(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 117851586 A
(43)申请公布日 2024.04.09

(21)申请号 202410038253.1

(22)申请日 2024.01.10

(71)申请人 京东城市(北京)数字科技有限公司
地址 100086 北京市海淀区知春路76号(写字楼)1号楼9层1-7-5号
申请人 京东科技信息技术有限公司

(72)发明人 申磊 陈蒙

(74)专利代理机构 北京律智知识产权代理有限公司 11438
专利代理师 韦萍

(51)Int.Cl.
G06F 16/34(2019.01)
G06F 40/35(2020.01)
G06F 18/214(2023.01)
G06F 40/289(2020.01)

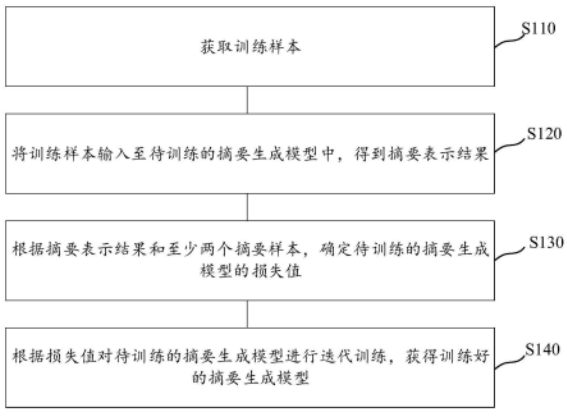
权利要求书2页 说明书12页 附图7页

(54)发明名称

摘要生成模型的训练方法及装置、存储介质、电子设备

(57)摘要

本公开涉及人工智能技术领域,提供了一种摘要生成模型的训练方法、摘要生成模型的训练装置、计算机存储介质、电子设备,其中,摘要生成模型的训练方法包括:获取训练样本,训练样本包括多个对话样本及每个对话样本对应的至少两个摘要样本;至少两个摘要样本包括正摘要样本和负摘要样本,负摘要样本是对正摘要样本进行至少一个维度的语义变换操作得到的;将每个对话样本输入至待训练的摘要生成模型中,得到摘要表示结果;根据摘要表示结果和至少两个摘要样本,确定待训练的摘要生成模型的损失值;根据损失值对待训练的摘要生成模型进行迭代训练,获得训练好的摘要生成模型。本公开能够保证模型输出结果与原始对话的事实一致性。



CN 117851586 A

1.一种摘要生成模型的训练方法,其特征在于,包括:

获取训练样本,所述训练样本包括多个对话样本及每个所述对话样本对应的至少两个摘要样本;所述至少两个摘要样本包括正摘要样本和负摘要样本,所述负摘要样本是对所述正摘要样本进行至少一个维度的语义变换操作得到的;

将每个所述对话样本输入至待训练的摘要生成模型中,得到摘要表示结果;

根据所述摘要表示结果和所述至少两个摘要样本,确定所述待训练的摘要生成模型的损失值;

根据所述损失值对所述待训练的摘要生成模型进行迭代训练,获得训练好的摘要生成模型。

2.根据权利要求1所述的方法,其特征在于,所述正摘要样本包括第一正摘要样本和第二正摘要样本;所述第一正摘要样本是预先标注的、所述对话样本对应的标准摘要,所述标准摘要是用指定语言表征的;

所述第二正摘要样本通过以下方式获得:

将所述第一正摘要样本翻译为中间语言文本;

将所述中间语言文本回译为所述指定语言,得到所述第二正摘要样本。

3.根据权利要求2所述的方法,其特征在于,所述负摘要样本通过以下方式获得:

通过第一自然语言处理模型对所述对话样本进行自然语言处理,获得所述对话样本的抽象含义表示;

对所述抽象含义表示进行至少一个维度的语义变换操作,获得变换后的抽象含义表示;

通过第二自然语言处理模型将所述变换后的抽象含义表示转换为目标对话样本;

通过第三自然语言处理模型对所述目标对话样本进行摘要提取处理,获得所述负摘要样本。

4.根据权利要求3所述的方法,其特征在于,所述对所述抽象含义表示进行至少一个维度的语义变换操作,获得变换后的抽象含义表示,包括:

对所述抽象语义表示中的实体信息进行替换操作,获得变换后的抽象语义表示;

对所述抽象语义表示中的情感极性进行变换操作,获得变换后的抽象语义表示;

对所述抽象语义表示中的非主干信息进行删减操作,获得变换后的抽象语义表示;

对所述抽象语义表示中的答案语句进行删除操作,获得变换后的抽象语义表示。

5.根据权利要求1所述的方法,其特征在于,所述根据所述摘要表示结果和所述至少两个摘要样本,确定所述待训练的摘要生成模型的损失值,包括:

根据所述至少两个摘要样本之间的差异程度,确定所述待训练的摘要生成模型的第一损失值;

根据所述摘要表示结果与所述正摘要样本之间的差异程度,确定所述待训练的摘要生成模型的第二损失值;

根据所述第一损失值和所述第二损失值,确定所述待训练的摘要生成模型的损失值。

6.根据权利要求5所述的方法,其特征在于,所述根据所述至少两个摘要样本之间的差异程度,确定所述待训练的摘要生成模型的第一损失值,包括:

根据所述第一正摘要样本和所述第二正摘要样本之间的差异程度,以及,所述第一正

摘要样本和所述负摘要样本之间的差异程度,确定所述待训练的摘要生成模型的第一损失值。

7.根据权利要求5所述的方法,其特征在于,所述根据所述摘要表示结果与所述正摘要样本之间的差异程度,确定所述待训练的摘要生成模型的第二损失值,包括:

根据所述摘要表示结果和所述第一正摘要样本之间的差异程度,确定所述待训练的摘要生成模型的第二损失值。

8.根据权利要求1所述的方法,其特征在于,所述根据所述损失值对所述待训练的摘要生成模型进行迭代训练,得到训练好的摘要生成模型,包括:

根据所述损失值对所述待训练的摘要生成模型进行迭代训练,直至所述损失值满足预设的收敛条件,得到训练好的摘要生成模型。

9.一种摘要生成模型的训练装置,其特征在于,包括:

训练样本获取模块,用于获取训练样本,所述训练样本包括多个对话样本及每个所述对话样本对应的至少两个摘要样本;所述至少两个摘要样本包括正摘要样本和负摘要样本,所述负摘要样本是对所述正摘要样本进行至少一个维度的语义变换操作得到的;

摘要表示模块,用于将每个所述对话样本输入至待训练的摘要生成模型中,得到摘要表示结果;

损失值确定模块,用于根据所述摘要表示结果和所述至少两个摘要样本,确定所述待训练的摘要生成模型的损失值;

模型训练模块,用于根据所述损失值对所述待训练的摘要生成模型进行迭代训练,获得训练好的摘要生成模型。

10.一种计算机存储介质,其上存储有计算机程序,其特征在于,所述计算机程序被处理器执行时实现权利要求1~8中任意一项所述的摘要生成模型的训练方法。

11.一种电子设备,其特征在于,包括:

处理器;以及

存储器,用于存储所述处理器的可执行指令;

其中,所述处理器配置为经由执行所述可执行指令来执行权利要求1~8中任意一项所述的摘要生成模型的训练方法。

摘要生成模型的训练方法及装置、存储介质、电子设备

技术领域

[0001] 本公开涉及人工智能技术领域,特别涉及一种摘要生成模型的训练方法、摘要生成模型的训练装置、计算机存储介质及电子设备。

背景技术

[0002] 对话摘要指的是能够涵盖对话中所有重要信息的短文本,通过对话摘要,可以对对话内容进行总结。大量的真实对话场景中可以利用到这项技术,例如会议活动和医疗诊断。

[0003] 相关技术中,一般是通过最大似然估计的方式的来训练摘要生成模型,最大似然概率函数是二元组函数,基于样本的预测表示和真实表示构建。但是,由于最大似然概率函数只是反映统计关系,不能准确反映语义关系,会影响摘要生成模型的准确度。

[0004] 鉴于此,本领域亟需开发一种新的摘要生成模型的训练方法及装置。

[0005] 需要说明的是,上述背景技术部分公开的信息仅用于加强对本公开的背景的理解。

发明内容

[0006] 本公开的目的在于提供一种摘要生成模型的训练方法、摘要生成模型的训练装置、计算机存储介质及电子设备,进而至少在一定程度上克服由于相关技术的限制而导致的模型准确度不足的技术问题。

[0007] 本公开的其他特性和优点将通过下面的详细描述变得显然,或部分地通过本公开的实践而习得。

[0008] 根据本公开的第一方面,提供一种摘要生成模型的训练方法,包括:获取训练样本,所述训练样本包括多个对话样本及每个所述对话样本对应的至少两个摘要样本;所述至少两个摘要样本包括正摘要样本和负摘要样本,所述负摘要样本是对所述正摘要样本进行至少一个维度的语义变换操作得到的;将每个所述对话样本输入至待训练的摘要生成模型中,得到摘要表示结果;根据所述摘要表示结果和所述至少两个摘要样本,确定所述待训练的摘要生成模型的损失值;根据所述损失值对所述待训练的摘要生成模型进行迭代训练,获得训练好的摘要生成模型。

