(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 114625866 A
(43)申请公布日 2022.06.14

(21)申请号 202210238223.6

(22)申请日 2022.03.11

(71)申请人 腾讯科技(深圳)有限公司
地址 518044 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 刘维 吴焕钦 牟文晶

(74)专利代理机构 北京同达信恒知识产权代理有限公司 11291
专利代理师 朱佳

(51)Int.Cl.
G06F 16/34(2019.01)
G06F 16/33(2019.01)
G06F 40/211(2020.01)
G06F 40/216(2020.01)
G06N 20/00(2019.01)

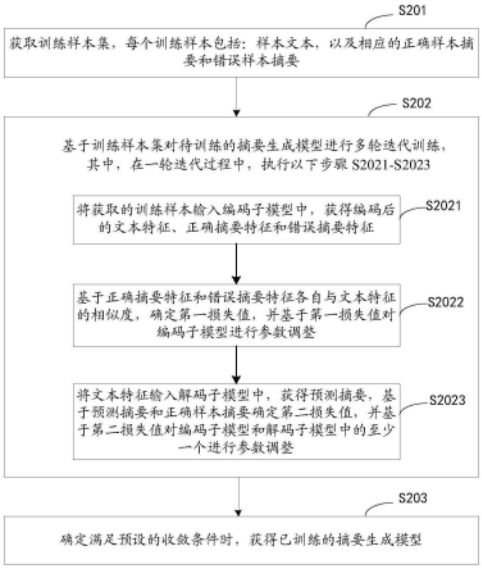
权利要求书2页 说明书17页 附图8页

(54)发明名称

训练摘要生成模型的方法、装置、设备及介质

(57)摘要

本申请提供一种训练摘要生成模型的方法、装置、设备及介质，涉及人工智能技术领域，可应用于各种场景，包括但不限于云技术、人工智能、智慧交通、辅助驾驶等。方法包括：基于训练样本集对待训练的摘要生成模型进行多轮迭代训练，在一轮迭代过程中，将训练样本输入编码子模型中，获得编码后的文本特征、正确摘要特征和错误摘要特征；基于正确摘要特征和错误摘要特征各自与文本特征的相似度，确定第一损失值，并基于第一损失值对编码子模型进行参数调整；将文本特征输入解码子模型中，获得预测摘要，基于由预测摘要和正确样本摘要确定的第二损失值，对编码子模型和解码子模型进行参数调整。本申请实施例可以提高摘要生成模型的事实一致性。



CN 114625866 A

1.一种训练摘要生成模型的方法,其特征在于,所述摘要生成模型包括编码子模型和解码子模型,所述方法包括:

基于训练样本集对待训练的摘要生成模型进行多轮迭代训练,输出已训练的摘要生成模型,每个训练样本包括:样本文本,以及相应的正确样本摘要和错误样本摘要;其中,在一轮迭代过程中,执行以下操作:

将获取的训练样本输入所述编码子模型中,获得编码后的文本特征、正确摘要特征和错误摘要特征;

基于所述正确摘要特征和所述错误摘要特征各自与所述文本特征的相似度,确定第一损失值,并基于所述第一损失值对所述编码子模型进行参数调整;

将所述文本特征输入所述解码子模型中,获得预测摘要,基于所述预测摘要和所述正确样本摘要确定第二损失值,并基于所述第二损失值对所述编码子模型和所述解码子模型中的至少一个进行参数调整。

2.根据权利要求1所述的方法,其特征在于,所述基于所述预测摘要和所述正确样本摘要确定第二损失值,包括:

基于所述预测摘要和所述正确样本摘要确定正确预测概率值,以及基于所述预测摘要和所述错误样本摘要确定错误预测概率值;

基于所述正确预测概率值和所述错误预测概率值,确定所述第二损失值。

3.根据权利要求1或2所述的方法,其特征在于,每个训练样本中的错误样本摘要通过如下方式获得:

从所述训练样本中的正确样本摘要中查找关键词集合;

针对所述关键词集合中的多个关键词,分别执行以下操作:若一个关键词的替换概率达到预设概率,则将所述一个关键词替换为相应的替换词;其中,所述替换词是从所述训练样本中的样本文本中选择的;

将替换后的正确样本摘要作为所述训练样本中的错误样本摘要。

4.根据权利要求3所述的方法,其特征在于,所述从所述训练样本中的正确样本摘要中查找关键词集合,包括:

从所述训练样本中的样本文本中,获得所述正确样本摘要中的至少一个摘要语句各自对应的上下文语句;

针对所述至少一个摘要语句,迭代执行多次以下操作:将所述至少一个摘要语句分别进行裁剪处理,获得所述至少一个摘要语句各自对应的词语片段,并将至少一个词语片段拼接成摘要片段,以及将每个词语片段作为新的一个摘要语句;

针对获得的多个摘要片段,分别执行以下操作:基于一个摘要片段以及所述正确样本摘要对应的至少一个上下文语句,确定所述一个摘要片段的评估值;

基于所述多个摘要片段各自的评估值,从所述多个摘要片段中选择目标摘要片段,并将所述目标摘要片段中的多个词语组成所述关键词集合。

5.根据权利要求4所述的方法,其特征在于,所述基于一个摘要片段以及所述正确样本摘要对应的至少一个上下文语句,确定所述一个摘要片段的评估值,包括:

针对所述至少一个上下文语句,分别执行以下操作:将一个上下文语句和所述一个摘要片段,输入自回归语言模型,获得第一语言预测概率;

基于获得的至少一个第一语言预测概率，获得所述一个摘要片段的相关性数值；

将所述一个摘要片段输入所述自回归语言模型，获得第二语言预测概率，并基于所述第二语言预测概率获得所述一个摘要片段的压缩率数值；

将所述一个摘要片段的相关性数值和压缩率数值，作为所述一个摘要片段的评估值。

6.根据权利要求5所述的方法，其特征在于，所述基于所述多个摘要片段各自的评估值，从所述多个摘要片段中选择目标摘要片段，包括：

从所述多个摘要片段中，选择压缩率数值满足预设条件的多个候选摘要片段；

将所述多个候选摘要片段中，相关性数值最大的候选摘要片段作为所述目标摘要片段。

7.一种训练摘要生成模型的装置，其特征在于，所述摘要生成模型包括编码子模型和解码子模型，所述装置包括：

训练模块，用于基于训练样本集对待训练的摘要生成模型进行多轮迭代训练，输出已训练的摘要生成模型，每个训练样本包括：样本文本，以及相应的正确样本摘要和错误样本摘要；其中，在一轮迭代过程中，执行以下操作：

将获取的训练样本输入所述编码子模型中，获得编码后的文本特征、正确摘要特征和错误摘要特征；

基于所述正确摘要特征和所述错误摘要特征各自与所述文本特征的相似度，确定第一损失值，并基于所述第一损失值对所述编码子模型进行参数调整；

将所述文本特征输入所述解码子模型中，获得预测摘要，基于所述预测摘要和所述正确样本摘要确定第二损失值，并基于所述第二损失值对所述编码子模型和所述解码子模型中的至少一个进行参数调整。

8.一种电子设备，其特征在于，其包括处理器和存储器，其中，所述存储器存储有程序代码，当所述程序代码被所述处理器执行时，使得所述处理器执行权利要求1～6中任一所述方法的步骤。

9.一种计算机可读存储介质，其特征在于，其包括程序代码，当所述程序代码在电子设备上运行时，所述程序代码用于使所述电子设备执行权利要求1～6中任一所述方法的步骤。

10.一种计算机程序产品，其特征在于，其包括计算机指令，所述计算机指令存储在计算机可读存储介质中；当计算机设备的处理器从所述计算机可读存储介质读取所述计算机指令时，所述处理器执行该计算机指令，使得所述计算机设备执行权利要求1～6中任一所述方法的步骤。

训练摘要生成模型的方法、装置、设备及介质

技术领域

[0001] 本申请涉及人工智能技术领域，尤其涉及一种训练摘要生成模型的方法、装置、设备及介质。

背景技术

[0002] 随着自然语言处理技术的不断发展，基于自然语言处理技术的摘要生成模型得到了广泛应用。摘要生成模型旨在获取文本的关键信息，并生成包含关键信息的简短摘要。

[0003] 摘要生成模型通常采用序列到序列模型，该模型包括编码器和解码器，编码器用于将文本进行文本编码得到向量，解码器用于从编码得到的向量中提取语义信息，以生成文本摘要。

[0004] 虽然，摘要生成模型便于快速生成文本的摘要，但是容易出现生成的摘要与文本中的事实不符的情况；例如：文本中包含了这样一个事实“人物A于2010年导演了一部电影，由人物B主演”，但是摘要生成模型可能会生成这样一个事实性的错误“人物B导演了一部电影”。

[0005] 因此，如何保证摘要生成模型生成的摘要与对应文本的事实一致性，是需要解决的问题。

发明内容

[0006] 本申请实施例提供一种训练摘要生成模型的方法、装置、电子设备及存储介质，用于实现摘要生成模型生成的摘要与对应文本的事实一致性。

[0007] 一方面，本申请实施例提供一种训练摘要生成模型的方法，包括：

[0008] 基于训练样本集对待训练的摘要生成模型进行多轮迭代训练，输出已训练的摘要生成模型，每个训练样本包括：样本文本，以及相应的正确样本摘要和错误样本摘要；其中，在一轮迭代过程中，执行以下操作：

[0009] 将获取的训练样本输入所述编码子模型中，获得编码后的文本特征、正确摘要特征和错误摘要特征；

[0010] 基于所述正确摘要特征和所述错误摘要特征各自与所述文本特征的相似度，确定第一损失值，并基于所述第一损失值对所述编码子模型进行参数调整；

[0011] 将所述文本特征输入所述解码子模型中，获得预测摘要，基于所述预测摘要和所述正确样本摘要确定第二损失值，并基于所述第二损失值对所述编码子模型和所述解码子模型中的至少一个进行参数调整。

