



(21) 申请号 202110699455.7

G06F 18/25 (2023.01)

(22) 申请日 2021.06.23

G06F 18/214 (2023.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 113591908 A

(56) 对比文件

CN 112182166 A, 2021.01.05

CN 109299262 A, 2019.02.01

CN 111382243 A, 2020.07.07

(43) 申请公布日 2021.11.02

US 2005198026 A1, 2005.09.08

(73) 专利权人 北京百度网讯科技有限公司

N. Srinivasan; V. Vaidehi. Timed

地址 100085 北京市海淀区上地十街10号

Coloured Petri Net Model for Misuse

百度大厦2层

Intrusion Detection. IEEE. 2007, 全文.

(72) 发明人 李心雨 韩翠云 施茜 黄佳艳

徐文峰; 杨艳; 张春风. 融合实体上下文特征

裴明

的深度文本语义匹配模型. 武汉大学学报(理学版). 2020, (第05期), 全文.

(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

张敏, 高剑峰, 马少平. 基于链接描述文本及其

专利代理师 王萌

其上下文的Web信息检索. 计算机研究与发展

(51) Int. Cl.

. 2004, (第01期), 全文.

G06F 18/22 (2023.01)

审查员 景京

G06F 18/2411 (2023.01)

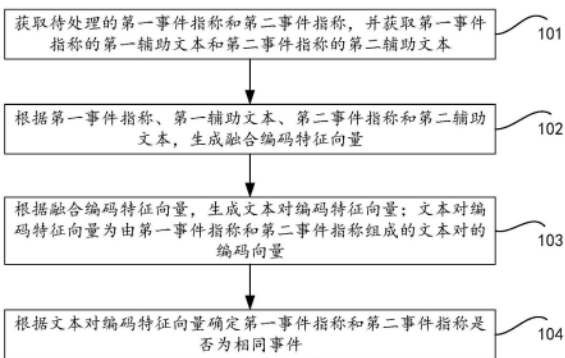
权利要求书5页 说明书13页 附图6页

(54) 发明名称

事件指称匹配方法、模型训练方法、装置和电子设备

(57) 摘要

本公开提供了一种事件指称匹配方法、模型训练方法以及装置, 涉及人工智能领域, 尤其涉及知识图谱领域和深度学习领域。具体实施方案为: 获取待处理的第一、第二事件指称, 并获取第一事件指称的第一辅助文本和第二事件指称的第二辅助文本; 根据第一事件指称、第一辅助文本、第二事件指称和第二辅助文本, 生成融合编码特征向量; 融合编码特征向量为融合了第一事件指称、第一辅助文本、第二事件指称和第二辅助文本的文本信息、结构化信息的编码特征向量; 根据融合编码特征向量, 生成文本对编码特征向量; 文本对编码特征向量为由第一事件指称和第二事件指称组成的文本对的编码向量; 根据文本对编码特征向量确定第一和第二事件指称是否为相同事件。



1. 一种事件指称匹配方法,包括:

获取待处理的第一事件指称和第二事件指称,并获取所述第一事件指称的第一辅助文本和所述第二事件指称的第二辅助文本,其中,所述第一事件指称和所述第二事件指称是分别从多媒体信息来源中挖掘到的简短的事件描述文本,所述第一辅助文本是利用结构化信息为第一事件指称确定有效性辅助文本,所述第二辅助文本是利用结构化信息为第二事件指称确定有效性辅助文本;

根据所述第一事件指称、所述第一辅助文本、所述第二事件指称和所述第二辅助文本,生成融合编码特征向量;所述融合编码特征向量为融合了所述第一事件指称、所述第一辅助文本、所述第二事件指称和所述第二辅助文本的文本信息、结构化信息的编码特征向量;

根据所述融合编码特征向量,生成文本对编码特征向量;所述文本对编码特征向量为由所述第一事件指称和第二事件指称组成的文本对的编码向量;

根据所述文本对编码特征向量确定所述第一事件指称和所述第二事件指称是否为相同事件;

