,图2是本公开实施例所示出的一种结合大模型技术的摘要生成方法的原理示意图。具体地,微调数据生成模块:用于通过确定多个摘要生成大模型处理煤炭行业的科技文献,以生成原始摘要,同时通过评价生成模块计算原始摘要中评分,以选取出评分大于设定阈值时对应的备用摘要生成大模型;同时获取到科技文献的问答对以反馈至微调数据生成模块作为微调数据,对备用摘要生成大模型进行煤炭开采垂直领域调整,以得到目标摘要生成大模型;然后获取目标摘要生成大模型处理煤炭行业的科技文献时对应的备用摘要;同时基于图谱数据生成模块,获取科技文献对应的知识图谱,以通过知识图谱的图谱数据对煤炭行业的文本进行梳理调整,以得到目标摘要。由此通过知识图谱的向量梳理及大模型微调,显著提高科技文献摘要的质量、效率及逻辑性。

[0040] 与上述实施例提供的结合大模型技术的摘要生成方法相对应,本公开还提供一种结合大模型技术的摘要生成装置,由于本公开实施例提供的结合大模型技术的摘要生成装置与上述实施例提供的结合大模型技术的摘要生成方法相对应,因此在结合大模型技术的摘要生成方法的实施方式也适用于本公开实施例提供的结合大模型技术的摘要生成装置,在本公开实施例中不再详细描述。

[0041] 图3是本公开第三实施例所示出的结合大模型技术的摘要生成装置的结构示意图。

[0042] 如图3所示,该结合大模型技术的摘要生成装置300包括:确定模块301、选取模块302、第一调整模块303、获取模块304、第二调整模块305。

[0043] 其中,确定模块301,用于确定多个摘要生成大模型,所述摘要生成大模型用于生成煤炭行业的科技文献的原始摘要;选取模块302,用于选取出多个摘要生成大模型生成的原始摘要中评分大于设定阈值时对应的备用摘要生成大模型;第一调整模块303,用于获取所述科技文献的问答对,并通过所述问答对对所述备用摘要生成大模型进行调整,以得到目标摘要生成大模型;获取模块304,用于获取所述目标摘要生成大模型处理所述煤炭行业

的科技文献时对应的备用摘要；第二调整模块305,用于基于所述科技文献对应的知识图谱对所述备用摘要进行调整,以得到目标摘要。

[0044] 作为本公开实施例的一种可能的实现方式,所述选取模块302,具体用于:获取各所述摘要生成大模型生成的原始摘要的第一评分;其中,所述第一评分是通过所述原始摘要与备用摘要之间的文本质量评估指标计算得到;获取各所述摘要生成大模型生成的原始摘要的第二评分;其中,所述第二评分是通过预设的用户评分机制、所述问答对的浏览数、采用数、讨论数计算得到;获取各所述摘要生成大模型的评分平均值和权重,所述评分平均值为各所述摘要生成大模型生成的多个原始摘要的评分的均值;基于所述第一评分、第二评分、问答对对应知识服务平台的活跃度参数、评分平均值和权重,计算出各所述摘要生成大模型生成的原始摘要的评分;选取出所述原始摘要的评分大于设定阈值时对应的备用摘要生成大模型。

[0045] 作为本公开实施例的一种可能的实现方式,所述问答对对应知识服务平台的活跃度参数是通过所述知识服务平台的月活跃数、所述问答对的浏览数、采用数、讨论数计算得到。

[0046] 作为本公开实施例的一种可能的实现方式,所述第一调整模块303,具体用于:从所述知识服务平台中获取所述科技文献对应的多个初始问答对;根据各初始问答对的浏览数、采用数、讨论数,确定出各初始问答对的选用概率;将选用概率大于设定概率阈值的问答对作为科技文献的问答对。

[0047] 作为本公开实施例的一种可能的实现方式,所述第二调整模块305,具体用于:获取所述科技文献对应的知识图谱;确定所述备用摘要中煤炭行业的文本;将所述知识图谱作为向量对所述煤炭行业的文本进行梳理调整,以得到目标摘要。