[0012] 一方面，本申请实施例提供一种训练摘要生成模型的装置，所述摘要生成模型包括编码子模型和解码子模型，所述装置包括：

[0013] 训练模块，用于基于训练样本集对待训练的摘要生成模型进行多轮迭代训练，输出已训练的摘要生成模型，每个训练样本包括：样本文本，以及相应的正确样本摘要和错误样本摘要；其中，在一轮迭代过程中，执行以下操作：

[0014] 将获取的训练样本输入所述编码子模型中，获得编码后的文本特征、正确摘要特征和错误摘要特征；

[0015] 基于所述正确摘要特征和所述错误摘要特征各自与所述文本特征的相似度，确定第一损失值，并基于所述第一损失值对所述编码子模型进行参数调整；

[0016] 将所述文本特征输入所述解码子模型中，获得预测摘要，基于所述预测摘要和所述正确样本摘要确定第二损失值，并基于所述第二损失值对所述编码子模型和所述解码子模型中的至少一个进行参数调整。

[0017] 在一种可能的实施例中，所述基于所述预测摘要和所述正确样本摘要确定第二损失值时，所述训练模块还用于：

[0018] 基于所述预测摘要和所述正确样本摘要确定正确预测概率值，以及基于所述预测摘要和所述错误样本摘要确定错误预测概率值；

[0019] 基于所述正确预测概率值和所述错误预测概率值，确定所述第二损失值。

[0020] 在一种可能的实施例中，还包括获取模块，用于通过如下方式获得每个训练样本中的错误样本摘要：

[0021] 从所述训练样本中的正确样本摘要中查找关键词集合；

[0022] 针对所述关键词集合中的多个关键词，分别执行以下操作：若一个关键词的替换概率达到预设概率，则将所述一个关键词替换为相应的替换词；其中，所述替换词是从所述训练样本中的样本文本中选择的；

[0023] 将替换后的正确样本摘要作为所述训练样本中的错误样本摘要。

[0024] 在一种可能的实施例中，所述从所述训练样本中的正确样本摘要中查找关键词集合时，所述获取模块还用于：

[0025] 从所述训练样本中的样本文本中，获得所述正确样本摘要中的至少一个摘要语句各自对应的上下文语句；

[0026] 针对所述至少一个摘要语句，迭代执行多次以下操作：将所述至少一个摘要语句分别进行裁剪处理，获得所述至少一个摘要语句各自对应的词语片段，并将至少一个词语片段拼接成摘要片段，以及将每个词语片段作为新的一个摘要语句；

[0027] 针对获得的多个摘要片段，分别执行以下操作：基于一个摘要片段以及所述正确样本摘要对应的至少一个上下文语句，确定所述一个摘要片段的评估值；

[0028] 基于所述多个摘要片段各自的评估值，从所述多个摘要片段中选择目标摘要片段，并将所述目标摘要片段中的多个词语组成所述关键词集合。

[0029] 在一种可能的实施例中，所述基于一个摘要片段以及所述正确样本摘要对应的至少一个上下文语句，确定所述一个摘要片段的评估值时，所述获取模块还用于：

[0030] 针对所述至少一个上下文语句，分别执行以下操作：将一个上下文语句和所述一个摘要片段，输入自回归语言模型，获得第一语言预测概率；

[0031] 基于获得的至少一个第一语言预测概率，获得所述一个摘要片段的相关性数值；

[0032] 将所述一个摘要片段输入所述自回归语言模型，获得第二语言预测概率，并基于所述第二语言预测概率获得所述一个摘要片段的压缩率数值；

[0033] 将所述一个摘要片段的相关性数值和压缩率数值，作为所述一个摘要片段的评估值。

[0034] 在一种可能的实施例中，所述基于所述多个摘要片段各自的评估值，从所述多个摘要片段中选择目标摘要片段时，所述获取模块还用于：

[0035] 从所述多个摘要片段中，选择压缩率数值满足预设条件的多个候选摘要片段；

[0036] 将所述多个候选摘要片段中，相关性数值最大的候选摘要片段作为所述目标摘要片段。

[0037] 一方面，本申请实施例提供一种电子设备，其包括处理器和存储器，其中，所述存储器存储有程序代码，当所述程序代码被所述处理器执行时，使得所述处理器执行上述任一种训练摘要生成模型的方法的步骤。

[0038] 一方面，本申请实施例提供一种计算机存储介质，所述计算机存储介质存储有计算机指令，当所述计算机指令在计算机上运行时，使得计算机执行上述任一种训练摘要生成模型的方法的步骤。

[0039] 一方面，本申请实施例提供一种计算机程序产品，其包括计算机指令，所述计算机指令存储在计算机可读存储介质中；当计算机设备的处理器从所述计算机可读存储介质读取所述计算机指令时，所述处理器执行该计算机指令，使得所述计算机设备执行上述任一种训练摘要生成模型的方法的步骤。

[0040] 由于本申请实施例采用上述技术方案，至少具有如下技术效果：

[0041] 本申请实施例的方案中，采用样本文本、正确样本摘要和错误样本摘要构成的训练样本对摘要生成模型的编码子模型进行对比学习；具体地，将训练样本输入编码子模型中，获得编码后的文本特征、正确摘要特征和错误摘要特征，然后，基于正确摘要特征和错误摘要特征各自与文本特征的相似度，确定第一损失值，并基于第一损失值对编码子模型进行参数调整，使得正确样本摘要的编码和样本文本的编码更相似，错误样本摘要的编码和样本文本的编码更不相似；进一步地，将文本特征输入解码子模型中获得预测摘要，基于预测摘要和正确样本摘要确定第二损失值，并基于第二损失值对编码子模型和所述解码子模型中的至少一个进行参数调整；这样，使得训练完成的编码子模型可以正确编码文本中的事实信息，降低事实错误，进而使训练完成的解码子模型输出与文本具有事实一致性的文本摘要。

[0042] 本申请的其它特征和优点将在随后的说明书中阐述，并且，部分地从说明书中变得显而易见，或者通过实施本申请而了解。本申请的目的和其他优点可通过在所写的说明书、权利要求书、以及附图中所特别指出的结构来实现和获得。

附图说明

[0043] 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要介绍，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，对于本领域的普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

[0044] 图1为本申请实施例提供的一种训练摘要生成模型的方法的应用场景示意图；

[0045] 图2为本申请实施例提供的一种训练摘要生成模型的方法流程图；

[0046] 图3为本申请实施例提供的一种训练样本的示意图；

[0047] 图4为本申请实施例提供的一种编码器对比学习的逻辑示意图；

[0048] 图5为本申请实施例提供的一种解码器对比学习的逻辑示意图；

[0049] 图6为本申请实施例提供的一种构造错误样本摘要的方法流程图；

[0050] 图7为本申请实施例提供的一种构造错误样本摘要的逻辑示意图；

[0051] 图8为本申请实施例提供的一种构造的错误样本摘要的样例示意图；

[0052] 图9为本申请实施例提供的一种训练摘要生成模型的方法的逻辑示意图；

[0053] 图10为本申请实施例提供的一种摘要生成模型的应用场景示意图；

[0054] 图11为本申请实施例提供的一种训练摘要生成模型的装置的结构框图；

[0055] 图12为本申请实施例提供的一种电子设备的结构示意图；

[0056] 图13为本申请实施例中的另一种电子设备的结构示意图。

具体实施方式

[0057] 为了使本申请的目的、技术方案和优点更加清楚，下面将结合附图对本申请作进一步地详细描述，显然，所描述的实施例仅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它实施例，都属于本申请保护的范围。

[0058] 为了便于本领域技术人员更好地理解本申请的技术方案，下面对本申请涉及的名词进行介绍。

[0059] 摘要生成模型：采用序列到序列架构，也称为编码器和解码器架构，编码器用于将原始文本进行文本编码得到向量，解码器用于从这个向量中提取信息、获取语义，以生成文本摘要。

[0060] 对比学习(Contrastive Learning)：是一种自监督学习方法，用于在没有标签的情况下，通过让模型学习相似的数据或不同的数据来学习数据集的一般特征。

[0061] 自回归语言模型：根据上文内容预测下一个可能跟随的单词，或者根据下文内容预测上一个可能跟随的单词，即自左向右或者自右向左的语言模型任务。

[0062] 下文中所用的词语“示例性”的意思为“用作例子、实施例或说明性”。作为“示例性”所说明的任何实施例不必解释为优于或好于其它实施例。

[0063] 文中的术语“第一”、“第二”仅用于描述目的，而不能理解为明示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此，限定有“第一”、“第二”的特征可以明示或者隐含地包括一个或者更多个该特征，在本申请实施例的描述中，除非另有说明，“多个”的含义是两个或两个以上。

[0064] 本申请实施例涉及人工智能(ArtificialIntelligence,AI)技术领域，基于人工智能中的自然语言处理(Nature Language processing,NLP)技术而设计。

[0065] 人工智能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。

[0066] 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、自动驾驶、智慧交通等几大方向。

[0067] 机器学习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，那么深度学习则是机器学习的核心，是实现机器学习的一种技术。机器学习通常包括深度学习、强化学习、迁移学习、归纳学习等技术，深度学习则包括移动视觉神经网络Mobilenet、卷积神经网络(Convolutional Neural Networks，CNN)、深度置信网络、递归神经网络、自动编码器、生成对抗网络等技术。

[0068] 自然语言处理(Nature Language processing，NLP)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。