所述根据所述第一事件指称、所述第一辅助文本、所述第二事件指称和所述第二辅助文本,生成融合编码特征向量,包括:

获取预先经过训练的事件指称匹配模型;所述事件指称匹配模型包括结构化信息输入层;

将所述第一事件指称、所述第一辅助文本、所述第二事件指称和所述第二辅助文本进行拼接,将拼接后的文本作为所述事件指称匹配模型的输入;

通过所述结构化信息输入层对输入文本分别进行文本编码、分句编码、触发词编码和实体编码,并将文本编码结果和分句编码结果、触发词编码结果和实体编码结果进行融合处理,获得所述融合编码特征向量。

2. 根据权利要求1所述的方法,其中,所述获取所述第一事件指称的第一辅助文本,包括:

从所述第一事件指称所属的第一多媒体文本中选取第一候选事件句;

获取所述第一多媒体文本的标题信息;

对所述第一事件指称、所述第一候选事件句和所述第一多媒体文本的标题信息分别进行实体识别;

根据实体识别结果,从所述第一候选事件句和所述第一多媒体文本的标题信息中选择一个作为所述第一事件指称的第一辅助文本。

3. 根据权利要求1所述的方法,其中,获取所述第二事件指称的第二辅助文本,包括:

从所述第二事件指称所属的第二多媒体文本中选取第二候选事件句;

获取所述第二多媒体文本的标题信息;

对所述第二事件指称、所述第二候选事件句和所述第二多媒体文本的标题信息分别进行实体识别;

根据实体识别结果,从所述第二候选事件句和所述第二多媒体文本的标题信息中选择一个作为所述第二事件指称的第二辅助文本。

4. 根据权利要求1所述的方法,其中,所述事件指称匹配模型还包括编码层;所述根据所述融合编码特征向量,生成文本对编码特征向量,包括:

将所述融合编码特征向量作为所述编码层的输入,通过所述编码层采用所述融合编码特征向量进行事件信息交互,并对事件信息进行编码,获得所述文本对编码特征向量。

5.根据权利要求1所述的方法,其中,所述事件指称匹配模型还包括输出层;所述根据所述文本对编码特征向量确定所述第一事件指称和所述第二事件指称是否为相同事件,包括:

将所述文本对编码特征向量作为所述输出层的输入,通过所述输出层对所述文本对编码特征向量进行预测;

获取所述输出层输出的预测结果,并根据所述预测结果确定所述第一事件指称和所述第二事件指称是否为相同事件。

6.根据权利要求4至5中任一项所述的方法,通过以下方式预先训练所述事件指称匹配模型:

获取事件指称对样本和所述事件指称对样本对应的标签;所述事件指称对样本包括第一事件指称样本和第二事件指称样本;所述标签用于指示所述第一事件指称样本与所述第二事件指称样本是否指代相同事件;

获取所述第一事件指称样本的第一辅助文本样本和所述第二事件指称样本的第二辅助文本样本;

构建事件指称匹配模型;所述事件指称匹配模型包括结构化信息输入层、编码层和输出层;

利用所述第一事件指称样本、所述第一辅助文本样本、所述第二事件指称样本、所述第二辅助文本样本和所述标签,对所述事件指称匹配模型进行训练。

7.一种事件指称匹配模型的训练方法,包括:

获取事件指称对样本和所述事件指称对样本对应的标签;所述事件指称对样本包括第一事件指称样本和第二事件指称样本;所述标签用于指示所述第一事件指称样本与所述第二事件指称样本是否指代相同事件;

获取所述第一事件指称样本的第一辅助文本样本和所述第二事件指称样本的第二辅助文本样本,其中,所述第一事件指称和所述第二事件指称是分别从多媒体信息来源中挖掘到的简短的事件描述文本,所述第一辅助文本是利用结构化信息为第一事件指称确定有效性辅助文本,所述第二辅助文本是利用结构化信息为第二事件指称确定有效性辅助文本;

构建事件指称匹配模型;所述事件指称匹配模型包括结构化信息输入层、编码层和输出层;