[0048] 本公开实施例的结合大模型技术的摘要生成装置,确定多个摘要生成大模型,摘要生成大模型用于生成煤炭行业的科技文献的原始摘要;选取出多个摘要生成大模型生成的原始摘要中评分大于设定阈值时对应的备用摘要生成大模型;获取科技文献的问答对,并通过问答对对备用摘要生成大模型进行调整,以得到目标摘要生成大模型;获取目标摘要生成大模型处理煤炭行业的科技文献时对应的备用摘要;基于科技文献对应的知识图谱对备用摘要进行调整,以得到目标摘要。由此,通过问答对对备用摘要生成大模型的调整以及知识图谱对备用摘要的调整,使摘要不断优化,贴近煤炭开采垂直领域,显著提高科技文献摘要的质量和效率。

[0049] 在示例性实施例中,还提出了一种电子设备。

[0050] 其中,电子设备包括:

处理器;

用于存储处理器可执行指令的存储器;

其中,处理器被配置为执行指令,以实现如前述任一实施例提出的结合大模型技术的摘要生成方法。

[0051] 作为一种示例,图4是本公开一示例性实施例所示出的电子设备400的结构示意图,如图4所示,上述电子设备400,还可以包括:

存储器410及处理器420,连接不同组件(包括存储器410和处理器420)的总线430,存储器410存储有计算机程序,当处理器420执行程序时实现本公开实施例所述的结合大模

型技术的摘要生成方法。

[0052] 总线430表示几类总线结构中的一种或多种,包括存储器总线或者存储器控制器,外围总线,图形加速端口,处理器或者使用多种总线结构中的任意总线结构的局域总线。举例来说,这些体系结构包括但不限于工业标准体系结构 (ISA) 总线,微通道体系结构 (MAC) 总线,增强型ISA总线、视频电子标准协会 (VESA) 局域总线以及外围组件互连 (PCI) 总线。

[0053] 电子设备400典型地包括多种电子设备可读介质。这些介质可以是任何能够被电子设备400访问的可用介质,包括易失性和非易失性介质,可移动的和不可移动的介质。

[0054] 存储器410还可以包括易失性存储器形式的计算机系统可读介质,例如随机存取存储器 (RAM) 440和/或高速缓存存储器450。服务器可以进一步包括其它可移动/不可移动的、易失性/非易失性计算机系统存储介质。仅作为举例,存储系统460可以用于读写不可移动的、非易失性磁介质 (图4未显示,通常称为“硬盘驱动器”)。尽管图4中未示出,可以提供用于对可移动非易失性磁盘 (例如“软盘”) 读写的磁盘驱动器,以及对可移动非易失性光盘 (例如CD-ROM, DVD-ROM或者其它光介质) 读写的光盘驱动器。在这些情况下,每个驱动器可以通过一个或者多个数据介质接口与总线430相连。存储器410可以包括至少一个程序产品,该程序产品具有一组 (例如至少一个) 程序模块,这些程序模块被配置以执行本公开各实施例的功能。

[0055] 具有一组 (至少一个) 程序模块470的程序/实用工具480,可以存储在例如存储器410中,这样的程序模块470包括但不限于操作系统、一个或者多个应用程序、其它程序模块以及程序数据,这些示例中的每一个或某种组合中可能包括网络环境的实现。程序模块470通常执行本公开所描述的实施例中的功能和/或方法。

[0056] 电子设备400也可以与一个或多个外部设备490 (例如键盘、指向设备、显示器491等) 通信,还可与一个或者多个使得用户能与该电子设备400交互的设备通信,和/或与使得该电子设备400能与一个或多个其它计算设备进行通信的任何设备 (例如网卡,调制解调器等等) 通信。这种通信可以通过输入/输出 (I/O) 接口492进行。并且,电子设备400还可以通过网络适配器493与一个或者多个网络 (例如局域网 (LAN), 广域网 (WAN) 和/或公共网络,例如因特网) 通信。如图所示,网络适配器493通过总线430与电子设备400的其它模块通信。应当明白,尽管图中未示出,可以结合电子设备400使用其它硬件和/或软件模块,包括但不限于:微代码、设备驱动器、冗余处理单元、外部磁盘驱动阵列、RAID系统、磁带驱动器以及数据备份存储系统等。