[0069] 本申请实施例中的摘要生成模型和自回归语言模型基于机器学习以及自然语言处理技术构建。

[0070] 下面对本申请实施例的设计思想进行简要介绍：

[0071] 相关技术中，为了保证摘要生成模型生成的摘要与对应文本的事实一致性，通常在摘要生成模型的基础上添加额外的处理模块，例如预处理模块或者后处理模块，以分别改进编码过程或者解码过程。预处理模块为模型输入补充特征，改进编码过程，例如额外添加一个编码器来编码关系抽取和命名实体识别的结果、添加一个知识图谱模型抽取文本知识进行编码或者添加一个文本蕴含模型来为摘要模型补充特征。后处理模块主要分为纠正和排序两类方案，改进解码结果。纠正方案将事实一致性摘要看成文本纠错问题，对于生成的摘要结果额外设计一个序列到序列文本纠错模型，来纠正摘要中的事实性错误。排序方案则是先通过束搜索解码或者遍历实体候选替换的方式生成多个摘要结果候选，然后额外设计一个事实一致性排序模型来对摘要候选打分排序，选取得分最高的摘要候选作为最终结果。

[0072] 但是，上述方式需要增加额外的处理模块来实现摘要生成模型的事实一致性，通常的，这些额外的处理模块本身也是参数量较大的深度自然语言处理模型，需要额外的训练数据，且不与摘要生成模型一起端到端训练；这样主要造成了两个问题：一是模块本身的误差会累积到摘要结果当中，例如在一些同时具备预处理和后处理模块的方案中，模型预测链路非常长，所有模块的错误会累积导致摘要结果质量下降；二是额外的模块增加了线上推理的负担，提高了线上推理的时延，一些方案的模块本身甚至和摘要模型相同大小，这样的方案不利于模型部署落地。

[0073] 有鉴于此，本申请实施例提供一种训练摘要生成模型的方法、装置、设备及介质，采用样本文本、正确样本摘要和错误样本摘要构成的训练样本对摘要生成模型的编码子模型进行对比学习，使得正确样本摘要的编码和样本文本的编码更相似，错误样本摘要的编

码和样本文本的编码更不相似;这样,使得训练完成的编码器可以正确编码文本中的事实信息,降低事实错误,从而实现摘要生成模型的事实一致性。本申请实施例不需要增加额外的处理模块,通过改进编码器的训练过程来实现摘要生成模型的事实一致性。

[0074] 以下结合说明书附图对本申请的优选实施例进行说明,应当理解,此处所描述的优选实施例仅用于说明和解释本申请,并不用于限定本申请,并且在不冲突的情况下,本申请实施例及实施例中的特征可以相互组合。

[0075] 如图1所示,其为本申请实施例中的应用场景示意图。该应用场景示意图中包括多个终端设备100和服务器200。终端设备100与服务器200之间可以通过通信网络进行通信。可选地,通信网络可以是有线网络或无线网络。终端设备100与服务器200可以通过有线或无线通信方式进行直接或间接地连接,本申请在此不做限制。

[0076] 在本申请实施例中,终端设备100为用户使用的电子设备,该电子设备包括但不限于个人计算机、手机、平板电脑、笔记本、电子书阅读器、智能语音交互设备、智能家电、车载终端等设备;终端设备100可以安装各种应用,例如浏览器类应用、视频应用、资讯类应用等等。服务器200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(Content Delivery Network,内容分发网络)、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。

[0077] 其中,本申请实施例中的摘要生成模型的训练操作可以由终端设备或者服务器执行;已训练的摘要生成模型可以部署在终端设备上,也可以部署在服务器上。

[0078] 当摘要生成模型部署在终端设备上,而摘要生成模型的训练操作由服务器执行时,终端设备接收服务器发送的已训练的摘要生成模型,把已训练的用户偏好模型安装在本地上,这样,当终端设备获取到需要生成摘要的文本时,可以将文本输入摘要生成模型,获得文本摘要。

[0079] 服务器在对摘要生成模型进行训练时,基于训练样本集对待训练的摘要生成模型进行多轮迭代训练,每个训练样本包括:样本文本,以及相应的正确样本摘要和错误样本摘要;在一轮迭代过程中,执行以下操作:将获取的训练样本输入编码子模型中,获得编码后的文本特征、正确摘要特征和错误摘要特征;基于正确摘要特征和错误摘要特征各自与文本特征的相似度,确定第一损失值,并基于第一损失值对编码子模型进行参数调整;将文本特征输入解码子模型中,获得预测摘要,基于预测摘要和正确样本摘要确定第二损失值,并基于第二损失值对编码子模型和解码子模型中的至少一个进行参数调整。

[0080] 本申请实施例的已训练的摘要生成模型可以用于各种摘要生成场景,例如商品文案生成、评论自动生成、医疗/法律报告生成、新闻摘要生成、信息流文章标题生成、体育战报生成等等,实现文本的冗余信息过滤,提高用户的阅读体验和阅读效率。

[0081] 应当说明的是,图1是对本申请的训练摘要生成模型的方法的应用场景进行示例介绍,实际本申请实施例中的方法可以适用的应用场景并不限于此。本申请实施例可应用于各种场景,包括但不限于云技术、人工智能、智慧交通、辅助驾驶等。

[0082] 下面对本申请实施例的训练摘要生成模型的方法的具体实施方式进行介绍。

[0083] 本申请实施例的摘要生成模型可以是任何序列到序列模型,也可以称为编码器-解码器结构的模型,例如可以是BART(Bidirectional and Auto-Regressive

Transformers,双向自回归变压器)模型等。

[0084] 图2示出了本申请实施例提供的一种训练摘要生成模型的方法的示意图,该方法可以由终端设备执行,也可以由服务器执行。摘要生成模型包括编码子模型和解码子模型,如图2所示,其训练方法可以包括如下步骤:

[0085] 步骤S201,获取训练样本集,每个训练样本包括:样本文本,以及相应的正确样本摘要和错误样本摘要。

[0086] 其中,每个训练样本中的样本文本可以是各种内容的文章,例如:各种内容包括但不限于信息流内容、新闻内容、商品内容、医疗内容、法律内容等等。

[0087] 当样本文本为新闻内容时,其对应的正确样本摘要可以是新闻摘要,该新闻摘要中的事实信息与新闻内容一致,错误样本摘要可以与正确样本摘要的内容相近,但是存在事实性错误。

[0088] 可选地,错误样本摘要可以基于对应的正确样本摘要生成,例如:将正确样本摘要中的一些关键词替换为其他词语,其他词语可以是关键词的相近词语,比如,可以从样本文本中选择关键词的相近词语。本申请下面实施例中将进一步介绍错误样本摘要的构造过程。

[0089] 示例性的,如图3所示,样本文本的内容为“.....A ferocious leopard may have killed 15people in Nepal in a15-month span,.....The police chief suspects that asingle man-eating lropard is responsible for the deaths.....While cases of leopards killing domestic animals are common.....”;正确样本摘要为“A4-year-old boy is the latest victim of a man-eating leopard,.....He suspects one leopard is behind the deaths of 15people in the past 15months.....Leopards are common.....”;错误样本摘要为“A4-year-old boy is the latest victim of a sloth leopard,.....He suspected one leopard is behind the deaths of 15people in the past 20months.....boars are common.....”,其中,错误样本摘要是将正确样本摘要中的一些关键词替换为其他词语后获得的;具体地,将“man-eating”替换为“sloth”,将“suspects”替换为“suspected”,将“15months”中的“15”替换为“20”,将“Leopards”替换为“boars”。

[0090] 步骤S202,基于训练样本集对待训练的摘要生成模型进行多轮迭代训练,其中,在一轮迭代过程中,执行以下步骤S2021-S2023:

[0091] 步骤S2021,将获取的训练样本输入编码子模型中,获得编码后的文本特征、正确摘要特征和错误摘要特征。

[0092] 其中,编码子模型可以是编码器,用于将样本文本、正确样本摘要和错误样本摘要分别进行编码,获得它们各自编码后的隐层表示,将样本文本的隐藏表示作为文本特征,将正确样本摘要的隐层表示作为正确摘要特征,将错误样本摘要的隐层表示作为错误摘要特征,这些特征均可以表示为向量。

[0093] 步骤S2022,基于正确摘要特征和错误摘要特征各自与文本特征的相似度,确定第一损失值,并基于第一损失值对编码子模型进行参数调整。

[0094] 其中,正确摘要特征与文本特征的相似度为向量之间的相似度,例如余弦相似度,错误摘要特征与文本特征的相似度也可以是余弦相似度。此外,向量之间的相似度不限于

余弦相似度，还可以采用其他相似算法，下面以余弦相似度为例进行说明。

[0095] 具体地，在计算正确摘要特征和错误摘要特征各自与文本特征的相似度时，可以先将文本特征、正确摘要特征和错误摘要特征分别进行归一化处理，假设归一化后的文本特征表示为H1，归一化后的正确摘要特征表示为H2，归一化后的错误摘要特征表示为H3，则可以通过以下式(1)的对比学习损失函数计算第一损失值，以实现对比学习：

[0096] $$L_{编码器} = -log\frac{exp(\frac{H_1 \cdot H_2}{\gamma})}{exp(\frac{H_1 \cdot H_3}{\gamma})} \qquad (1)$$

[0097] 其中，γ 是控制对比学习的温度系数。

[0098] 下面以编码子模型为编码器为例，结合图4对编码器的对比学习过程进行示例性介绍。

[0099] 如图4所示，将样本文本、正确样本摘要和错误样本摘要输入编码器中，获得编码后的文本特征、正确摘要特征和错误摘要特征，然后，计算文本特征和正确摘要特征的相似度，以及计算文本特征和错误摘要特征的相似度，基于这两个相似度进行对比学习，即将这两个相似度代入对比学习损失函数，获得第一损失值，基于第一损失值对编码器进行参数调整。

[0100] 本申请实施例中，基于样本文本的正确样本摘要和错误样本摘要对编码子模型进行对比学习，使得编码子模型能够在给定样本文本的情况下，使得正确样本摘要的编码和样本文本的编码更相似，错误样本摘要的编码和样本文本的编码更不相似。这样，使得训练后的编码子模型可以正确编码文本中的事实信息，降低事实错误。