利用所述第一事件指称样本、所述第一辅助文本样本、所述第二事件指称样本、所述第二辅助文本样本和所述标签,对所述事件指称匹配模型进行训练;

所述利用所述第一事件指称样本、所述第一辅助文本样本、所述第二事件指称样本、所述第二辅助文本样本和所述标签,对所述事件指称匹配模型进行训练,包括:

将所述第一事件指称样本、所述第一辅助文本样本、所述第二事件指称样本和所述第二辅助文本样本进行拼接,并将拼接后的文本和所述标签作为所述事件指称匹配模型的输入;

通过所述结构化信息输入层对输入文本分别进行文本编码、分句编码、触发词编码和

实体编码,并将文本编码结果和分句编码结果、触发词编码结果和实体编码结果进行融合处理,获得融合编码特征向量;

通过所述编码层采用所述融合编码特征向量进行事件信息交互,并对事件信息进行编码,获得文本对编码特征向量;

通过所述输出层对所述文本对编码特征向量进行预测,获取所述输出层输出的预测结果;

计算所述预测结果和所述标签的损失值,并根据所述损失值训练所述事件指称匹配模型。

8. 根据权利要求7所述的方法,其中,所述获取所述第一事件指称样本的第一辅助文本样本,包括:

从所述第一事件指称样本所属的第一多媒体文本样本中选取第一候选事件句样本;

获取所述第一多媒体文本样本的标题信息;

对所述第一事件指称样本、所述第一候选事件句样本和所述第一多媒体文本样本的标题信息分别进行实体识别;

根据实体识别结果,从所述第一候选事件句样本和所述第一多媒体文本样本的标题信息中选择其中一个作为所述第一事件指称样本的第一辅助文本样本。

9. 一种事件指称匹配装置,包括:

获取模块,用于获取待处理的第一事件指称和第二事件指称,并获取所述第一事件指称的第一辅助文本和所述第二事件指称的第二辅助文本,其中,所述第一事件指称和所述第二事件指称是分别从多媒体信息来源中挖掘到的简短的事件描述文本,所述第一辅助文本是利用结构化信息为第一事件指称确定有效性辅助文本,所述第二辅助文本是利用结构化信息为第二事件指称确定有效性辅助文本;

第一生成模块,用于根据所述第一事件指称、所述第一辅助文本、所述第二事件指称和所述第二辅助文本,生成融合编码特征向量;所述融合编码特征向量为融合了所述第一事件指称、所述第一辅助文本、所述第二事件指称和所述第二辅助文本的文本信息、结构化信息的编码特征向量;

第二生成模块,用于根据所述融合编码特征向量,生成文本对编码特征向量;所述文本对编码特征向量为由所述第一事件指称和第二事件指称组成的文本对的编码向量;

确定模块,用于根据所述文本对编码特征向量确定所述第一事件指称和所述第二事件指称是否为相同事件;所述第一生成模块具体用于:

获取预先经过训练的事件指称匹配模型;所述事件指称匹配模型包括结构化信息输入层;

将所述第一事件指称、所述第一辅助文本、所述第二事件指称和所述第二辅助文本进行拼接,将拼接后的文本作为所述事件指称匹配模型的输入;

通过所述结构化信息输入层对输入文本分别进行文本编码、分句编码、触发词编码和实体编码,并将文本编码结果和分句编码结果、触发词编码结果和实体编码结果进行融合处理,获得所述融合编码特征向量。

10. 根据权利要求9所述的装置,其中,所述获取模块具体用于:

从所述第一事件指称所属的第一多媒体文本中选取第一候选事件句;