[0009] 在本公开的示例性实施例中,所述正摘要样本包括第一正摘要样本和第二正摘要样本;所述第一正摘要样本是预先标注的、所述对话样本对应的标准摘要,所述标准摘要是用指定语言表征的;所述第二正摘要样本通过以下方式获得:将所述第一正摘要样本翻译为中间语言文本;将所述中间语言文本回译为所述指定语言,得到所述第二正摘要样本。

[0010] 在本公开的示例性实施例中,所述负摘要样本通过以下方式获得:通过第一自然语言处理模型对所述对话样本进行自然语言处理,获得所述对话样本的抽象含义表示;对所述抽象含义表示进行至少一个维度的语义变换操作,获得变换后的抽象含义表示;通过第二自然语言处理模型将所述变换后的抽象含义表示转换为目标对话样本;通过第三自然

语言处理模型对所述目标对话样本进行摘要提取处理,获得所述负摘要样本。

[0011] 在本公开的示例性实施例中,所述对所述抽象含义表示进行至少一个维度的语义变换操作,获得变换后的抽象含义表示,包括:对所述抽象语义表示中的实体信息进行替换操作,获得变换后的抽象语义表示;对所述抽象语义表示中的情感极性进行变换操作,获得变换后的抽象语义表示;对所述抽象语义表示中的非主干信息进行删减操作,获得变换后的抽象语义表示;对所述抽象语义表示中的答案语句进行删除操作,获得变换后的抽象语义表示。

[0012] 在本公开的示例性实施例中,所述根据所述摘要表示结果和所述至少两个摘要样本,确定所述待训练的摘要生成模型的损失值,包括:根据所述至少两个摘要样本之间的差异程度,确定所述待训练的摘要生成模型的第一损失值;根据所述摘要表示结果与所述正摘要样本之间的差异程度,确定所述待训练的摘要生成模型的第二损失值;根据所述第一损失值和所述第二损失值,确定所述待训练的摘要生成模型的损失值。

[0013] 在本公开的示例性实施例中,所述根据所述至少两个摘要样本之间的差异程度,确定所述待训练的摘要生成模型的第一损失值,包括:根据所述第一正摘要样本和所述第二正摘要样本之间的差异程度,以及,所述第一正摘要样本和所述负摘要样本之间的差异程度,确定所述待训练的摘要生成模型的第一损失值。

[0014] 在本公开的示例性实施例中,所述根据所述摘要表示结果与所述正摘要样本之间的差异程度,确定所述待训练的摘要生成模型的第二损失值,包括:根据所述摘要表示结果和所述第一正摘要样本之间的差异程度,确定所述待训练的摘要生成模型的第二损失值。

[0015] 在本公开的示例性实施例中,所述根据所述损失值对所述待训练的摘要生成模型进行迭代训练,得到训练好的摘要生成模型,包括:根据所述损失值对所述待训练的摘要生成模型进行迭代训练,直至所述损失值满足预设的收敛条件,得到训练好的摘要生成模型。

[0016] 根据本公开的第二方面,提供一种摘要生成模型的训练装置,包括:训练样本获取模块,用于获取训练样本,所述训练样本包括多个对话样本及每个所述对话样本对应的至少两个摘要样本;所述至少两个摘要样本包括正摘要样本和负摘要样本,所述负摘要样本是对所述正摘要样本进行至少一个维度的语义变换操作得到的;摘要表示模块,用于将每个所述对话样本输入至待训练的摘要生成模型中,得到摘要表示结果;损失值确定模块,用于根据所述摘要表示结果和所述至少两个摘要样本,确定所述待训练的摘要生成模型的损失值;模型训练模块,用于根据所述损失值对所述待训练的摘要生成模型进行迭代训练,获得训练好的摘要生成模型。

[0017] 根据本公开的第三方面,提供一种计算机存储介质,其上存储有计算机程序,所述计算机程序被处理器执行时实现上述第一方面所述的摘要生成模型的训练方法。

[0018] 根据本公开的第四方面,提供一种电子设备,包括:处理器;以及存储器,用于存储所述处理器的可执行指令;其中,所述处理器配置为经由执行所述可执行指令来执行上述第一方面所述的摘要生成模型的训练方法。

[0019] 由上述技术方案可知,本公开示例性实施例中的摘要生成模型的训练方法、摘要生成模型的训练装置、计算机存储介质及电子设备至少具备以下优点和积极效果:

[0020] 在本公开的一些实施例所提供的技术方案中,一方面,鉴于本公开的训练样本中包含正摘要样本和负摘要样本,且负摘要样本是对所述正摘要样本进行至少一个维度的语

义变换操作得到的,从而,相比现有技术中仅通过实体替换来构造负样本的单一方式而言,本公开能够构造类型丰富的负摘要样本,从而有利于模型训练过程中学习更加全面,避免出现事实不一致等错误。另一方面,本公开通过将每个所述对话样本输入至待训练的摘要生成模型中,得到摘要表示结果,根据所述摘要表示结果和所述至少两个摘要样本,确定所述待训练的摘要生成模型的损失值,根据所述损失值对所述待训练的摘要生成模型进行迭代训练,获得训练好的摘要生成模型,能够根据摘要表示结果、正摘要样本和负摘要样本等多个维度的元素来构造模型损失值,从而,有助于从多个维度对模型的损失进行优化,以提升模型的准确度。

[0021] 本公开应当理解的是,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

[0022] 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。显而易见地,下面描述中的附图仅仅是本公开的一些实施例,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。

[0023] 图1示出本公开实施例中摘要生成模型的训练方法的流程示意图;

[0024] 图2示出本公开实施例中如何生成负摘要样本的流程示意图;

[0025] 图3示出本公开实施例中对话样本对应的抽象含义表示的示意图;

[0026] 图4示出本公开实施例中如何对对抽象含义表示进行至少一个维度的语义变换操作,获得变换后的抽象含义表示的流程示意图;

[0027] 图5示出本公开实施例中如何对抽象含义表示中的实体信息进行替换操作,获得变换后的抽象语义表示的示意图;

[0028] 图6示出本公开实施例中如何对抽象语义表示中的情感极性进行变换操作,获得变换后的抽象语义表示的示意图;

[0029] 图7示出本公开实施例中对抽象语义表示中的非主干信息进行删减操作,获得变换后的抽象语义表示的示意图;

[0030] 图8示出本公开实施例中如何对抽象语义表示中的答案语句进行删除操作,获得变换后的抽象语义表示的示意图;

[0031] 图9示出本公开实施例中如何生成负摘要样本的整体示意图;

[0032] 图10示出本公开实施例中如何根据摘要表示结果和至少两个摘要样本,确定待训练的摘要生成模型的损失值的流程示意图;

[0033] 图11示出本公开示例性实施例中摘要生成模型的训练装置的结构示意图;

[0034] 图12示出本公开示例性实施例中电子设备的结构示意图。

具体实施方式

[0035] 现在将参考附图更全面地描述示例实施方式。然而,示例实施方式能够以多种形式实施,且不应被理解为限于在此阐述的范例;相反,提供这些实施方式使得本公开将更加全面和完整,并将示例实施方式的构思全面地传达给本领域的技术人员。所描述的特征、结

构或特性可以以任何合适的方式结合在一个或更多实施方式中。在下面的描述中,提供许多具体细节从而给出对本公开的实施方式的充分理解。然而,本领域技术人员将意识到,可以实践本公开的技术方案而省略所述特定细节中的一个或更多,或者可以采用其它的方法、组元、装置、步骤等。在其它情况下,不详细示出或描述公知技术方案以避免喧宾夺主而使得本公开的各方面变得模糊。

[0036] 本说明书中使用用语“一个”、“一”、“该”和“所述”用以表示存在一个或多个要素/组成部分/等;用语“包括”和“具有”用以表示开放式的包括在内的意思并且是指除了列出的要素/组成部分/等之外还可存在另外的要素/组成部分/等;用语“第一”和“第二”等仅作为标记使用,不是对其对象的数量限制。

[0037] 此外,附图仅为本公开的示意性图解,并非一定是按比例绘制。图中相同的附图标记表示相同或类似的部分,因而将省略对它们的重复描述。附图中所示的一些方框图是功能实体,不一定必须与物理或逻辑上独立的实体相对应。

[0038] 对话摘要指的是能够涵盖对话中所有重要信息的短文本,通过对话摘要,可以对对话内容进行总结。大量的真实对话场景中可以利用到这项技术,例如会议活动和医疗诊断。

[0039] 产生与对话内容事实相符的总结是应用生成式摘要的基本前提,许多对话摘要工作都试图消除生成摘要中的事实性错误(幻觉)。这些错误可以分为内在错误和外在错误,内在错误指的是产生了与原始对话相矛盾的信息,外在错误指的是产生了原始对话中不存在的内容。