[0101] 步骤S2023，将文本特征输入解码子模型中，获得预测摘要，基于预测摘要和正确样本摘要确定第二损失值，并基于第二损失值对编码子模型和解码子模型中的至少一个进行参数调整。

[0102] 其中，解码子模型可以是解码器，在获得解码器输出的预测摘要后，可以采用解码子模型的损失函数(例如交叉熵损失函数等)计算预测摘要和正确样本摘要的第二损失值，基于第二损失值对解码子模型进行参数调整，同时还可以对编码子模型进行参数调整。

[0103] 步骤S203，确定满足预设的收敛条件时，获得已训练的摘要生成模型。

[0104] 例如，预设的收敛条件可以是第一损失值小于第一设定值，且第二损失值小于第二设定值；其中，第一设定值和第二设定值可以根据需要设置，在此不作限定。

[0105] 本申请实施例的方案中，采用样本文本、正确样本摘要和错误样本摘要构成的训练样本对摘要生成模型的编码子模型进行对比学习，使得训练后的编码子模型在给定文本的情况下，使得正确摘要的编码和文本的编码更相似，错误摘要的编码和文本的编码更不相似。这样，使得编码器可以正确编码文本中的事实信息，降低事实错误，进而使训练后的解码器输出与文本具有事实一致性的文本摘要。

[0106] 在一些实施例中，为了进一步提高摘要生成模型的事实一致性，可以对解码器也进行对比学习，使得预测摘要与正确摘要的损失最小，预测摘要与错误摘要的损失最大。

[0107] 此时，上述步骤S2023中基于预测摘要和正确样本摘要确定第二损失值，可以包括以下步骤A1-A2：

[0108] 步骤A1、基于预测摘要和正确样本摘要确定正确预测概率值，以及基于预测摘要

和错误样本摘要确定错误预测概率值。

[0109] 其中，可以通过解码子模型的损失函数计算预测摘要和正确样本摘要的正确损失值，并将正确损失值转换为正确预测概率值，可以理解的是，正确损失值越小，正确预测概率值越大；同样地，通过解码器的损失函数计算预测摘要和错误样本摘要的错误损失值，并将错误损失值转换为错误预测概率值，错误损失值越大，错误预测概率值越小。

[0110] 步骤A2、基于正确预测概率值和错误预测概率值，确定第二损失值。

[0111] 该步骤中，可以将正确预测概率值和错误预测概率值代入解码子模型的对比学习损失函数，计算得到第二损失值。

[0112] 下面以解码子模型为解码器为例，结合图4对解码器的对比学习过程进行示例性介绍。

[0113] 如图5所示，将编码后的文本特征输入解码器后，获得预测摘要，然后基于预测摘要和错误样本摘要计算错误预测概率值，基于预测摘要和正确样本摘要计算正确预测概率值。通过设置以下式(2)的对比学习损失函数实现对比学习，使得训练后的解码器解码出正确摘要的概率最大化，解码出错误摘要的概率最小化。

[0114] $L_{解码器}=\max(P1-P2+\eta,0)$ (2)

[0115] 其中，P1是解码器的正确预测概率值，P2是解码器的错误预测概率值，η是对比学习间隔参数。通过这样的训练能够使得解码器具备对事实错误的感知能力，能够区分易错的事实不一致结果，提高生成具有事实一致性的正确摘要的概率。

[0116] 本申请下面实施例对上述训练样本中的错误样本摘要的构造过程进行介绍。

[0117] 如图6所示，每个训练样本中的错误样本摘要的生成过程可以包括以下步骤S601-S603：

[0118] 步骤S601，从训练样本中的正确样本摘要中查找关键词集合。

[0119] 该步骤中，可以将正确样本摘要进行多次迭代裁剪，每次裁剪获得一个摘要片段，最终获得多个摘要片段，然后从多个摘要片段中选择最合适的摘要片段，并将选择的摘要片段中的多个词语作为关键词集合。

[0120] 可选地，步骤S601可以包括以下步骤B1-B4：

[0121] 步骤B1、从训练样本中的样本文本中，获得正确样本摘要中的至少一个摘要语句各自对应的上下文语句。

[0122] 其中，正确样本摘要可以包括一个或多个摘要语句，每个语句可以对应样本文本中的一个上下文语句。具体地，可以基于文本摘要评测指标(Recall-Oriented Understudy for Gisting Evaluation,ROUGE)从样本文本中查找正确样本摘要中的每个语句对应的上下文语句，例如：对于正确样本摘要中的每个摘要语句，计算样本文本中的各个文本语句分别与该摘要语句的ROUGE指标，将ROUGE指标的数值最大的一个文本语句作为该摘要语句的上下文语句。

[0123] 步骤B2、针对至少一个摘要语句，迭代执行多次以下操作：将至少一个摘要语句分别进行裁剪处理，获得至少一个摘要语句各自对应的词语片段，并将至少一个词语片段拼接成摘要片段，以及将每个词语片段作为新的一个摘要语句；

[0124] 其中，裁剪处理可以是随机裁剪，例如去掉所有停用词、随机删除一些词组和从句等简单规则处理。将正确样本摘要的至少一个摘要语句分别进行多次迭代裁剪处理，每次

裁剪处理完，将获得的至少一个裁剪语句合并成一个摘要片段。如果正确样本摘要包括一个摘要语句，则每次裁剪处理完可以获得一个摘要片段。裁剪处理的次数可以根据具体情况确定。

[0125] 示例性的，假设正确样本摘要包括一个摘要语句“读完《富兰克林自传》，我懂得了优秀的人都是怎样过好这一生的”，将这个摘要语句进行第一次裁剪后得到一个摘要片段“《富兰克林自传》、懂得、优秀的人、过好一生”，将这个裁剪后的摘要片段再次进行裁剪处理，获得第二个摘要片段“《富兰克林自传》、优秀的人”，以此类推。

[0126] 步骤B3、针对获得的多个摘要片段，分别执行以下操作：基于一个摘要片段以及正确样本摘要对应的至少一个上下文语句，确定一个摘要片段的评估值。

[0127] 其中，一个摘要片段的评估值可以包括相关性数值和压缩率数值，相关性数值表示摘要片段与各个上下文语句的相关程度，压缩率数值表示摘要片段的长度。

[0128] 可选地，上述步骤B3中基于一个摘要片段以及正确样本摘要对应的至少一个上下文语句，确定一个摘要片段的评估值，可以包括以下步骤B31-B34：

[0129] 步骤B31、针对至少一个上下文语句，分别执行以下操作：将一个上下文语句和一个摘要片段，输入自回归语言模型，获得第一语言预测概率。

[0130] 其中，将一个文本输入自回归语言模型后，可以获得这个文本中的每个词的预测概率，然后将所有词的预测概率分别取对数后再求和，就是这个文本的语言预测概率。

[0131] 为了确定一个摘要片段分别与至少一个上下文语句的相关性，将该一个摘要片段与一个上下语句拼接后，输入自回归语言模型，获得第一语言预测概率，将这个第一语言预测概作为该一个摘要片段与该一个上下文语句的相似性数值。

[0132] 步骤B32、基于获得的至少一个第一语言预测概率，获得一个摘要片段的相关性数值。

[0133] 如果正确样本摘要对应一个上下语句，则可以获得一个第一语言预测概率，将这个第一语言预测概率作为一个摘要片段的相关性数值；如果正确样本摘要对应多个上下语句，则可以获得多个第一语言预测概率，可以将多个第一语言预测概率取平均值，将该平均值作为一个摘要片段的相关性数值。

[0134] 步骤B33、将一个摘要片段输入自回归语言模型，获得第二语言预测概率，并基于第二语言预测概率获得一个摘要片段的压缩率数值。

[0135] 其中，在给定文本的情况下，最小化文本片段的语言预测概率等价于最大化压缩率。因此，可以将第二语言预测概率取反，获得摘要片段的压缩率数值；例如：第二语言预测概率为0.4，取反后为1-0.4＝0.6，即压缩率数值为0.6。

[0136] 步骤B34、将一个摘要片段的相关性数值和压缩率数值，作为一个摘要片段的评估值。

[0137] 步骤B4、基于多个摘要片段各自的评估值，从多个摘要片段中选择目标摘要片段，并将目标摘要片段中的多个词语组成关键词集合。

[0138] 该步骤中，可以从多个摘要片段中选择评估值满足条件的目标摘要片段。例如：当一个摘要片段的评估值包括相关性数值和压缩率数值时，可以在压缩率数值满足预设条件的情况下，最大化相关性数值。

[0139] 可选地，上述步骤S6014中基于多个摘要片段各自的评估值，从多个摘要片段中选

择目标摘要片段，可以包括以下步骤C1-C2：

[0140] C1、从多个摘要片段中，选择压缩率数值满足预设条件的多个候选摘要片段。

[0141] 其中，预设条件可以是：将多个摘要片段按照压缩率数值从大到小排列后，排在前k个的摘要片段。

[0142] C2、将多个候选摘要片段中，相关性数值最大的候选摘要片段作为目标摘要片段。

[0143] 考虑到摘要片段的压缩率越高，包含的关键词越少，对正确样本摘要的扰动越小，因此在选择目标摘要片段(即需要替换的关键词集合)时，应当是在压缩率尽可能高的情况下最大化相关性，这样，构造的错误样本摘要和正确样本摘要越相似，越难以区分，这样的错误样本摘要能够为摘要生成模型的学习提供更大的价值。

[0144] 基于此，从多个摘要片段中选择压缩率数值最大的前k个摘要片段作为候选摘要片段，再从这k个候选摘要片段中选择相关性数值最大的作为目标摘要片段。这样，可以使得目标摘要片段在压缩率尽可能高的情况下，相关性最大化。

[0145] 步骤S602，针对关键词集合中的多个关键词，分别执行以下操作：若一个关键词的替换概率达到预设概率，则将一个关键词替换为相应的替换词；其中，替换词是从训练样本中的样本文本中选择的。