[0057] 处理器420通过运行存储在存储器410中的程序,从而执行各种功能应用以及数据处理。

[0058] 需要说明的是,本实施例的电子设备的实施过程和技术原理参见前述对本公开实施例的结合大模型技术的摘要生成方法的解释说明,此处不再赘述。

[0059] 在示例性实施例中,还提供了一种包括指令的计算机可读存储介质,例如包括指令的存储器,上述指令可由电子设备的处理器执行以完成上述任一实施例提出的结合大模型技术的摘要生成方法。可选地,计算机可读存储介质可以是ROM、随机存取存储器 (RAM)、CD-ROM、磁带、软盘和光数据存储设备等。

[0060] 在示例性实施例中,还提供一种计算机程序产品,包括计算机程序/指令,所述计算机程序/指令被处理器执行时实现上述任一实施例提出的结合大模型技术的摘要生成方

法。

[0061] 本领域技术人员在考虑说明书及实践这里公开的发明后,将容易想到本公开的其它实施方案。本公开旨在涵盖本公开的任何变型、用途或者适应性变化,这些变型、用途或者适应性变化遵循本公开的一般性原理并包括本公开未公开的本技术领域中的公知常识或惯用技术手段。说明书和实施例仅被视为示例性的,本公开的真正范围和精神由下面的权利要求指出。

[0062] 应当理解的是,本公开并不局限于上面已经描述并在附图中示出的精确结构,并且可以在不脱离其范围进行各种修改和改变。本公开的范围仅由所附的权利要求来限制。