[0040] 为了解决对话摘要中的事实不一致问题,现有的方案一般可以分为以下两种:

[0041] 第一种是辅助算法的融合,例如:通过指代消歧模型来获得指代簇,并将这些指代簇整合到摘要生成模型中,以减少生成摘要时的错误指代关系;第二种是采用对比学习的思路,例如:通过修改实体信息来构建负样本。

[0042] 然而,上述方案具有以下缺陷:

[0043] 融合辅助算法的方案,可能会将来自辅助算法模块的错误引入到摘要生成中;

[0044] 而采用对比学习的方案一般主要基于词替换构造负摘要样本,该方案忽略了话语层面以及整轮对话层面会出现的事实不一致,导致构造的负样本类别不充分,从而,影响后续模型输出结果的准确度。

[0045] 在本公开的实施例中,首先提供了一种摘要生成模型的训练方法,至少在一定程度上克服相关技术中模型准确度不足的缺陷。

[0046] 图1示出本公开实施例中摘要生成模型的训练方法的流程示意图,该摘要生成模型的训练方法的执行主体可以是对摘要生成模型进行训练的服务器。

[0047] 参考图1,根据本公开的一个实施例的摘要生成模型的训练方法包括以下步骤:

[0048] 步骤S110,获取训练样本,训练样本包括多个对话样本及每个对话样本对应的至少两个摘要样本;至少两个摘要样本包括正摘要样本和负摘要样本,负摘要样本是对正摘要样本进行至少一个维度的语义变换操作得到的;

[0049] 步骤S120,将训练样本输入至待训练的摘要生成模型中,得到摘要表示结果;

[0050] 步骤S130,根据摘要表示结果和至少两个摘要样本,确定待训练的摘要生成模型的损失值;

[0051] 步骤S140,根据损失值对待训练的摘要生成模型进行迭代训练,获得训练好的摘要生成模型。

[0052] 在图1所示实施例所提供的技术方案中,一方面,鉴于本公开的训练样本中包含正摘要样本和负摘要样本,且负摘要样本是对所述正摘要样本进行至少一个维度的语义变换操作得到的,从而,相比现有技术中仅通过实体替换来构造负样本的单一方式而言,本公开能够构造类型丰富的负摘要样本,从而有利于模型训练过程中学习更加全面,避免出现事实不一致等错误。另一方面,本公开通过将每个所述对话样本输入至待训练的摘要生成模型中,得到摘要表示结果,根据所述摘要表示结果和所述至少两个摘要样本,确定所述待训练的摘要生成模型的损失值,根据所述损失值对所述待训练的摘要生成模型进行迭代训练,获得训练好的摘要生成模型,能够根据摘要表示结果、正摘要样本和负摘要样本等多个维度的元素来构造模型损失值,从而,有助于从多个维度对模型的损失进行优化,以提升模型的准确度。

[0053] 以下对图1中的各个步骤的具体实现过程进行详细阐述:

[0054] 在步骤S110中,获取训练样本。

[0055] 本步骤中,可以获取用于训练摘要生成模型的训练样本,该训练样本中可以包含多个对话样本及每个对话样本对应的至少两个摘要样本。

[0056] 其中,上述至少两个摘要样本包括正摘要样本和负摘要样本。正摘要样本是指与对话样本语义一致的摘要样本,负摘要样本是指与对话样本语义不一致的摘要样本。比如,一个对话样本是说某个电影很好,若一个摘要样本也是说该电影很好或类似的语义,比如极好、还可以等,则该摘要样本为正摘要样本,若一个摘要样本是说该电影不好,比如糟透了,则该摘要样本为负摘要样本。

[0057] 其中,上述正摘要样本包括第一正摘要样本和第二正摘要样本,第一正摘要样本可以是预先标注的、对话样本对应的标准摘要。

[0058] 示例性的,上述第二正摘要样本可以是通过以下方式获得的:

[0059] 可以先采用回译技术将上述第一正摘要样本翻译成中间语言文本,例如:法语文本,可以根据实际情况自行设定,本公开对此不作特殊限定,之后,再将中间语言文本回译为上述指定语言,以得到上述第二正摘要样本。

[0060] 示例性的,上述负摘要样本可以是通过以下方式获得的:

[0061] 参考图2,图2示出本公开实施例中如何生成负摘要样本的流程示意图,包含步骤S201-步骤S204:

[0062] 在步骤S201中,通过第一自然语言处理模型对对话样本进行自然语言处理,获得对话样本的抽象含义表示。

[0063] 本步骤中,示例性的,可以通过第一个T5模型(Text-to-Text Transfer Transformer)对上述对话样本进行自然语言处理,获得上述对话样本对应的抽象含义表示AMR。

[0064] AMR全称abstract meaning representation,是对语言的一种表示方法,它并不表示单个的句子,而是着重表示句子中的逻辑,也就是说它不存在语法和单词的概念,只有semantic relation和concept的概念,多个concept之间被semantic relation连接起来。AMR在句子中捕捉“谁在对谁做什么”。每个句子都表示为一个有根、有向、无环的图,每条边

代表一个语义关系semantic relation,每个叶子节点代表一个concept。

[0065] 参考图3,图3示出本公开实施例中对话样本对应的抽象含义表示的示意图,具体示出对话样本“Emma:I will try.I can't go on like this anymore.”对应的抽象含义表示,如图3所示,ARG0、ARG1、ARG2表示实体或概念,带有短划线的短句表示实体或概念之间的关系。通过AMR表示,能够从句法层面提供对话的核心概念和明确结构。

[0066] 在步骤S202中,对抽象含义表示进行至少一个维度的语义变换操作,获得变换后的抽象含义表示。

[0067] 本步骤中,参考图4,图4示出本公开实施例中如何对对抽象含义表示进行至少一个维度的语义变换操作,获得变换后的抽象含义表示的流程示意图,包含步骤S401-步骤S404:

[0068] 在步骤S401中,对抽象语义表示中的实体信息进行替换操作,获得变换后的抽象语义表示。

[0069] 本步骤中,上述实体信息可以是人名、物品名、时间等,可以根据实际情况自行设定,本公开对此不作特殊限定。

[0070] 参考图5,图5示出本公开实施例中如何对抽象含义表示中的实体信息进行替换操作,获得变换后的抽象语义表示的示意图,如图5所示:

[0071] 针对于对话样本“Doylan:Coming back on saturday for party would be the best”对应的AMR表示,可以将其中的人名“Doylan”替换为“Beau”,将其中的时间“saturday”替换为“friday”,以得到变换后的抽象语义表示。

[0072] 接着参考图4,在步骤S402中,对抽象语义表示中的情感极性进行变换操作,获得变换后的抽象语义表示。

[0073] 本步骤中,情感极性指的可以是话语的肯定和否定情感。

[0074] 参考图6,图6示出本公开实施例中如何对抽象语义表示中的情感极性进行变换操作,获得变换后的抽象语义表示的示意图,如图6所示:

[0075] 针对于对话样本“Monica:Son,that's not acceptable”对应的AMR表示,可以将其中表示否定的情感信息“not”删掉,以对其情感信息进行变换,获得变换后的抽象语义表示。

[0076] 接着参考图4,在步骤S403中,对抽象语义表示中的非主干信息进行删减操作,获得变换后的抽象语义表示。

[0077] 本步骤中,非主干信息可以是句子中的状语成分、补语成分,即除去主谓宾之外的成分,或者,除去主系表之外的成分。

[0078] 示例性的,参考图7,图7示出本公开实施例中对抽象语义表示中的非主干信息进行删减操作,获得变换后的抽象语义表示的示意图,如图7所示:

[0079] 针对对话样本“Caro:I have just received a most beautiful and apologetic text message from my prince charming!”对应的AMR表示,可以将其非主干信息“most beautiful and apologetic text、from my prince charming”等进行删减操作,获得变换后的抽象语义表示。

[0080] 接着参考图4,在步骤S404中,对抽象语义表示中的答案语句进行删除操作,获得变换后的抽象语义表示。

[0081] 本步骤中,答案语句指的是问答式对话中回答问题的一方所回应的语句。

[0082] 参考图8,图8示出本公开实施例中如何对抽象语义表示中的答案语句进行删除操作,获得变换后的抽象语义表示的示意图,如图8所示:

[0083] 针对对话样本“Marion:Have they found the scratches?Kim:No!They didn't say anything.Simon:So relieved.”所对应的AMR表示,可以将其中的答案语句“Kim:No!They didn't say anything.”删除,以获得变更后的抽象语义表示。

[0084] 接着参考图2,在步骤S203中,通过第二自然语言处理模型将变换后的抽象含义表示转换为目标对话样本。

[0085] 本步骤中,在获得变换后的抽象含义表示之后,可以通过第二个T5模型将变换后的抽象含义表示转换为目标对话样本。

[0086] 在步骤S204中,通过第三自然语言处理模型对目标对话样本进行摘要提取处理,获得负摘要样本。

[0087] 本步骤中,可以通过第三自然语言处理模型(例如:BART,可以根据实际情况自行设定,本公开对此不作特殊限定)对上述目标对话样本进行摘要提取处理,获得负摘要样本。