[0146] 其中，针对每个关键词，可以随机生成一个0-1的数值，如果生成数值达到预设数值，则需要进行替换，否则不需要进行替换；例如：预设数值为0.1或者0.2，在此不做限定。

[0147] 在确定一个关键词的替换概率达到预设概率后，可以从样本文本中选择与这个关键词的词向量最相似的替换词，并将这个关键词替换为该替换词。

[0148] 步骤S603，将替换后的正确样本摘要作为训练样本中的错误样本摘要。

[0149] 下面结合图7对本申请上述实施例中的错误样本摘要的构造过程进行说明。

[0150] 如图7所示，基于一个样本文本和对应的正确样本摘要构造错误样本摘要时，从样本文本中查找正确样本摘要对应的上下文语句(一个或多个)，同时，基于正确样本摘要生成多个摘要片段；针对每个摘要片段，将该摘要片段与每个上下文语句输入自回归语言模型，获得该摘要片段的评估值；基于多个摘要片段各自的评估值，从多个摘要片段中选择目标摘要片段，获得关键词集合；将关键词集合中的每个关键词按照替换概率进行替换，具体可以替换为样本文本中与关键词的词向量最相近的词语，将正确样本摘要进行替换后，获得错误样本摘要。

[0151] 示例性的，如图8所示，采用本申请上述实施例中的错误样本摘要的构造方法，构造的错误样本摘要的样例包括各种错误类型，图8中的正确摘要即正确样本摘要，负例摘要即错误样本摘要；例如：样例1为名词错误，样例2为数字错误，样例3为形容词错误，样例4为实体错误，样例5为词组错误，样例6为动词错误等等。可以看到，构造出的错误样本摘要具有多样化的特点，覆盖了多种类型错误，比人工设计错误样本摘要更全面高效。并且，本申请实施例可以构造出与正确样本摘要十分相似的错误样本摘要，构造的错误样本摘要与正确样本摘要越相似，越难以区分，这样，错误样本摘要能够为摘要生成模型的学习提供更大的价值。

[0152] 图9示出了本申请实施例的训练摘要生成模型的方法的逻辑示意图。

[0153] 如图9所示，在获得一个样本文本和对应的正确样本摘要后，基于样本文本和正确样本摘要构造错误样本摘要，将样本文本、正确样本摘要和错误样本摘要作为一个训练样

本;基于获得的训练样本集对编码器进行对比学习,以及对解码器进行对比学习,获得训练后的摘要生成模型;将训练后的摘要生成模型用于线上使用时,将线上文本输入该摘要生成模型,获得线上文本对应的文本摘要。

[0154] 本申请实施例的训练摘要生成模型的方法可以应用各种摘要生成场景,例如:商品文案生成、评论自动生成、医疗/法律报告生成、新闻摘要生成、信息流文章标题生成、体育战报生成等等,实现文本的冗余信息过滤,提高用户的阅读体验和阅读效率。

[0155] 示例性的,以信息流文章标题生成为例,如图10所示,将信息流文章“一个理性的文字治愈师,希望你能够在我的文字里,找到一点对于生活的慰藉.....那么,下面就和大家分享四本我觉得对我有较大帮助的书.....”输入训练后的摘要生成模型后,获得信息流文章标题“推荐四本好书,很适合大家阅读:丰富自己,永远不会吃亏”。

[0156] 下面对采用本申请实施例的方法(包括只进行编码器对比学习、只进行解码器对比学习和编码解码组合对比学习)训练得到的摘要生成模型,与未改进的摘要生成模型(基础方案),在常用的两个公开数据集上进行对比,由下表1和表2可知,采用本申请实施例的方法训练得到的摘要生成模型在四个常用的事实一致性指标上具均得到了提升。下表1展示了在公开数据集1上的对比结果:

[0157] 表1

[0158]

方案	问答事实指标1	问答事实指标2	三元组事实指标1	三元组事实指标2
基础方案	70.15	30.68	54.89	41.94

[0159]

编码器对比学习	72.28(+2.13)	30.67(-0.01)	57.05(+2.16)	46.52(+4.58)
解码器对比学习	73.22(+3.07)	30.79(+0.11)	58.19(+3.30)	48.36(+6.42)
编码解码组合对比学习	73.87(+3.23)	30.73(+0.05)	58.18(+3.29)	49.72(+7.78)

[0160] 下表2是公开数据集2上的对比结果:

[0161] 表2

[0162]

方案	问答事实指标1	问答事实指标2	三元组事实指标1	三元组事实指标2
基础方案	13.19	15.57	2.75	2.71
编码器对比学习	13.53(+0.34)	16.64(+1.07)	2.92(+0.17)	3.59(+0.88)
解码器对比学习	13.27(+0.05)	16.73(+1.16)	3.03(+0.28)	3.31(+0.60)
编码解码组合对比学习	13.48(+0.29)	16.86(+1.29)	3.31(+0.56)	4.34(+1.63)

[0163] 可以看到，在两个公开数据集上，编码器和解码器上的对比学习优化方案几乎在所有指标上均优于基础方案，且编码解码组合优化方案可以进一步提高摘要的事实一致性。

[0164] 此外，将本申请实施例的编码器对比学习方案和现有技术(在摘要生成模型的基础上，添加预处理模块进行编码优化或者添加后处理模块进行解码优化)进行对比，在可公平比较的公开数据集3上的对比结果如下表3所示：

[0165] 表3

[0166]

方案	问答事实指标1	问答事实指标2	三元组事实指标1	三元组事实指标2
现有编码优化方案	10.17	9.32	1.08	0.78
现有解码优化方案	12.15(+1.98)	15.02(+5.7)	2.97(+1.89)	3.73(+2.95)

[0167]

本申请的编码方案	12.11(+1.94)	15.83(+6.51)	3.05(+1.97)	1.82(+1.04)

[0168] 可以看到，本申请实施例的方案和现有解码方案都优于现有编码方案，但是本申请实施例的编码方案(即编码器对比学习)与现有解码方案相比，不需要引入额外的处理模块，不会增加线上推理负担。

[0169] 与本申请上述方法实施例基于同一发明构思，本申请实施例中还提供了一种训练摘要生成模型的装置，该装置解决问题的原理与上述实施例的方法相似，因此该装置的实施可以参见上述方法的实施，重复之处不再赘述。

[0170] 如图11所示，本申请实施例提供了一种训练摘要生成模型的装置，该摘要生成模型包括编码子模型和解码子模型，装置包括：

[0171] 训练模块111，用于基于训练样本集对待训练的摘要生成模型进行多轮迭代训练，输出已训练的摘要生成模型，每个训练样本包括：样本文本，以及相应的正确样本摘要和错误样本摘要；其中，在一轮迭代过程中，执行以下操作：

[0172] 将获取的训练样本输入编码子模型中，获得编码后的文本特征、正确摘要特征和

错误摘要特征；

[0173] 基于正确摘要特征和错误摘要特征各自与文本特征的相似度，确定第一损失值，并基于第一损失值对编码子模型进行参数调整；

[0174] 将文本特征输入解码子模型中，获得预测摘要，基于预测摘要和正确样本摘要确定第二损失值，并基于第二损失值对编码子模型和解码子模型中的至少一个进行参数调整。

[0175] 本申请实施例的方案中，采用样本文本、正确样本摘要和错误样本摘要构成的训练样本对摘要生成模型的编码子模型进行对比学习，使得训练后的编码子模型在给定文本的情况下，使得正确摘要的编码和文本的编码更相似，错误摘要的编码和文本的编码更不相似。这样，使得编码器可以正确编码文本中的事实信息，降低事实错误，进而使训练后的解码器输出与文本具有事实一致性的文本摘要。

[0176] 在一种可能的实施例中，基于预测摘要和正确样本摘要确定第二损失值时，训练模块还用于：

[0177] 基于预测摘要和正确样本摘要确定正确预测概率值，以及基于预测摘要和错误样本摘要确定错误预测概率值；

[0178] 基于正确预测概率值和错误预测概率值，确定第二损失值。

[0179] 在一种可能的实施例中，还包括获取模块，用于通过如下方式获得每个训练样本中的错误样本摘要：

[0180] 从训练样本中的正确样本摘要中查找关键词集合；

[0181] 针对关键词集合中的多个关键词，分别执行以下操作：若一个关键词的替换概率达到预设概率，则将一个关键词替换为相应的替换词；其中，替换词是从训练样本中的样本文本中选择的；

[0182] 将替换后的正确样本摘要作为训练样本中的错误样本摘要。

[0183] 在一种可能的实施例中，从训练样本中的正确样本摘要中查找关键词集合时，获取模块还用于：

[0184] 从训练样本中的样本文本中，获得正确样本摘要中的至少一个摘要语句各自对应的上下文语句；

[0185] 针对至少一个摘要语句，迭代执行多次以下操作：将至少一个摘要语句分别进行裁剪处理，获得至少一个摘要语句各自对应的词语片段，并将至少一个词语片段拼接成摘要片段，以及将每个词语片段作为新的一个摘要语句；