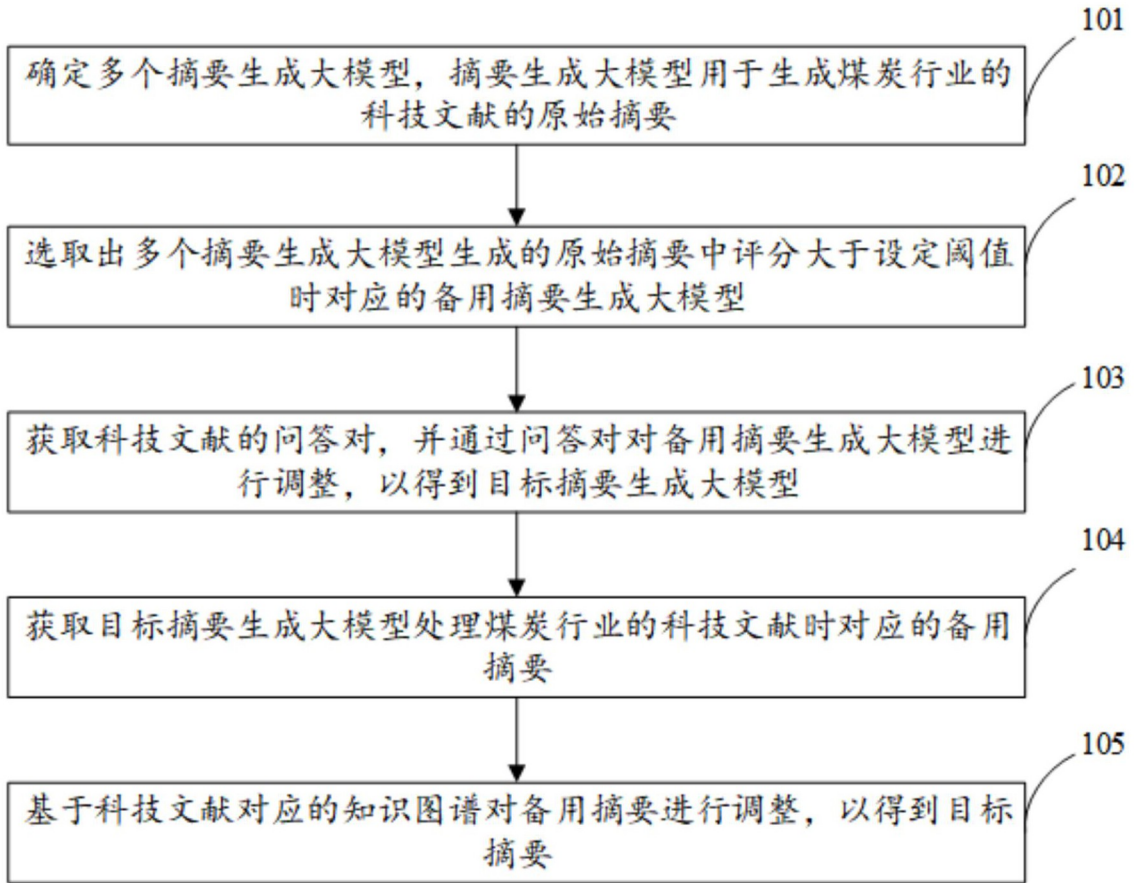


图1

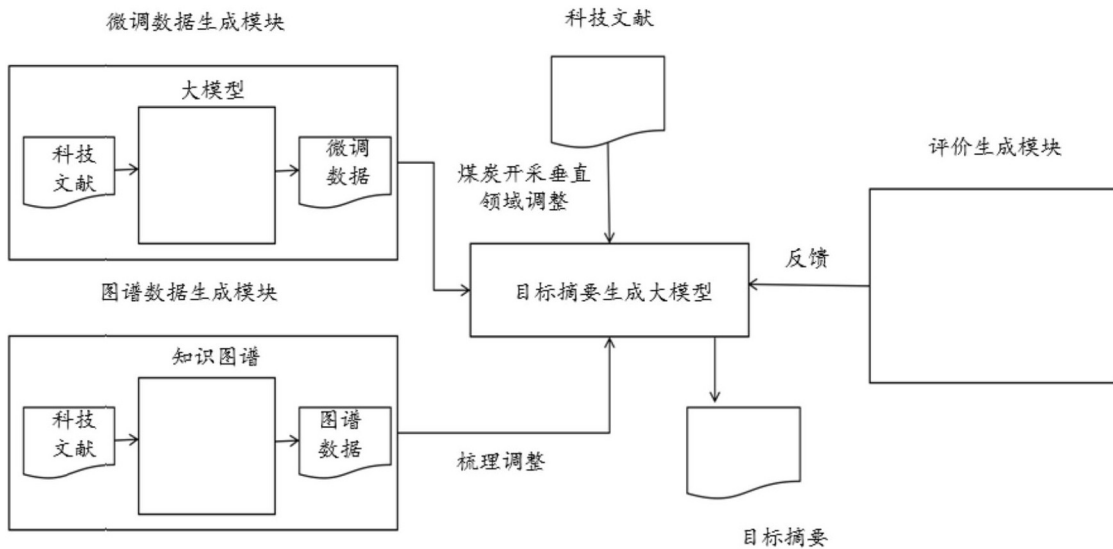


图2