[0088] 参考图9,图9示出本公开实施例中如何生成负摘要样本的整体示意图,如图9所示:

[0089] 针对对话样本“Mark:What are we doing tonight?Predrinks at your place?Daniel:My sister is visiting.Mark:All right all right drinks at my place.Daniel:Okay”,本公开可以先通过第一个T5模型将上述对话样本转换为AMR表示,之后,可以对该AMR表示进行至少一个维度的语义变换处理,得到转换后的AMR表示。之后,可以利用第二个T5模型将该转换后的AMR表示转换为目标对话,进而,可以利用BART模型基于该目标对话提取出负摘要样本。

[0090] 示例性的,在正摘要样本为“Daniel's sister is visiting him.Daniel and Mark will meet at Mark's place for drinks.”的情况下,生成的负摘要样本可以是“Mark and Daniel are going to meet for drinks at Daniel's place”。

[0091] 基于本公开所提供的负摘要样本的生成方案,能够从实体、话语、对话这三个不同的层面对对话样本进行干扰,从而构造类型丰富的负摘要样本。

[0092] 接着参考图1,在步骤S120中,将每个对话样本输入至待训练的摘要生成模型中,得到摘要表示结果。

[0093] 本步骤中,可以将每个对话样本输入至待训练的摘要生成模型中,根据待训练的摘要生成模型的输出,获得摘要表示结果。

[0094] 示例性的,上述待训练的摘要生成模型可以是BART模型,可以根据实际情况自行设定,本公开对此不作特殊限定。

[0095] 在步骤S130中,根据摘要表示结果和至少两个摘要样本,确定待训练的摘要生成模型的损失值。

[0096] 本步骤中,参考图10,图10示出本公开实施例中如何根据摘要表示结果和至少两个摘要样本,确定待训练的摘要生成模型的损失值的流程示意图,包含步骤S1001-步骤S1003:

[0097] 在步骤S1001中,根据至少两个摘要样本之间的差异程度,确定待训练的摘要生成模型的第一损失值。

[0098] 本步骤中,可以根据第一正摘要样本和第二正摘要样本之间的差异程度,以及,第一正摘要样本和负摘要样本之间的差异程度,确定待训练的摘要生成模型的第一损失值,即第一损失值是上述至少两个摘要样本之间的对比损失(contrastive loss),鉴于对比损失函数是基于三元组构建的损失函数,以更好地对比正负样本之间的关系。

[0099] 示例性的,可以基于以下公式1确定上述第一损失值:

[0100] $$\mathcal{L}_{CL} = -\sum_{s_i \neq s_j} \log \frac{\exp\left(\mathrm{sim}(g_i, g_j)\right)/\tau}{\sum_{s_i \neq s_t} \exp(\mathrm{sim}(g_i, g_t))/\tau} \qquad \text{公式 1}$$

[0101] 其中,$\mathcal{L}_{CL}$表示上述第一损失值,s_i是第一正摘要样本,s_j是第二正摘要样本,s_t是负摘要样本。g_i、g_j、g_t是s_i、s_j和s_t被输入至待训练的摘要生成模型之后各自对应的特征表示。τ是一个超参数,被设置为1.0,可以根据实际情况自行设定,本公开对此不作特殊限定。$\mathrm{sim}(\cdot)$是余弦相似度。

[0102] 在步骤S1002中,根据摘要表示结果与正摘要样本之间的差异程度,确定待训练的摘要生成模型的第二损失值。

[0103] 本步骤中,可以根据摘要表示结果和第一正摘要样本之间的差异程度,确定待训练的摘要生成模型的第二损失值$\mathcal{L}_{CE}$。

[0104] 示例性的,可以计算摘要表示结果与第一正摘要样本之间的相似度(例如:余弦相似度、欧式距离等,可以根据实际情况自行设定,本公开对此不作特殊限定),根据该相似度确定上述待训练的摘要生成模型的第二损失值$\mathcal{L}_{CE}$。

[0105] 在步骤S 1 003中,根据第一损失值和第二损失值,确定待训练的摘要生成模型的损失值。

[0106] 本步骤中,可以基于以下公式2确定上述待训练的摘要生成模型的损失值:

[0107] $$\mathcal{L} = \mathcal{L}_{CE} + \lambda \mathcal{L}_{CL} \qquad \text{公式 2}$$

[0108] 其中,$\mathcal{L}$表示上述损失值,λ为预设的系数,可以根据实际情况自行设定,本公开对此不作特殊限定。

[0109] 通过基于对比损失和交叉熵损失共同构建模型的损失,便于从多个维度全面的对模型的损失进行衡量和改进,以提高摘要生成模型的准确度。

[0110] 接着参考图1,在步骤S140中,根据损失值对待训练的摘要生成模型进行迭代训练,获得训练好的摘要生成模型。

[0111] 本步骤中,可以根据损失值对待训练的摘要生成模型进行迭代训练,直至损失值满足预设的收敛条件,得到训练好的摘要生成模型。

[0112] 在训练好上述摘要生成模型之后,在模型的应用阶段,可以将待处理的对话信息输入上述摘要生成模型中,通过该摘要生成模型对该对话信息进行处理,自动输出该对话信息对应的摘要。基于摘要生成模型生成的摘要,可以生成原始的文档中没有的新的词语、短语等,以对对话信息中的关键信息进行简要总结。

[0113] 基于以上技术方案,本公开至少具有以下技术效果:

[0114] 第一,将AMR图表示引入到对话摘要生成任务中,可对对话信息进行更好的表征,同时建立对话与摘要间更密切的关系;

[0115] 第二,基于AMR图进行实体、话语、对话级别的操作,从而构建类别更丰富的负摘要样本,能够覆盖更多的事实不一致性错误;

[0116] 第三,利用多种负摘要样本进行对比学习训练,提升模型生成事实一致性摘要的能力。

[0117] 本公开还提供了一种摘要生成模型的训练装置,图11示出本公开示例性实施例中摘要生成模型的训练装置的结构示意图;如图11所示,摘要生成模型的训练装置1100可以包括训练样本获取模块1110、摘要表示模块1120、损失值确定模块1130和模型训练模块1140。其中:

[0118] 训练样本获取模块1110,用于获取训练样本,所述训练样本包括多个对话样本及每个所述对话样本对应的至少两个摘要样本;所述至少两个摘要样本包括正摘要样本和负摘要样本,所述负摘要样本是对所述正摘要样本进行至少一个维度的语义变换操作得到的;

[0119] 摘要表示模块1120,用于将每个所述对话样本输入至待训练的摘要生成模型中,得到摘要表示结果;

[0120] 损失值确定模块1130,用于根据所述摘要表示结果和所述至少两个摘要样本,确定所述待训练的摘要生成模型的损失值;

[0121] 模型训练模块1140,用于根据所述损失值对所述待训练的摘要生成模型进行迭代训练,获得训练好的摘要生成模型。

[0122] 在本公开的示例性实施例中,所述正摘要样本包括第一正摘要样本和第二正摘要样本;所述第一正摘要样本是预先标注的、所述对话样本对应的标准摘要,所述标准摘要是用指定语言表征的;所述第二正摘要样本通过以下方式获得:将所述第一正摘要样本翻译为中间语言文本;将所述中间语言文本回译为所述指定语言,得到所述第二正摘要样本。

[0123] 在本公开的示例性实施例中,所述负摘要样本通过以下方式获得:通过第一自然语言处理模型对所述对话样本进行自然语言处理,获得所述对话样本的抽象含义表示;对所述抽象含义表示进行至少一个维度的语义变换操作,获得变换后的抽象含义表示;通过第二自然语言处理模型将所述变换后的抽象含义表示转换为目标对话样本;通过第三自然语言处理模型对所述目标对话样本进行摘要提取处理,获得所述负摘要样本。

[0124] 在本公开的示例性实施例中,训练样本获取模块1110对所述抽象含义表示进行至少一个维度的语义变换操作,获得变换后的抽象含义表示,包括:对所述抽象语义表示中的实体信息进行替换操作,获得变换后的抽象语义表示;对所述抽象语义表示中的情感极性进行变换操作,获得变换后的抽象语义表示;对所述抽象语义表示中的非主干信息进行删减操作,获得变换后的抽象语义表示;对所述抽象语义表示中的答案语句进行删除操作,获得变换后的抽象语义表示。

[0125] 在本公开的示例性实施例中,损失值确定模块1130根据所述摘要表示结果和所述至少两个摘要样本,确定所述待训练的摘要生成模型的损失值,包括:根据所述至少两个摘要样本之间的差异程度,确定所述待训练的摘要生成模型的第一损失值;根据所述摘要表示结果与所述正摘要样本之间的差异程度,确定所述待训练的摘要生成模型的第二损失