[0186] 针对获得的多个摘要片段，分别执行以下操作：基于一个摘要片段以及正确样本摘要对应的至少一个上下文语句，确定一个摘要片段的评估值；

[0187] 基于多个摘要片段各自的评估值，从多个摘要片段中选择目标摘要片段，并将目标摘要片段中的多个词语组成关键词集合。

[0188] 在一种可能的实施例中，基于一个摘要片段以及正确样本摘要对应的至少一个上下文语句，确定一个摘要片段的评估值时，获取模块还用于：

[0189] 针对至少一个上下文语句，分别执行以下操作：将一个上下文语句和一个摘要片段，输入自回归语言模型，获得第一语言预测概率；

[0190] 基于获得的至少一个第一语言预测概率，获得一个摘要片段的相关性数值；

[0191] 将一个摘要片段输入自回归语言模型，获得第二语言预测概率，并基于第二语言预测概率获得一个摘要片段的压缩率数值；

[0192] 将一个摘要片段的相关性数值和压缩率数值，作为一个摘要片段的评估值。

[0193] 在一种可能的实施例中，基于多个摘要片段各自的评估值，从多个摘要片段中选择目标摘要片段时，获取模块还用于：

[0194] 从多个摘要片段中，选择压缩率数值满足预设条件的多个候选摘要片段；

[0195] 将多个候选摘要片段中，相关性数值最大的候选摘要片段作为目标摘要片段。

[0196] 为了描述的方便，以上各部分按照功能划分为各模块(或单元)分别描述。当然，在实施本申请时可以把各模块(或单元)的功能在同一个或多个软件或硬件中实现。

[0197] 关于上述实施例中的装置，其中各个模块的具体执行方式已经在有关该方法的实施例中进行了详细描述，此处将不做详细阐述说明。

[0198] 在介绍了本申请示例性实施方式的训练摘要生成模型的方法和装置之后，接下来，介绍根据本申请的另一示例性实施方式的电子设备。

[0199] 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、或计算为了描述的方便，以上各部分按照功能划分为各模块分别描述。当然，在实施本申请时可以把各模块的功能在同一个或多个软件或硬件中实现。

[0200] 所属技术领域的技术人员能够理解，本申请的各个方面可以实现为系统、方法或程序产品。因此，本申请的各个方面可以具体实现为以下形式，即：完全的硬件实施方式、完全的软件实施方式(包括固件、微代码等)，或硬件和软件方面结合的实施方式，这里可以统称为“电路”、“模块”或“系统”。

[0201] 与本申请上述方法实施例基于同一发明构思，本申请实施例中还提供了一种电子设备，该电子设备解决问题的原理与上述实施例的方法相似，因此该电子设备的实施可以参见上述方法的实施，重复之处不再赘述。

[0202] 参阅图12所示，电子设备1200可以至少包括处理器1201、以及存储器1202。其中，存储器1202存储有程序代码，当程序代码被处理器1201执行时，使得处理器1201执行上述任意一种训练摘要生成模型的方法中的步骤。

[0203] 在一些可能的实施方式中，根据本申请的电子设备可以至少包括至少一个处理器、以及至少一个存储器。其中，存储器存储有程序代码，当程序代码被处理器执行时，使得处理器执行本说明书上述描述的根据本申请各种示例性实施方式的训练摘要生成模型的方法中的步骤。例如，处理器可以执行如图2中所示的步骤。

[0204] 在示例性实施例中，本申请还提供了一种包括程序代码的存储介质，例如包括程序代码的存储器1202，上述程序代码可由电子设备1200的处理器1201执行以完成上述训练摘要生成模型的方法。可选地，存储介质可以是非临时性计算机可读存储介质，例如，非临时性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。

[0205] 下面参照图13来描述根据本申请的这种实施方式的电子设备130。图13的电子设备130仅仅是一个示例，不应对本申请实施例的功能和使用范围带来任何限制。

[0206] 如图13，电子设备130以通用电子设备的形式表现。电子设备130的组件可以包括但不限于：上述至少一个处理单元131、上述至少一个存储单元132、连接不同系统组件(包

括存储单元132和处理单元131)的总线133。

[0207] 总线133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、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。

[0208] 存储单元132可以包括易失性存储器形式的可读介质,例如随机存取存储器(RAM)1321和/或高速缓存存储器1322,还可以进一步包括只读存储器(ROM)1323。

[0209] 存储单元132还可以包括具有一组(至少一个)程序模块1324的程序/实用工具1325,这样的程序模块1324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。

[0210] 电子设备130也可以与一个或多个外部设备134(例如键盘、指向设备等)通信,还可与一个或者多个使得用户能与电子设备130交互的设备通信,和/或与使得该电子设备130能与一个或多个其它电子设备进行通信的任何设备(例如路由器、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135进行。并且,电子设备130还可以通过网络适配器136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136通过总线133与用于电子设备130的其它模块通信。应当理解,尽管图中未示出,可以结合电子设备130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、冗余处理器、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。

[0211] 与上述方法实施例基于同一发明构思,本申请实施例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任意一种训练摘要生成模型的方法的步骤。

[0212] 程序产品可以采用一个或多个可读介质的任意组合。可读介质可以是可读信号介质或者可读存储介质。可读存储介质例如可以是但不限于电、磁、光、电磁、红外线、或半导体的系统、装置或器件,或者任意以上的组合。可读存储介质的更具体的例子(非穷举的列表)包括:具有一个或多个导线的电连接、便携式盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、光纤、便携式紧凑盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。

[0213] 本申请的实施方式的程序产品可以采用便携式紧凑盘只读存储器(CD-ROM)并包括程序代码,并可以在计算装置上运行。然而,本申请的程序产品不限于此,在本文件中,可读存储介质可以是任何包含或存储程序的有形介质,该程序可以被命令执行系统、装置或者器件使用或者与其结合使用。

[0214] 可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号,其中承载了可读程序代码。这种传播的数据信号可以采用多种形式,包括但不限于电磁信号、光信号或上述的任意合适的组合。可读信号介质还可以是可读存储介质以外的任何可读介质,该可读介质可以发送、传播或者传输用于由命令执行系统、装置或者器件使用或者与其结合使用的程序。

[0215] 可读介质上包含的程序代码可以用任何适当的介质传输,包括但不限于无线、有线、光缆、RF等等,或者上述的任意合适的组合。

[0216] 尽管已描述了本申请的优选实施例,但本领域内的技术人员一旦得知了基本创造性概念,则可对这些实施例做出另外的变更和修改。所以,所附权利要求意欲解释为包括优选实施例以及落入本申请范围的所有变更和修改。

[0217] 显然,本领域的技术人员可以对本申请进行各种改动和变型而不脱离本申请的精神和范围。这样,倘若本申请的这些修改和变型属于本申请权利要求及其等同技术的范围之内,则本申请也意图包含这些改动和变型在内。

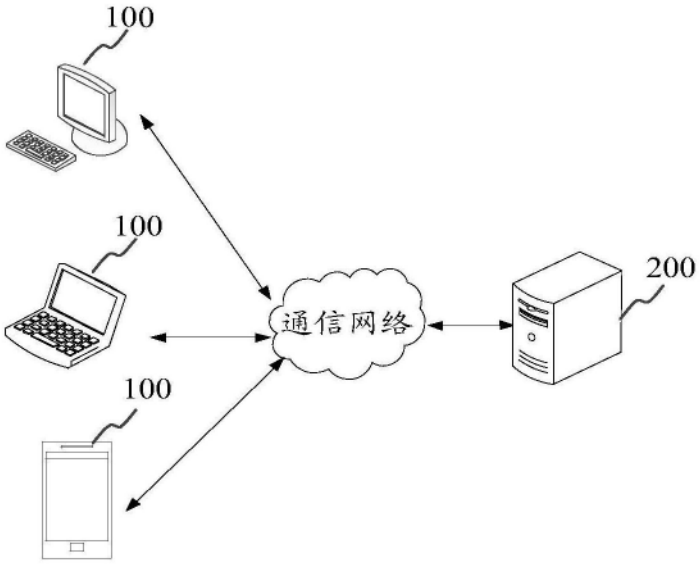


图1

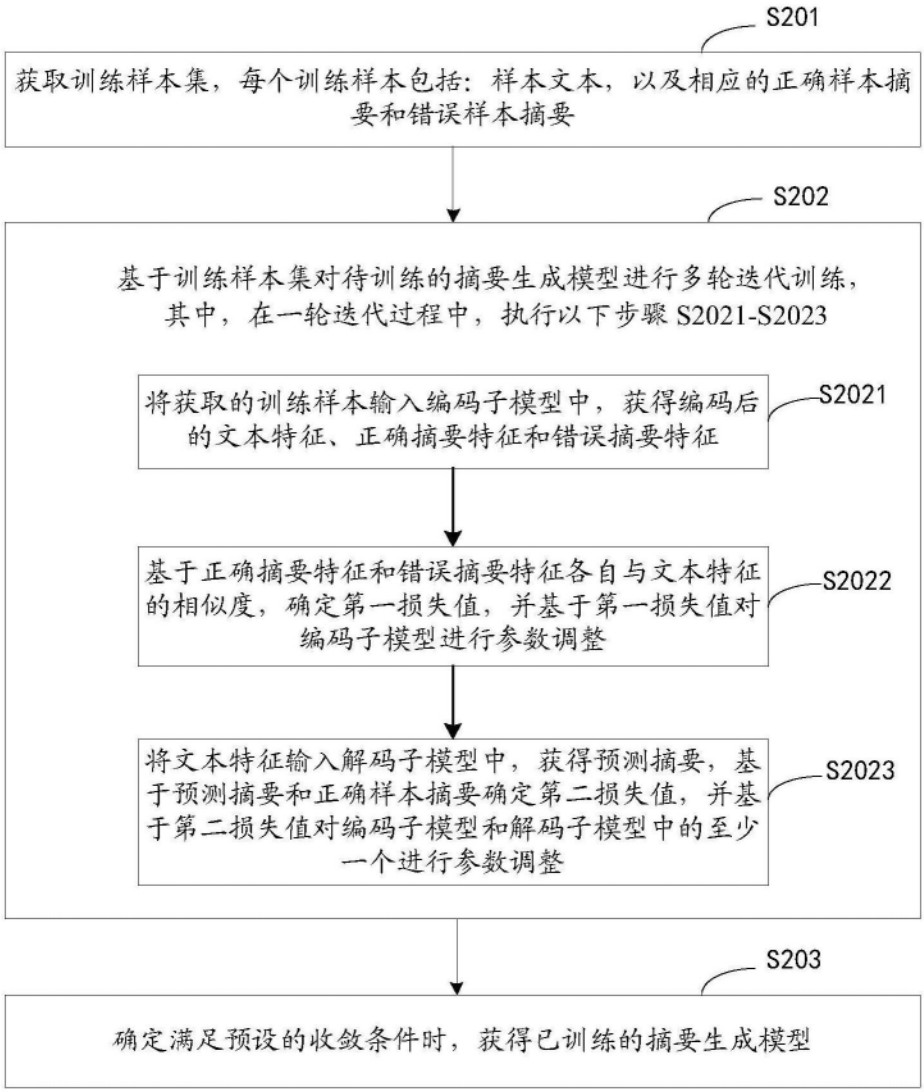

S201
获取训练样本集，每个训练样本包括：样本文本，以及相应的正确样本摘要和错误样本摘要
S202
基于训练样本集对待训练的摘要生成模型进行多轮迭代训练，其中，在一轮迭代过程中，执行以下步骤 S2021-S2023
将获取的训练样本输入编码子模型中，获得编码后的文本特征、正确摘要特征和错误摘要特征
S2021
基于正确摘要特征和错误摘要特征各自与文本特征的相似度，确定第一损失值，并基于第一损失值对编码子模型进行参数调整
S2022
将文本特征输入解码子模型中，获得预测摘要，基于预测摘要和正确样本摘要确定第二损失值，并基于第二损失值对编码子模型和解码子模型中的至少一个进行参数调整
S2023
S203
确定满足预设的收敛条件时，获得已训练的摘要生成模型