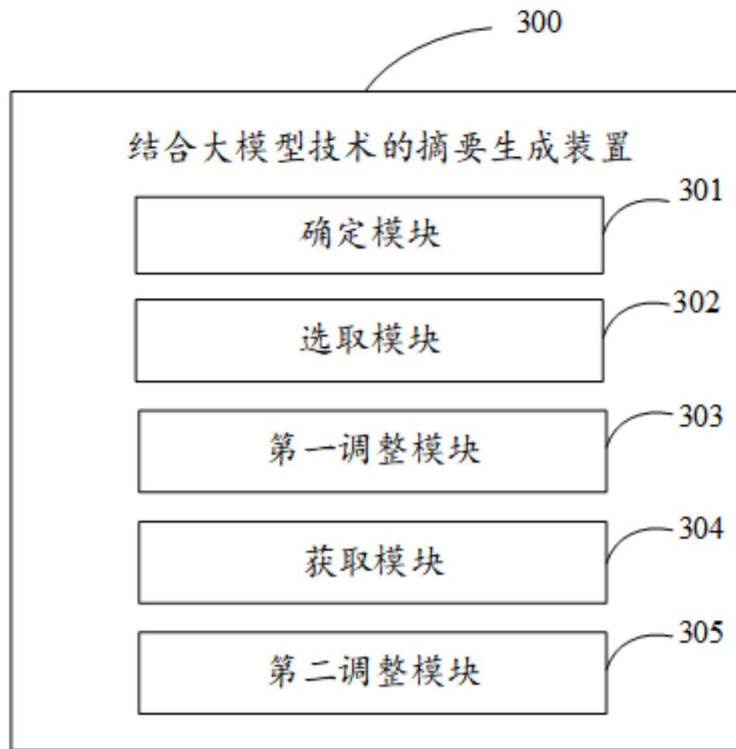


图3

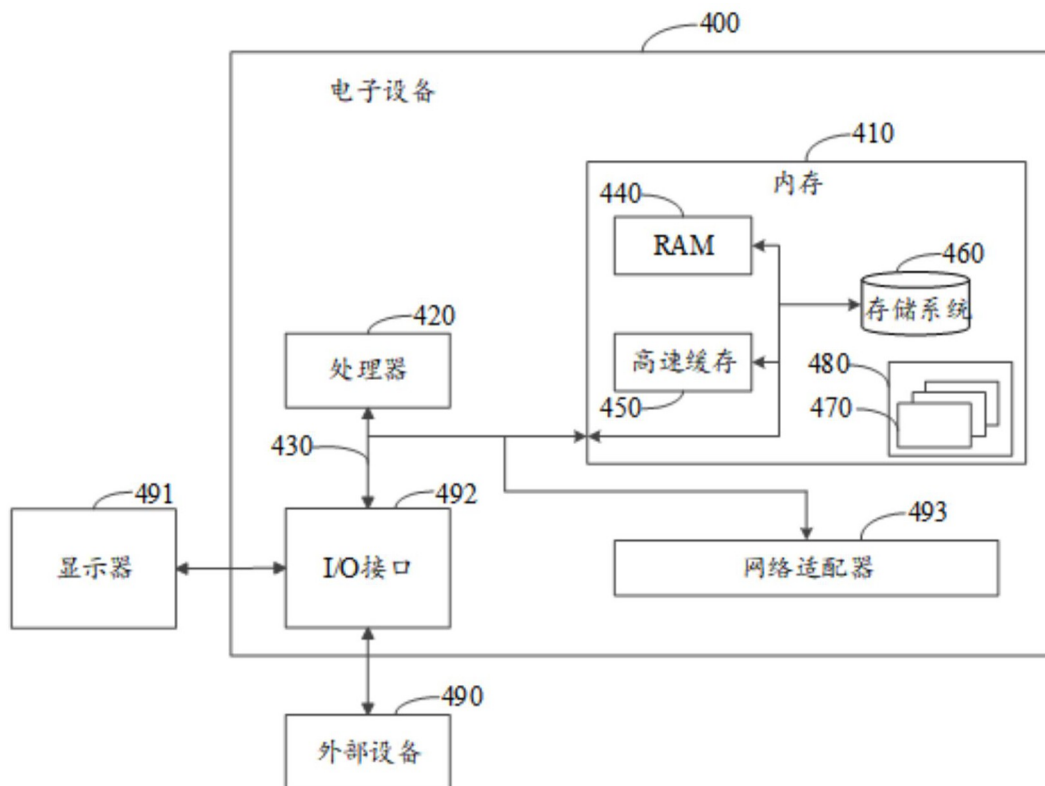


图4