值;根据所述第一损失值和所述第二损失值,确定所述待训练的摘要生成模型的损失值。

[0126] 在本公开的示例性实施例中,损失值确定模块1130根据所述至少两个摘要样本之间的差异程度,确定所述待训练的摘要生成模型的第一损失值,包括:根据所述第一正摘要样本和所述第二正摘要样本之间的差异程度,以及,所述第一正摘要样本和所述负摘要样本之间的差异程度,确定所述待训练的摘要生成模型的第一损失值。

[0127] 在本公开的示例性实施例中,损失值确定模块1130根据所述摘要表示结果与所述正摘要样本之间的差异程度,确定所述待训练的摘要生成模型的第二损失值,包括:根据所述摘要表示结果和所述第一正摘要样本之间的差异程度,确定所述待训练的摘要生成模型的第二损失值。

[0128] 在本公开的示例性实施例中,模型训练模块1140根据所述损失值对所述待训练的摘要生成模型进行迭代训练,得到训练好的摘要生成模型,包括:根据所述损失值对所述待训练的摘要生成模型进行迭代训练,直至所述损失值满足预设的收敛条件,得到训练好的摘要生成模型。

[0129] 上述摘要生成模型的训练装置中各模块的具体细节已经在对应的摘要生成模型的训练方法中进行了详细的描述,因此此处不再赘述。

[0130] 应当注意,尽管在上文详细描述中提及了用于动作执行的设备的若干模块或者单元,但是这种划分并非强制性的。实际上,根据本公开的实施方式,上文描述的两个或更多模块或者单元的特征和功能可以在一个模块或者单元中具体化。反之,上文描述的一个模块或者单元的特征和功能可以进一步划分为由多个模块或者单元来具体化。

[0131] 此外,尽管在附图中以特定顺序描述了本公开中方法的各个步骤,但是,这并非要求或者暗示必须按照该特定顺序来执行这些步骤,或是必须执行全部所示的步骤才能实现期望的结果。附加的或备选的,可以省略某些步骤,将多个步骤合并为一个步骤执行,以及/或者将一个步骤分解为多个步骤执行等。

[0132] 通过以上的实施方式的描述,本领域的技术人员易于理解,这里描述的示例实施方式可以通过软件实现,也可以通过软件结合必要的硬件的方式来实现。因此,根据本公开实施方式的技术方案可以以软件产品的形式体现出来,该软件产品可以存储在一个非易失性存储介质(可以是CD-ROM,U盘,移动硬盘等)中或网络上,包括若干指令以使得一台计算设备(可以是个人计算机、服务器、移动终端、或者网络设备等)执行根据本公开实施方式的方法。

[0133] 本申请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可以是上述实施例中描述的电子设备中所包含的;也可以是单独存在,而未装配入该电子设备中。

[0134] 计算机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。计算机可读存储介质的更具体的例子可以包括但不限于:具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、光纤、便携式紧凑磁盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本公开中,计算机可读存储介质可以是任何包含或存储程序的有形介质,该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。

[0135] 计算机可读存储介质可以发送、传播或者传输用于由指令执行系统、装置或者器

件使用或者与其结合使用的程序。计算机可读存储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、电线、光缆、RF等等,或者上述的任意合适的组合。

[0136] 计算机可读存储介质承载有一个或者多个程序,当上述一个或者多个程序被一个该电子设备执行时,使得该电子设备实现如上述实施例中所述的方法。

[0137] 此外,在本公开实施例中还提供了一种能够实现上述方法的电子设备。

[0138] 所属技术领域的技术人员能够理解,本公开的各个方面可以实现为系统、方法或程序产品。因此,本公开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”。

[0139] 下面参照图12来描述根据本公开的这种实施方式的电子设备1200。图12显示的电子设备1200仅仅是一个示例,不应对本公开实施例的功能和使用范围带来任何限制。

[0140] 如图12所示,电子设备1200以通用计算设备的形式表现。电子设备1200的组件可以包括但不限于:上述至少一个处理单元1210、上述至少一个存储单元1220、连接不同系统组件(包括存储单元1220和处理单元1210)的总线1230以及显示单元1240。

[0141] 其中,所述存储单元存储有程序代码,所述程序代码可以被所述处理单元1210执行,使得所述处理单元1210执行本说明书上述“示例性方法”部分中描述的根据本公开各种示例性实施方式的步骤。例如,所述处理单元1210可以执行如图1中所示的:步骤S110,获取训练样本,所述训练样本包括多个对话样本及每个所述对话样本对应的至少两个摘要样本;所述至少两个摘要样本包括正摘要样本和负摘要样本,所述负摘要样本是对所述正摘要样本进行至少一个维度的语义变换操作得到的;步骤S120,将所述训练样本输入至待训练的摘要生成模型中,得到摘要表示结果;步骤S130,根据所述摘要表示结果和所述至少两个摘要样本,确定所述待训练的摘要生成模型的损失值;步骤S140,根据所述损失值对所述待训练的摘要生成模型进行迭代训练,获得训练好的摘要生成模型。

[0142] 存储单元1220可以包括易失性存储单元形式的可读介质,例如随机存取存储单元(RAM)12201和/或高速缓存存储单元12202,还可以进一步包括只读存储单元(ROM)12203。

[0143] 存储单元1220还可以包括具有一组(至少一个)程序模块12205的程序/实用工具12204,这样的程序模块12205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。

[0144] 总线1230可以为表示几类总线结构中的一种或多种,包括存储单元总线或者存储单元控制器、外围总线、图形加速端口、处理单元或者使用多种总线结构中的任意总线结构的局域总线。

[0145] 电子设备1200也可以与一个或多个外部设备1300(例如键盘、指向设备、蓝牙设备等)通信,还可与一个或者多个使得用户能与该电子设备1200交互的设备通信,和/或与使得该电子设备1200能与一个或多个其它计算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1250进行。并且,电子设备1200还可以通过网络适配器1260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1260通过总线1230与电子设备1200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1200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。

[0146] 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他实施例。本申请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权利要求指出。

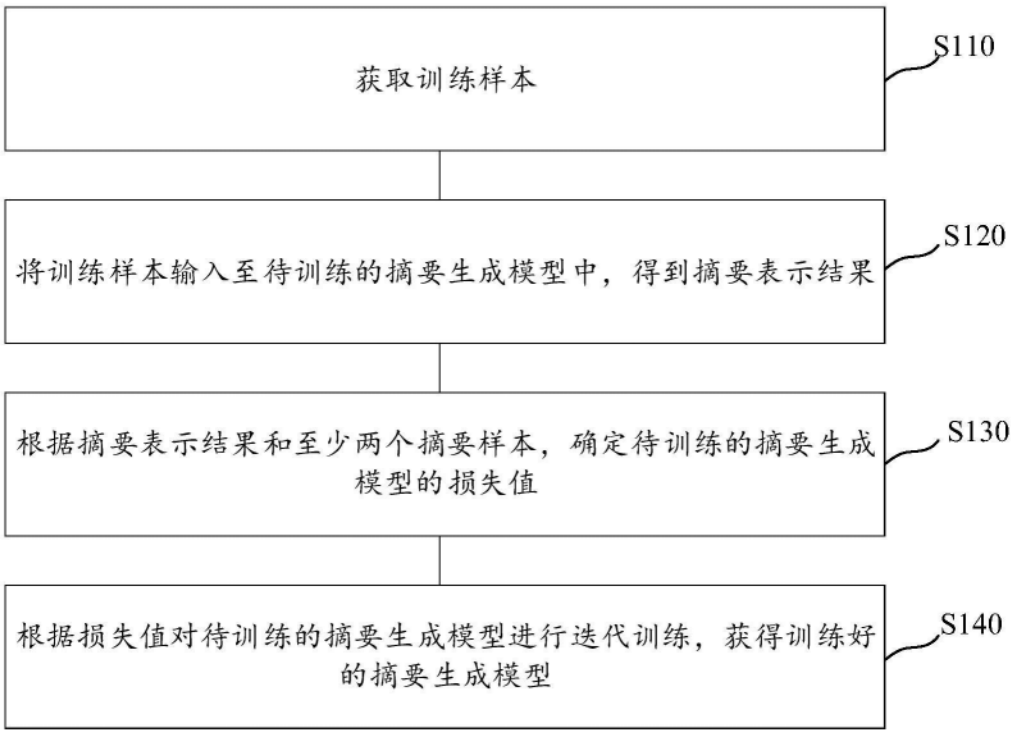

获取训练样本
S110
将训练样本输入至待训练的摘要生成模型中，得到摘要表示结果
S120
根据摘要表示结果和至少两个摘要样本，确定待训练的摘要生成模型的损失值
S130
根据损失值对待训练的摘要生成模型进行迭代训练，获得训练好的摘要生成模型
S140