图2

样本文本	.....A ferocious leopard may have killed 15 people in Nepal in a 15-month span, ·····The police chief suspects that a single man-eating leopard is responsible for the deaths. ·····The district administration has announced a Rs 25,000 (about $300) reward to anyone who captures or kills the leopard·····Leopards are common in the low mountain areas, as compared to the high Himalayas, across the country. While cases of leopards killing domestic animals are common·····
正确样本摘要	A 4-year-old boy is the latest victim of a **man-eating** leopard, a local police chief says . He **suspects** one leopard is behind the deaths of 15 people in the past **15** months . A reward has been offered to anyone who captures or kills the man-eating creature . **Leopards** are common in low mountain areas of Nepal but usually eat wild prey like deer .
错误样本摘要	A 4-year-old boy is the latest victim of a **sloth** leopard, a local police chief says . He **suspected** one leopard is behind the deaths of 15 people in the past **20** months . A reward has been offered to than who captures or kills the man-eating creature . **boars** are common in low mountain areas of Nepal but usually eat wild prey like deer .

图3

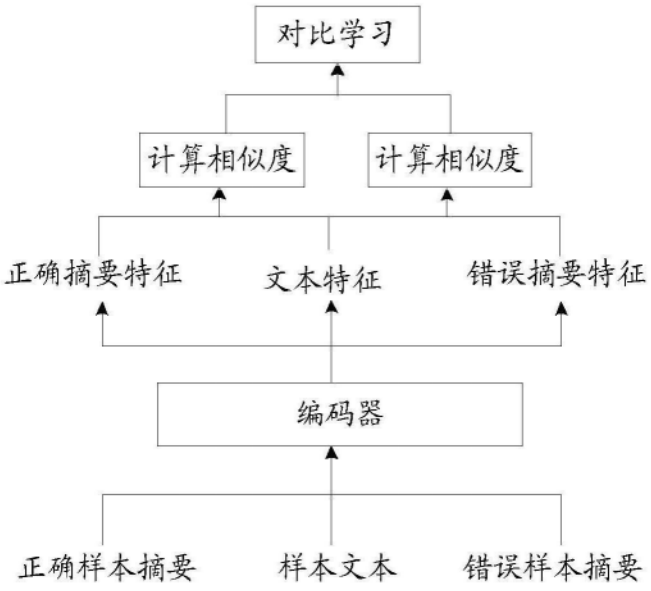


图4

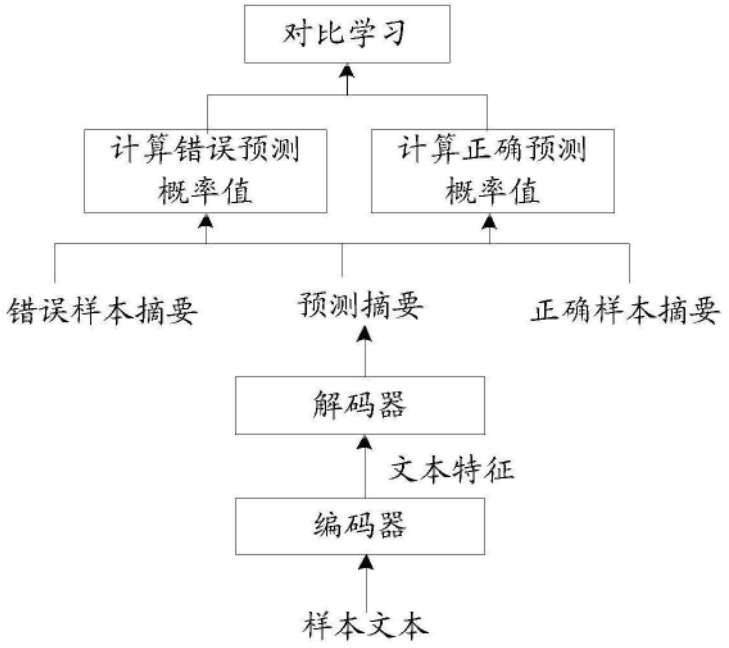

对比学习
计算错误预测概率值
计算正确预测概率值
错误样本摘要
预测摘要
正确样本摘要
解码器
文本特征
编码器
样本文本

图5

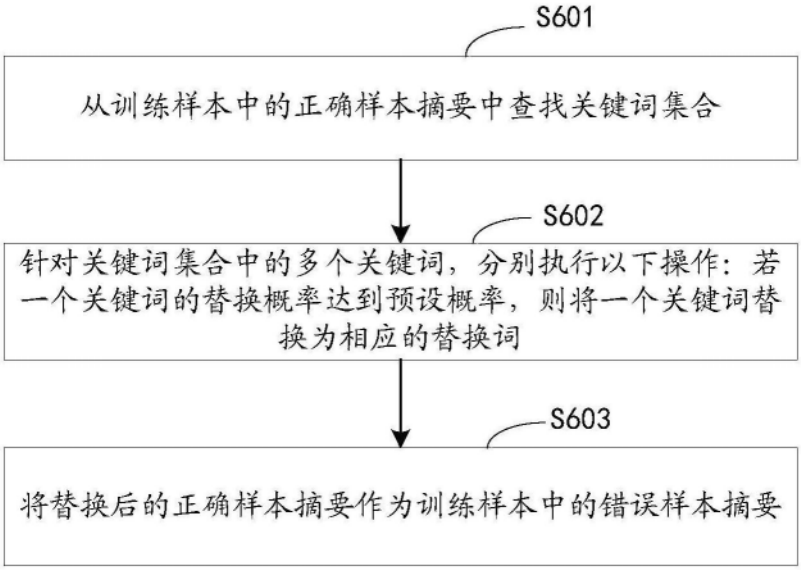

S601
从训练样本中的正确样本摘要中查找关键词集合
S602
针对关键词集合中的多个关键词，分别执行以下操作：若一个关键词的替换概率达到预设概率，则将一个关键词替换为相应的替换词
S603
将替换后的正确样本摘要作为训练样本中的错误样本摘要