图1

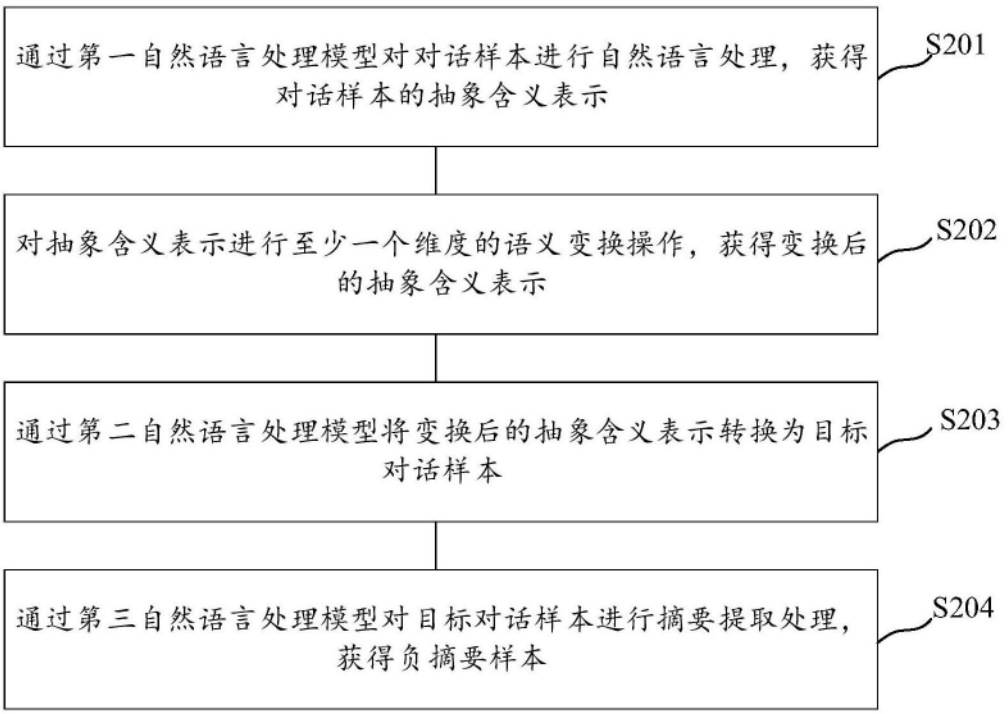

通过第一自然语言处理模型对对话样本进行自然语言处理，获得对话样本的抽象含义表示
S201
对抽象含义表示进行至少一个维度的语义变换操作，获得变换后的抽象含义表示
S202
通过第二自然语言处理模型将变换后的抽象含义表示转换为目标对话样本
S203
通过第三自然语言处理模型对目标对话样本进行摘要提取处理，获得负摘要样本
S204

图2

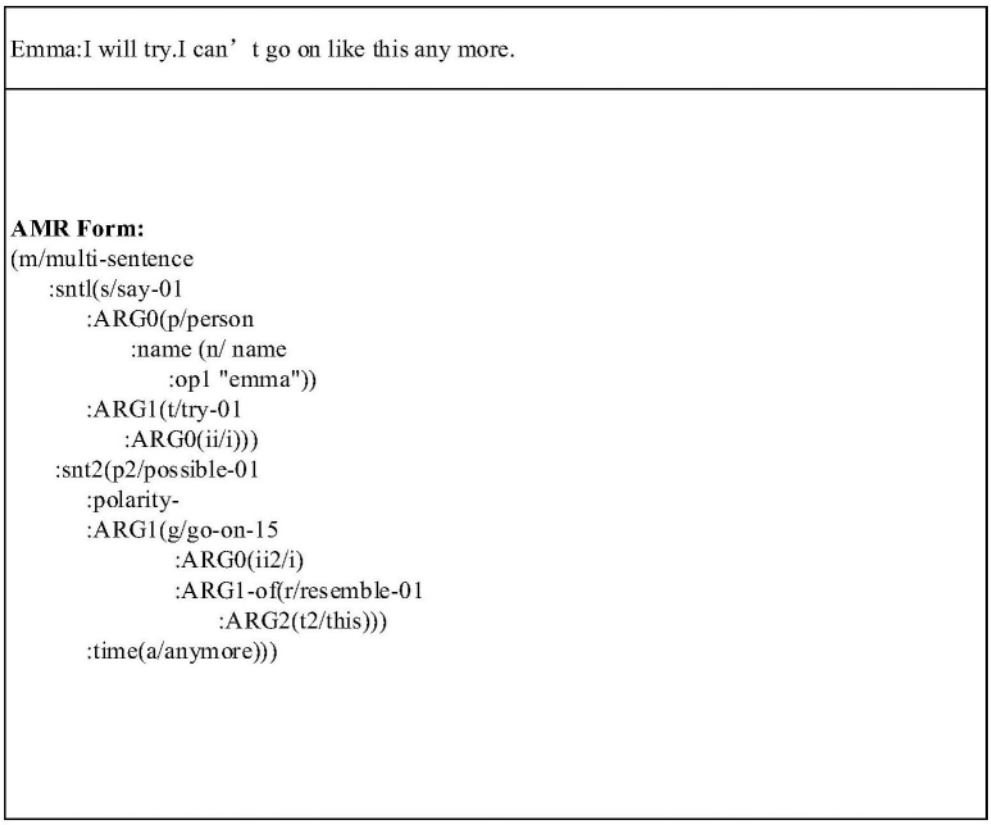

Emma:I will try.I can’ t go on like this any more.
AMR Form:
(m/multi-sentence
:sntl(s/say-01
:ARG0(p/person
:name (n/ name
:op1 "emma"))
:ARG1(t/try-01
:ARG0(ii/i)))
:snt2(p2/possible-01
:polarity-
:ARG1(g/go-on-15
:ARG0(ii2/i)
:ARG1-of(r/resemble-01
:ARG2(t2/this)))
:time(a/anymore)))

图3

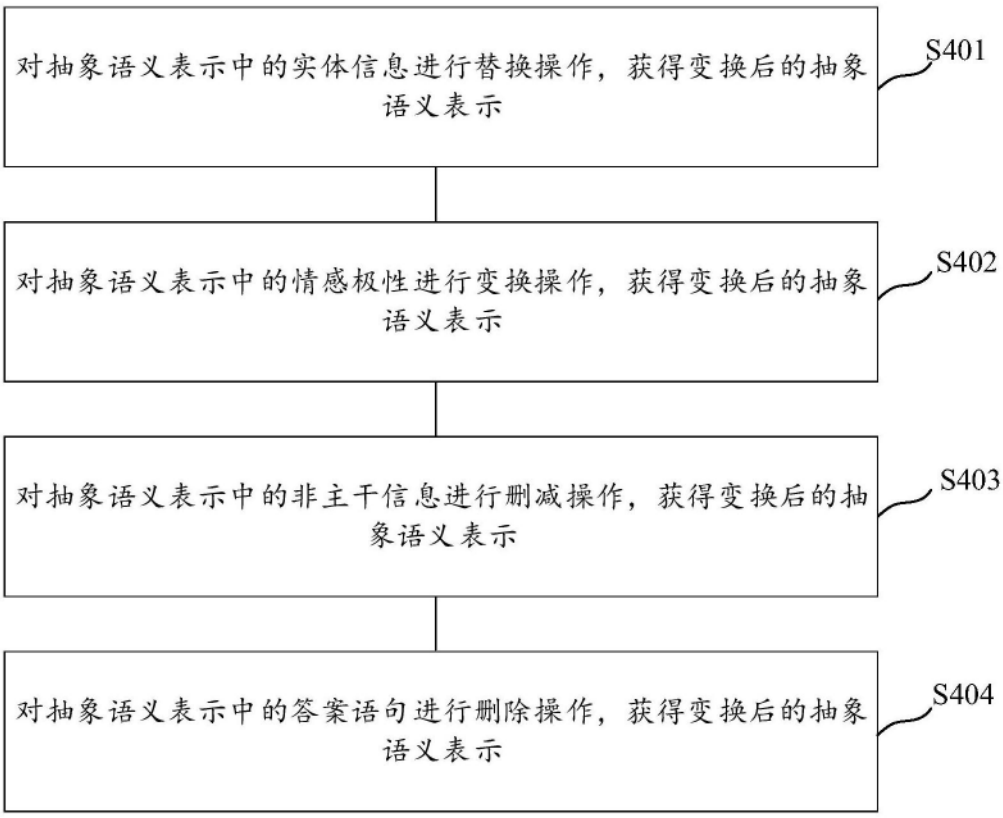

对抽象语义表示中的实体信息进行替换操作，获得变换后的抽象语义表示
S401
对抽象语义表示中的情感极性进行变换操作，获得变换后的抽象语义表示
S402
对抽象语义表示中的非主干信息进行删减操作，获得变换后的抽象语义表示
S403
对抽象语义表示中的答案语句进行删除操作，获得变换后的抽象语义表示
S404

图4

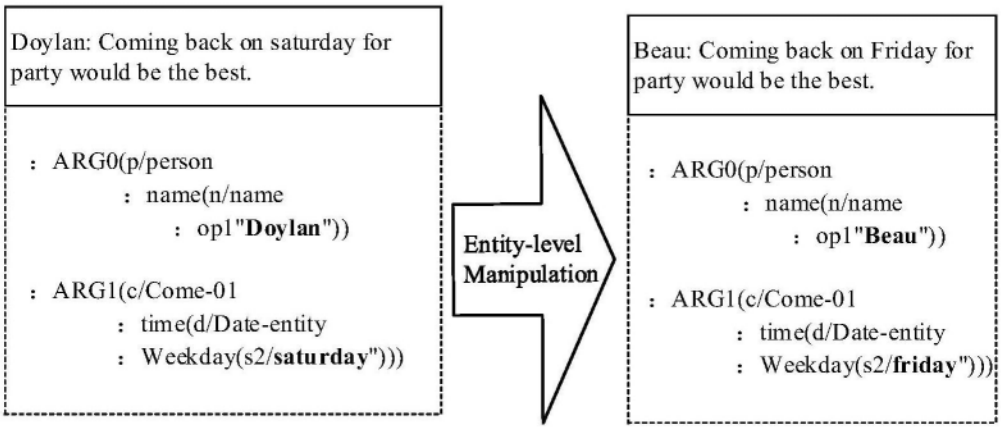

Doylan: Coming back on saturday for party would be the best.
: ARG0(p/person
: name(n/name
: op1"Doylan"))
: ARG1(c/Come-01
: time(d/Date-entity
: Weekday(s2/saturday")))
Entity-level Manipulation
Beau: Coming back on Friday for party would be the best.
: ARG0(p/person
: name(n/name
: op1"Beau"))
: ARG1(c/Come-01
: time(d/Date-entity
: Weekday(s2/friday")))

图5

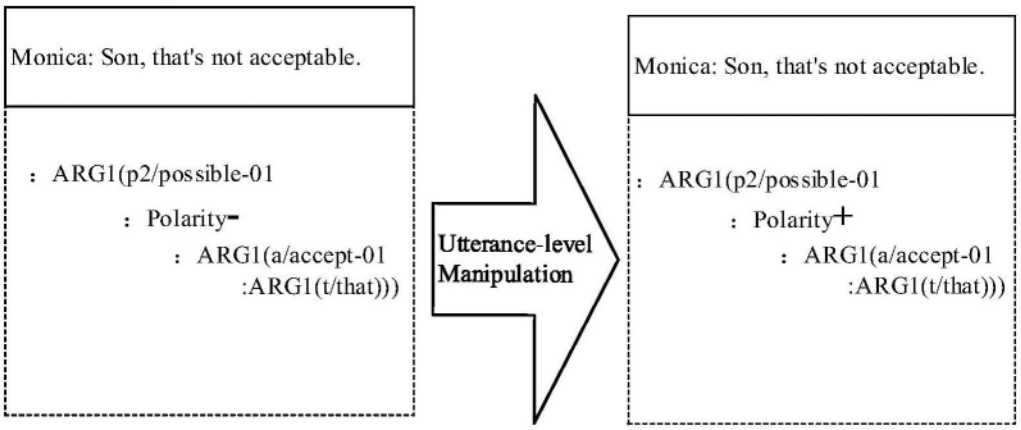

Monica: Son, that's not acceptable.
: ARG1(p2/possible-01
: Polarity-
: ARG1(a/accept-01
:ARG1(t/that)))
Utterance-level Manipulation
Monica: Son, that's not acceptable.
: ARG1(p2/possible-01
: Polarity+
: ARG1(a/accept-01
:ARG1(t/that)))

图6

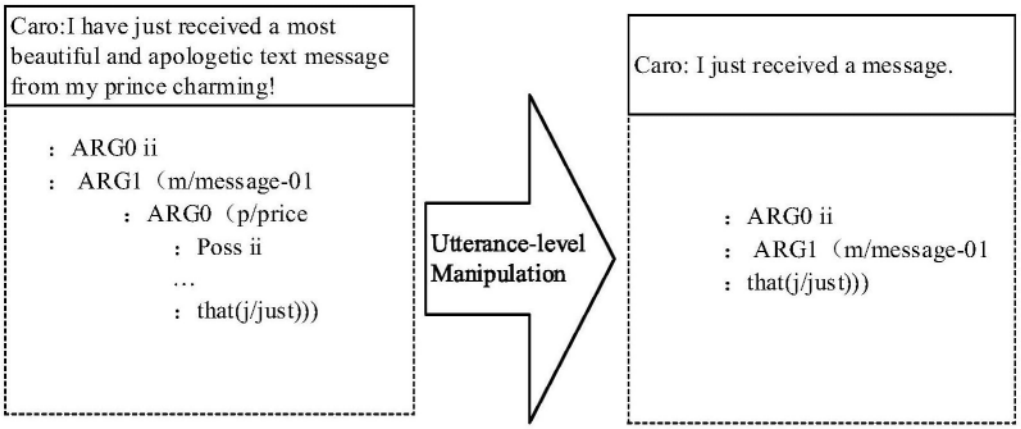

Caro:I have just received a most beautiful and apologetic text message from my prince charming!
: ARG0 ii
: ARG1 (m/message-01
: ARG0 (p/price
: Poss ii
…
: that(j/just)))
Utterance-level Manipulation
Caro: I just received a message.
: ARG0 ii
: ARG1 (m/message-01
: that(j/just)))

图7

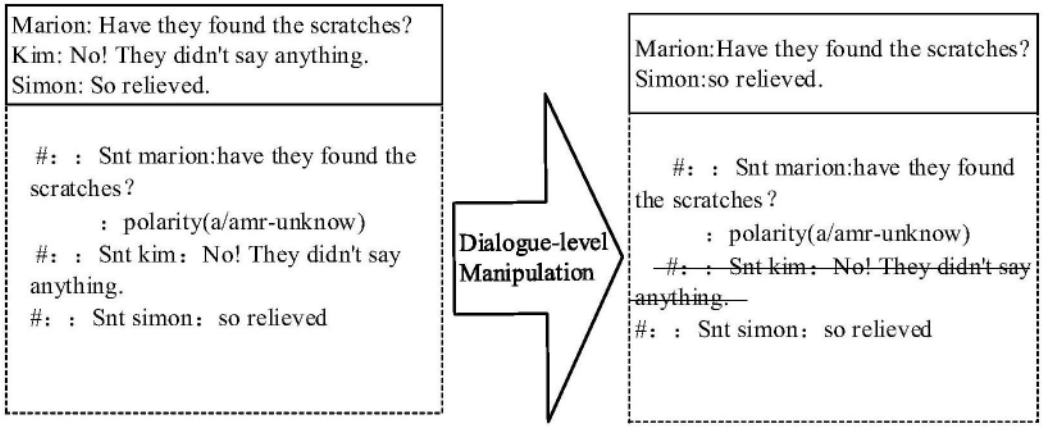

Marion: Have they found the scratches?
Kim: No! They didn't say anything.
Simon: So relieved.
#: : Snt marion:have they found the scratches?
: polarity(a/amr-unknow)
#: : Snt kim: No! They didn't say anything.
#: : Snt simon: so relieved
Dialogue-level Manipulation
Marion:Have they found the scratches?
Simon:so relieved.
#: : Snt marion:have they found the scratches?
: polarity(a/amr-unknow)
#: : Snt kim: No! They didn't say anything.
#: : Snt simon: so relieved