图6

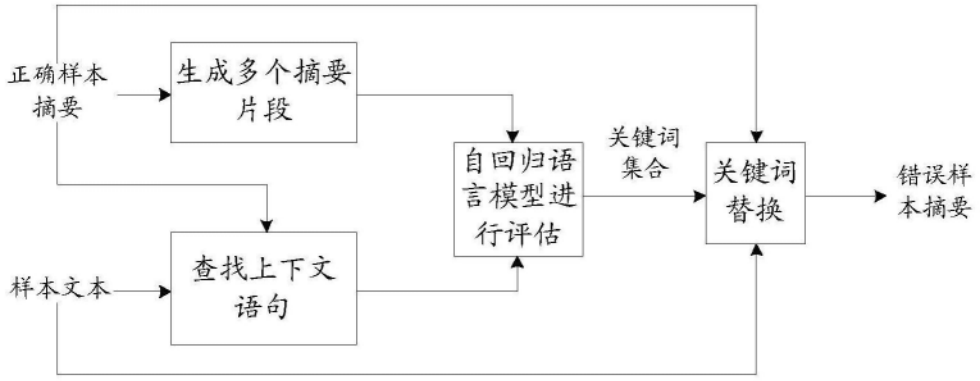

正确样本
摘要
生成多个摘要
片段
样本文本
查找上下文
语句
自回归语
言模型进
行评估
关键词
集合
关键词
替换
错误样
本摘要

图7

样例1 名词错误	正确摘要	Syrian official: Obama climbed to the top of the tree, "doesn't know how to get down" Obama sends a letter to the heads of the House and Senate.Obama to seek congressional approval on military action against Syria . Aim is to determine whether CW were used, not by whom, says U.N. spokesman.
	构造错误摘要	Syrian official: Obama climbed to the top of the tree, "doesn't know how to get down" Obama sends a letter to the heads of the House and Senate.Obama mounting seek congressional request on force action against Syria . Aim is to determine whether showtime were used, not by whom, says U.N. statement.
	替换内容	'approval --> request', 'military --> force', 'CW --> showtime', 'spokesman --> statement'
样例2 数字错误	正确摘要	The 15 new cardinals will be installed on February 14.They come from countries such as Myanmar and Tonga.No Americans made the list this time or the previous time in Francis' papacy.
	构造错误摘要	The 14 new cardinals will be installed on February 19 . They come from countries such as cambodian and Tonga . No Americans made the list this time or the previous coming in Francis' papacy .
	替换内容	'15 --> 14', '14 --> 19', 'Myanmar --> cambodian', 'time --> coming'
样例3 形容词 错误	正确摘要	Protesters converge on Hong Kong's Victoria Park for a candlelight vigil.It is the 22nd anniversary of the bloody crackdown on pro-democracy protesters.The vigil is held after recent efforts to quash anti-government demonstrations.
	构造错误摘要	Protesters converge on malaysia Kong's Victoria Park for a candlelight vigil.It is the 22nd anniversary of the bloody crackdown on pro-democracy protesters.The vigil is in after recent efforts to quash pro-democracy demonstrations.
	替换内容	'held --> in', 'anti-government --> pro-democracy'
样例4 实体错误	正确摘要	Man Haron Monis had history of "infatuation with extremism and mental instability," Abbott says . Hostage-taker was granted political asylum in Australia in 2001; was on bail for violent offending . His former lawyer says he campaigned against "the victimization of Muslims and Islamists" On Monis' apparent website, there is a pledge of allegiance to ISIS .
	构造错误摘要	Man Haron kalantar had history of "infatuation with radicalism and mental instability," Abbott says . Hostage-taker was granted political asylum in australian in 2001; was on bail for violent offending . His former lawyer says he campaigned against "the deviance of Mus lims and Islamists" On Monis' obvious website, there is a pledge of allegiance to ISIS .
	替换内容	'Monis --> kalantar', 'extremism --> radicalism', 'Australia --> australian', 'victimization --> deviance', 'apparent --> obvious'
样例5 词组错误	正确摘要	Larry Norman was Christian rock musician before genre existed . His first solo album, "Upon This Rock," came out in late 1969 . Norman's fans include U2, Guns N' Roses and Bob Dylan . More than 300 versions of Norman's songs have been recorded by other artists .
	构造错误摘要	Larry Norman was Christian rock vocalist before genre existed . His first solo album, "Upon This Rock," came out in early 1966 . Norman's crowd include U2, Guns N' Roses and Bob mccartney . More than 300 versions of Norman's songs have been recorded by other artists .
	替换内容	'musician --> vocalist', 'late --> early', '1969 --> 1966', 'fans --> crowd', 'Dylan --> mccartney'
样例6 动词错误	正确摘要	Customer booked an international flight through Cheapoair.com . Ticket for one leg of her flight was not accepted, and she had to buy a new one . The online agency did not know why her ticket was rejected . Troubleshooter contacted Cheapoair, and it refunded the customer's money .
	构造错误摘要	Customer booked an international flight But Cheapoair.com . Ticket for one leg of her flight was not accepted, and she had to buy a new one . The online agency did not know why her ticket was argued . Troubleshooter consulted Cheapoair, and it refunded the customer's money .
	替换内容	'rejected --> argued', 'contacted --> consulted'

图8

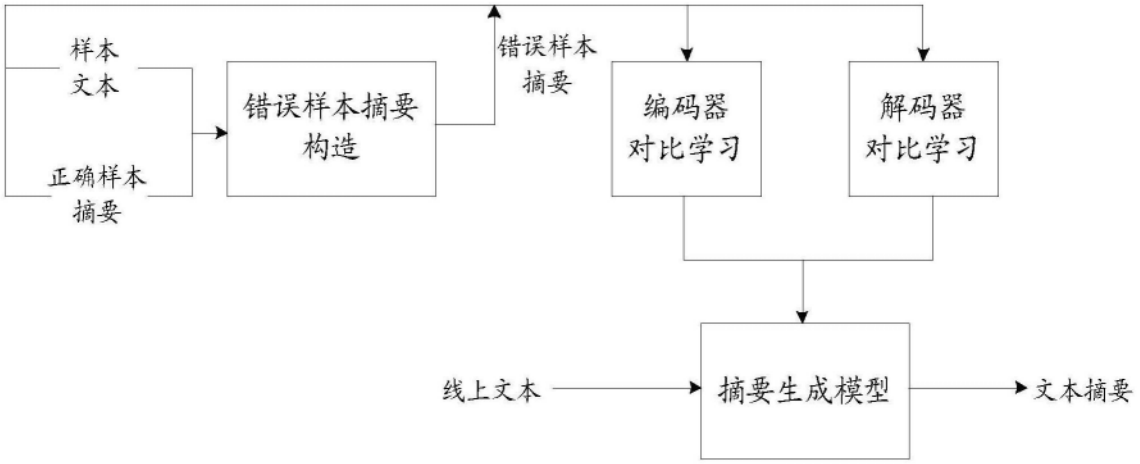

样本
文本
正确样本
摘要
错误样本摘要构造
错误样本
摘要
编码器
对比学习
解码器
对比学习
线上文本
摘要生成模型
文本摘要

图9


“一个理性的文字治愈师，希望你能够在我的文字里，找到一点对于生活的慰藉.....那么下面就和大家分享四本我觉得对我有较大帮助的书.....”
信息流文章
摘要生成模型
“推荐四本好书，很适合大家阅读：丰富自己，永远不会吃亏”
信息流文章摘要

图10

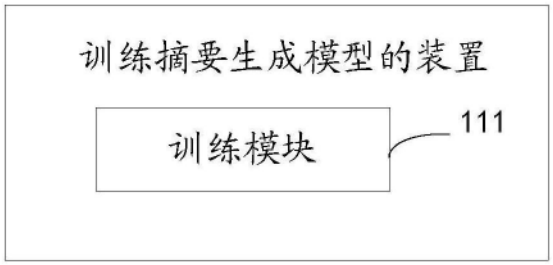

训练摘要生成模型的装置
训练模块
111

图11

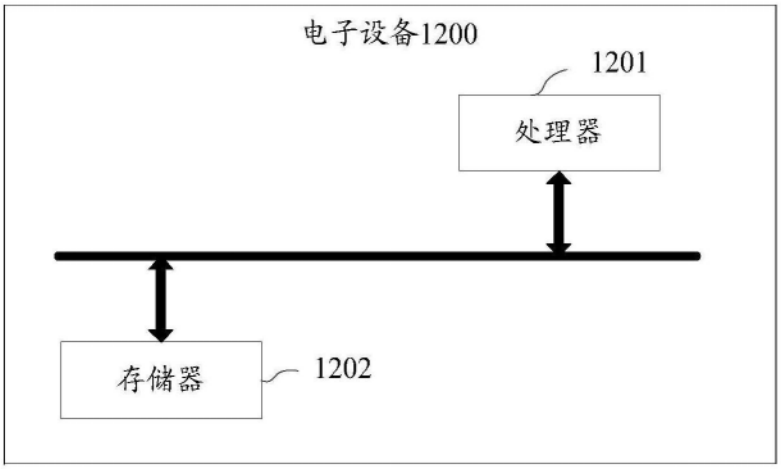

电子设备1200
1201
处理器
存储器
1202

图12

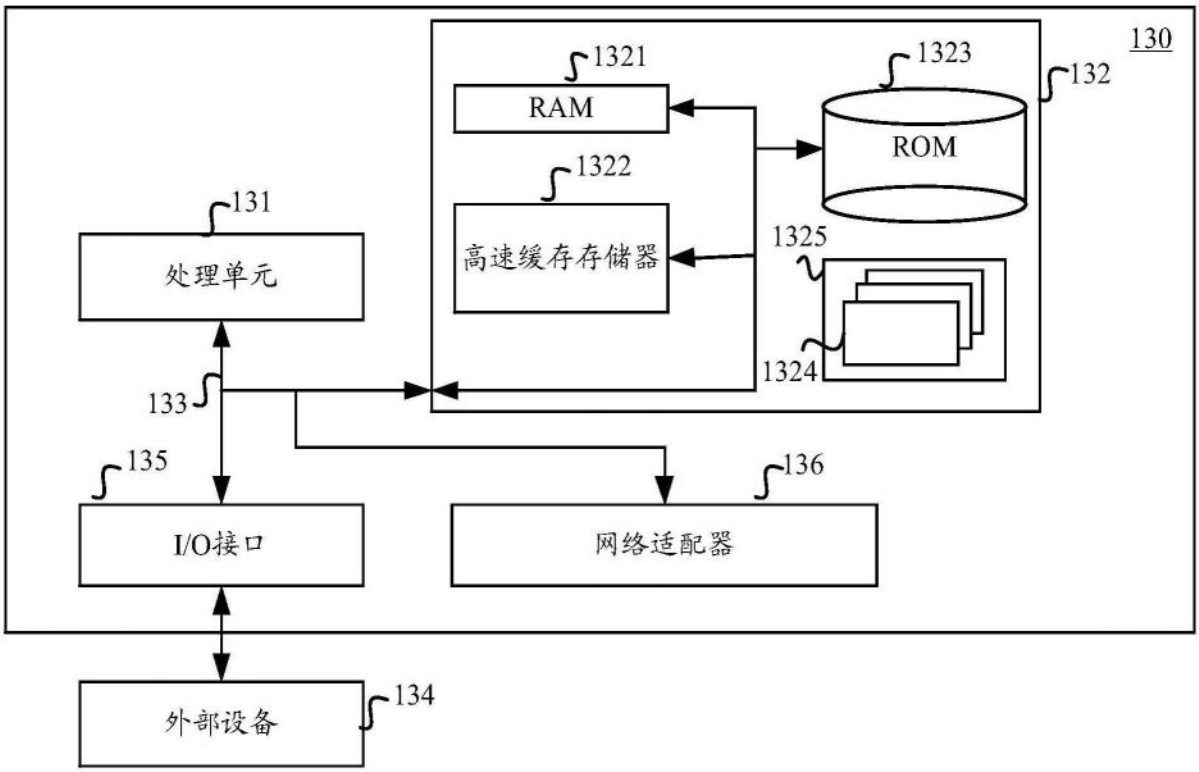

130
1321
RAM
1323
ROM
132
1322
高速缓存存储器
1325
1324
131
处理单元
133
135
I/O接口
136
网络适配器
外部设备
134

图13