图8

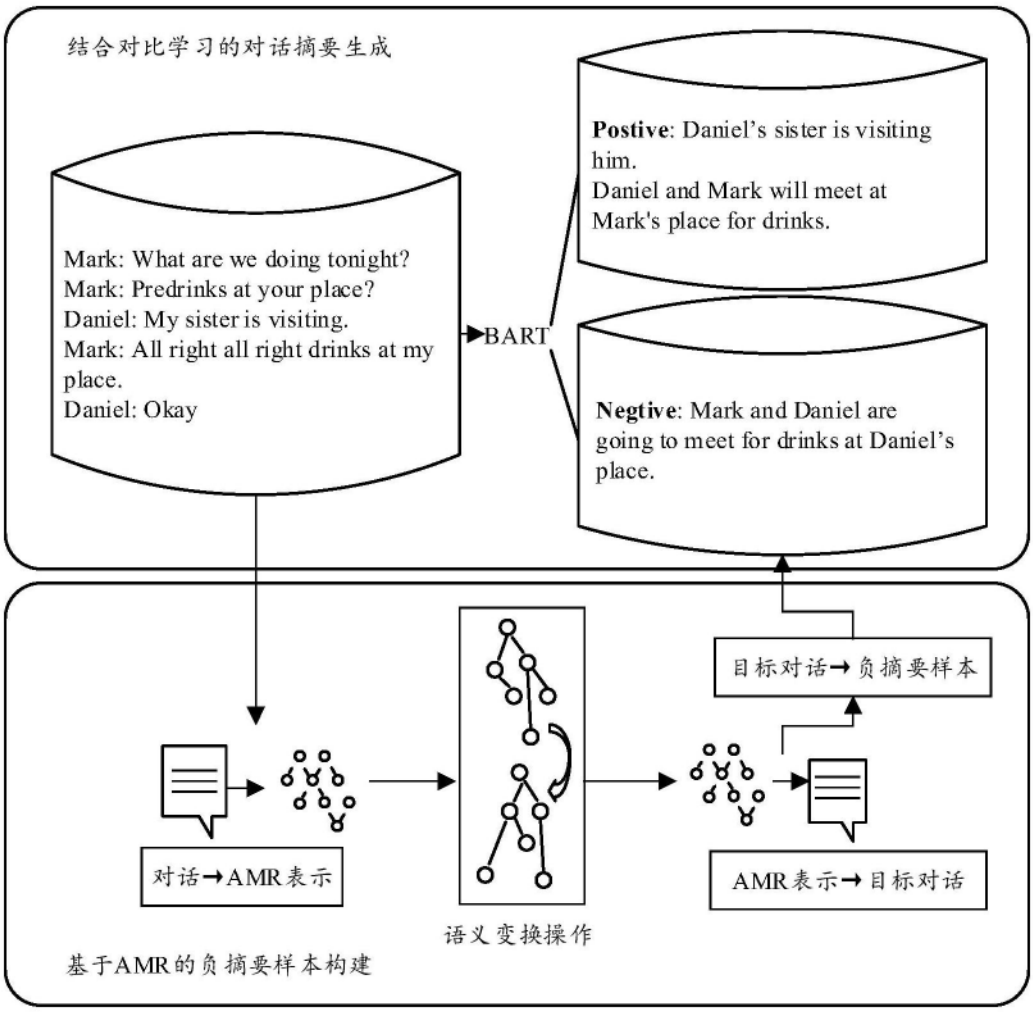

结合对比学习的对话摘要生成
Mark: What are we doing tonight?
Mark: Predrinks at your place?
Daniel: My sister is visiting.
Mark: All right all right drinks at my place.
Daniel: Okay
BART
Postive: Daniel's sister is visiting him.
Daniel and Mark will meet at Mark's place for drinks.
Negtive: Mark and Daniel are going to meet for drinks at Daniel's place.
目标对话→负摘要样本
对话→AMR表示
语义变换操作
AMR表示→目标对话
基于AMR的负摘要样本构建

图9

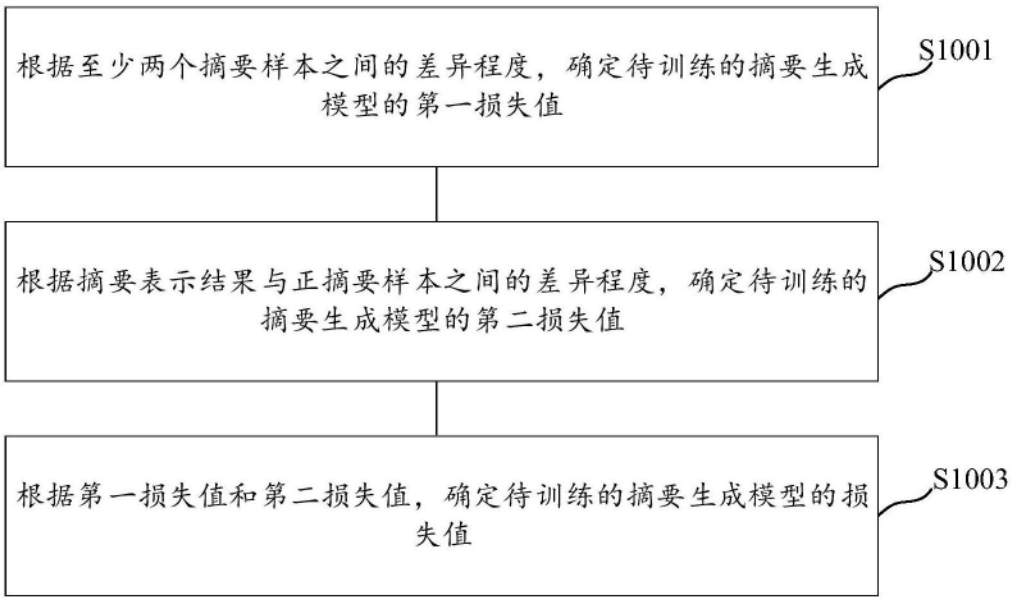

根据至少两个摘要样本之间的差异程度，确定待训练的摘要生成模型的第一损失值
S1001
根据摘要表示结果与正摘要样本之间的差异程度，确定待训练的摘要生成模型的第二损失值
S1002
根据第一损失值和第二损失值，确定待训练的摘要生成模型的损失值
S1003

图10

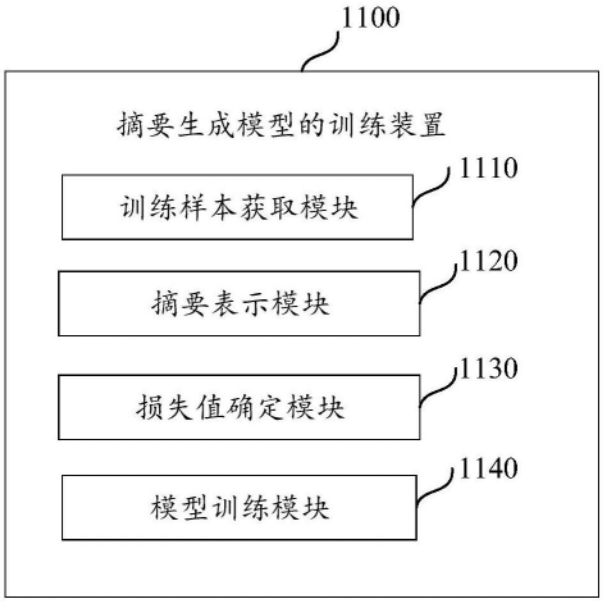

1100
摘要生成模型的训练装置
训练样本获取模块
1110
摘要表示模块
1120
损失值确定模块
1130
模型训练模块
1140

图11

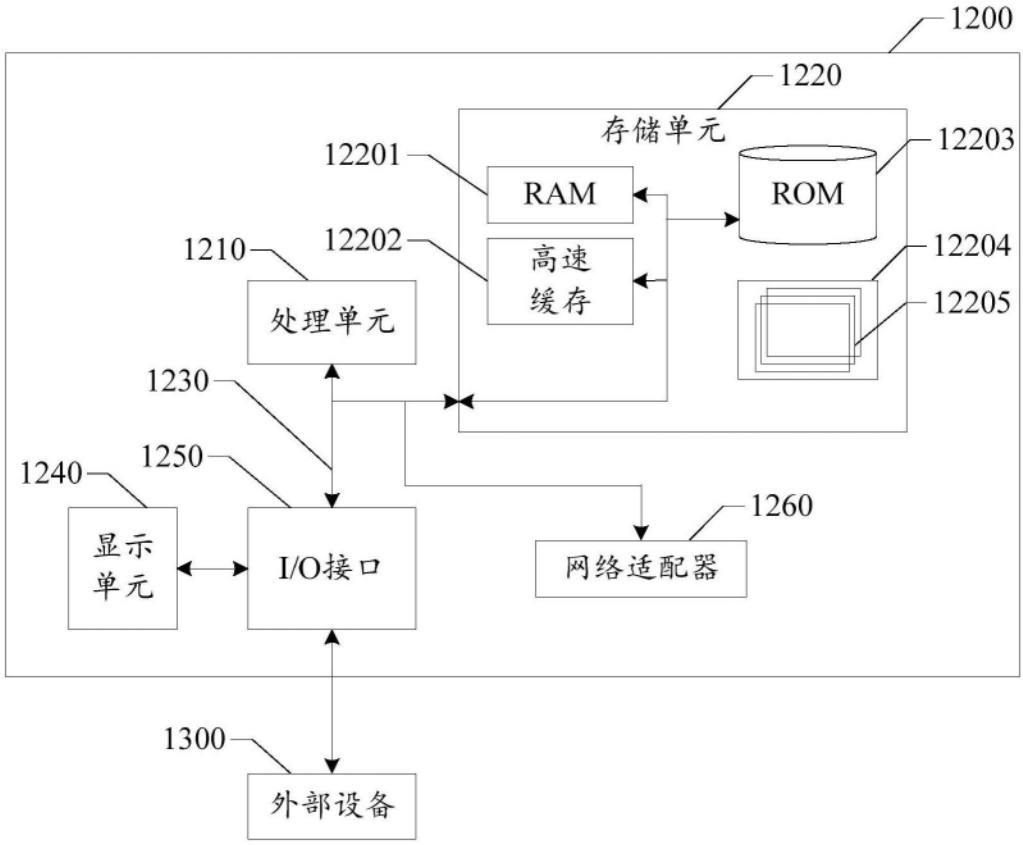

1200
1220
存储单元
12201
RAM
ROM
12203
12202
高速
缓存
12204
1210
处理单元
12205
1230
1250
1240
显示
单元
I/O接口
网络适配器
1260
1300
外部设备

图12