获取所述第一多媒体文本的标题信息；

对所述第一事件指称、所述第一候选事件句和所述第一多媒体文本的标题信息分别进行实体识别；

根据实体识别结果，从所述第一候选事件句和所述第一多媒体文本的标题信息中选择一个作为所述第一事件指称的第一辅助文本。

11. 根据权利要求9所述的装置，其中，所述获取模块具体用于：

从所述第二事件指称所属的第二多媒体文本中选取第二候选事件句；

获取所述第二多媒体文本的标题信息；

对所述第二事件指称、所述第二候选事件句和所述第二多媒体文本的标题信息分别进行实体识别；

根据实体识别结果，从所述第二候选事件句和所述第二多媒体文本的标题信息中选择一个作为所述第二事件指称的第二辅助文本。

12. 根据权利要求9所述的装置，其中，所述事件指称匹配模型还包括编码层；所述第二生成模块具体用于：

将所述融合编码特征向量作为所述编码层的输入，通过所述编码层采用所述融合编码特征向量进行事件信息交互，并对事件信息进行编码，获得所述文本对编码特征向量。

13. 根据权利要求9所述的装置，其中，所述事件指称匹配模型还包括输出层；所述确定模块具体用于：

将所述文本对编码特征向量作为所述输出层的输入，通过所述输出层对所述文本对编码特征向量进行预测；

获取所述输出层输出的预测结果，并根据所述预测结果确定所述第一事件指称和所述第二事件指称是否为相同事件。

14. 根据权利要求12至13中任一项所述的装置，还包括模型训练模块，所述模型训练模块具体用于：

获取事件指称对样本和所述事件指称对样本对应的标签；所述事件指称对样本包括第一事件指称样本和第二事件指称样本；所述标签用于指示所述第一事件指称样本与所述第二事件指称样本是否指代相同事件；

获取所述第一事件指称样本的第一辅助文本样本和所述第二事件指称样本的第二辅助文本样本；

构建事件指称匹配模型；所述事件指称匹配模型包括结构化信息输入层、编码层和输出层；

利用所述第一事件指称样本、所述第一辅助文本样本、所述第二事件指称样本、所述第二辅助文本样本和所述标签，对所述事件指称匹配模型进行训练。

15. 一种事件指称匹配模型的训练装置，包括：

第一获取模块，用于获取事件指称对样本和所述事件指称对样本对应的标签；所述事件指称对样本包括第一事件指称样本和第二事件指称样本；所述标签用于指示所述第一事件指称样本与所述第二事件指称样本是否指代相同事件；

第二获取模块，用于获取所述第一事件指称样本的第一辅助文本样本和所述第二事件指称样本的第二辅助文本样本，其中，所述第一事件指称和所述第二事件指称是分别从多

媒体信息来源中挖掘到的简短的事件描述文本,所述第一辅助文本是利用结构化信息为第一事件指称确定有效性辅助文本,所述第二辅助文本是利用结构化信息为第二事件指称确定有效性辅助文本;

构建模块,用于构建事件指称匹配模型;所述事件指称匹配模型包括结构化信息输入层、编码层和输出层;

训练模块,用于利用所述第一事件指称样本、所述第一辅助文本样本、所述第二事件指称样本、所述第二辅助文本样本和所述标签,对所述事件指称匹配模型进行训练;

所述训练模块具体用于:

将所述第一事件指称样本、所述第一辅助文本样本、所述第二事件指称样本和所述第二辅助文本样本进行拼接,并将拼接后的文本和所述标签作为所述事件指称匹配模型的输入;

通过所述结构化信息输入层对输入文本分别进行文本编码、分句编码、触发词编码和实体编码,并将文本编码结果和分句编码结果、触发词编码结果和实体编码结果进行融合处理,获得融合编码特征向量;

通过所述编码层采用所述融合编码特征向量进行事件信息交互,并对事件信息进行编码,获得文本对编码特征向量;

通过所述输出层对所述文本对编码特征向量进行预测,获取所述输出层输出的预测结果;

计算所述预测结果和所述标签的损失值,并根据所述损失值训练所述事件指称匹配模型。

16. 根据权利要求15所述的装置,其中,所述第二获取模块具体用于:

从所述第一事件指称样本所属的第一多媒体文本样本中选取第一候选事件句样本;

获取所述第一多媒体文本样本的标题信息;

对所述第一事件指称样本、所述第一候选事件句样本和所述第一多媒体文本样本的标题信息分别进行实体识别;

根据实体识别结果,从所述第一候选事件句样本和所述第一多媒体文本样本的标题信息中选择其中一个作为所述第一事件指称样本的第一辅助文本样本。

17. 一种电子设备,包括:

至少一个处理器;以及

与所述至少一个处理器通信连接的存储器;其中,

所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行权利要求1至8中任一项所述的方法。

18. 一种存储有计算机指令的非瞬时计算机可读存储介质,其中,所述计算机指令用于使所述计算机执行权利要求1至8中任一项所述的方法。

事件指称匹配方法、模型训练方法、装置和电子设备

技术领域

[0001] 本公开涉及人工智能领域,尤其涉及知识图谱领域和深度学习领域,特别的涉及一种事件指称匹配方法、事件指称匹配模型的训练方法、装置、电子设备和存储介质,可用于事件指称归一场景。

背景技术

[0002] 事件指称归一是判别两个事件指称是否对应现实世界中的相同事件的任务,其中事件指称是指从新闻标题等多媒体信息来源当中挖掘到的相对简短的事件描述文本。

[0003] 事件指称归一(又称为事件指称匹配)目前普遍采用基于文本匹配的事件指称归一技术进行解决,这种方案下,事件指称归一问题会被建模成普通的文本匹配任务,利用基于深度学习的文本匹配方法判别任意给定的两个事件指称是否指代相同事件。

发明内容

[0004] 本公开提供了一种用于事件指称归一场景的事件指称匹配方法、事件指称匹配模型的训练方法、装置、电子设备和存储介质。

[0005] 根据本公开的第一方面,提供了一种事件指称匹配方法,包括:

[0006] 获取待处理的第一事件指称和第二事件指称,并获取所述第一事件指称的第一辅助文本和所述第二事件指称的第二辅助文本;

[0007] 根据所述第一事件指称、所述第一辅助文本、所述第二事件指称和所述第二辅助文本,生成融合编码特征向量;所述融合编码特征向量为融合了所述第一事件指称、所述第一辅助文本、所述第二事件指称和所述第二辅助文本的文本信息、结构化信息的编码特征向量;

[0008] 根据所述融合编码特征向量,生成文本对编码特征向量;所述文本对编码特征向量为由所述第一事件指称和第二事件指称组成的文本对的编码向量;

[0009] 根据所述文本对编码特征向量确定所述第一事件指称和所述第二事件指称是否为相同事件。

[0010] 根据本公开的第二方面,提供了一种事件指称匹配模型的训练方法,包括:

[0011] 获取事件指称对样本和所述事件指称对样本对应的标签;所述事件指称对样本包括第一事件指称样本和第二事件指称样本;所述标签用于指示所述第一事件指称样本与所述第二事件指称样本是否指代相同事件;

[0012] 获取所述第一事件指称样本的第一辅助文本样本和所述第二事件指称样本的第二辅助文本样本;

[0013] 构建事件指称匹配模型;所述事件指称匹配模型包括结构化信息输入层、编码层和输出层;

[0014] 利用所述第一事件指称样本、所述第一辅助文本样本、所述第二事件指称样本、所述第二辅助文本样本和所述标签,对所述事件指称匹配模型进行训练。

[0015] 根据本公开的第三方法,提供了一种事件指称匹配装置,包括:

[0016] 获取模块,用于获取待处理的第一事件指称和第二事件指称,并获取所述第一事件指称的第一辅助文本和所述第二事件指称的第二辅助文本;

[0017] 第一生成模块,用于根据所述第一事件指称、所述第一辅助文本、所述第二事件指称和所述第二辅助文本,生成融合编码特征向量;所述融合编码特征向量为融合了所述第一事件指称、所述第一辅助文本、所述第二事件指称和所述第二辅助文本的文本信息、结构化信息的编码特征向量;

[0018] 第二生成模块,用于根据所述融合编码特征向量,生成文本对编码特征向量;所述文本对编码特征向量为由所述第一事件指称和第二事件指称组成的文本对的编码向量;

[0019] 确定模块,用于根据所述文本对编码特征向量确定所述第一事件指称和所述第二事件指称是否为相同事件。

[0020] 根据本公开的第四方面,提供了一种事件指称匹配模型的训练装置,包括:

[0021] 第一获取模块,用于获取事件指称对样本和所述事件指称对样本对应的标签;所述事件指称对样本包括第一事件指称样本和第二事件指称样本;所述标签用于指示所述第一事件指称样本与所述第二事件指称样本是否指代相同事件;

[0022] 第二获取模块,用于获取所述第一事件指称样本的第一辅助文本样本和所述第二事件指称样本的第二辅助文本样本;

[0023] 构建模块,用于构建事件指称匹配模型;所述事件指称匹配模型包括结构化信息输入层、编码层和输出层;

[0024] 训练模块,用于利用所述第一事件指称样本、所述第一辅助文本样本、所述第二事件指称样本、所述第二辅助文本样本和所述标签,对所述事件指称匹配模型进行训练。

[0025] 根据本公开的第五方面,提供了一种电子设备,包括:

[0026] 至少一个处理器;以及

[0027] 与所述至少一个处理器通信连接的存储器;其中,

[0028] 所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行前述第一方面所述的事件指称匹配方法。

[0029] 根据本公开的第六方面,提供另一种电子设备,包括:

[0030] 至少一个处理器;以及

[0031] 与所述至少一个处理器通信连接的存储器;其中,

[0032] 所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行前述第二方面所述的事件指称匹配模型的训练方法。

[0033] 根据本公开的第七方面,提供了一种存储有计算机指令的非瞬时计算机可读存储介质,其中,所述计算机指令用于使所述计算机执行前述第一方面所述的事件指称匹配方法。

[0034] 根据本公开的第八方面,提供了另一种存储有计算机指令的非瞬时计算机可读存储介质,其中,所述计算机指令用于使所述计算机执行前述第二方面所述的事件指称匹配模型的训练方法。

[0035] 根据本公开的第九方面,提供了一种计算机程序产品,包括计算机程序,所述计算机程序在被处理器执行时实现根据前述第一方面所述的事件指称匹配方法。

[0036] 根据本公开的第十方面,提供了另一种计算机程序产品,包括计算机程序,所述计算机程序在被处理器执行时实现根据前述第二方面所述的事件指称匹配模型的训练方法。

[0037] 根据本公开的技术方案,可利用结构化信息为事件指称确定有效性辅助文本,大大降低了事件指称描述不全的现象;在生成事件对的文本对编码特征向量时,融合了两个事件的事件指称、辅助文本和结构化信息的编码特征向量,进而利用融合了事件指称、辅助文本和结构化信息的编码特征向量预测两个事件是否指代相同事件,可以有效提升事件指称归一任务的判别效果,即提升事件指称匹配的判别效果,可以推动事件检测、事件库构建等产品的实际落地。

[0038] 应当理解,本部分所描述的内容并非旨在标识本公开的实施例的关键或重要特征,也不用于限制本公开的范围。本公开的其它特征将通过以下的说明书而变得容易理解。

附图说明

[0039] 附图用于更好地理解本方案,不构成对本公开的限定。其中:

[0040] 图1为本公开实施例所提供的一种事件指称匹配方法的流程图;

[0041] 图2是根据本公开实施例的获取第一辅助文本的流程图;

[0042] 图3是根据本公开实施例的获取第二辅助文本的流程图;

[0043] 图4为本公开实施例所提供的另一种事件指称匹配方法的流程图;

[0044] 图5为本公开实施例所提供的事件指称匹配模型的示例图;

[0045] 图6为本公开实施例所提供的一种事件指称匹配模型的训练方法的流程图;

[0046] 图7为本公开实施例所提供的一种事件指称匹配装置的结构示意图;

[0047] 图8为本公开实施例所提供的另一种事件指称匹配装置的结构示意图;

[0048] 图9为本公开实施例所提供的一种事件指称匹配模型的训练装置的结构示意图;

[0049] 图10是本公开实施例所提供的一种电子设备的框图。

具体实施方式

[0050] 以下结合附图对本公开的示范性实施例做出说明,其中包括本公开实施例的各种细节以助于理解,应当将它们认为仅仅是示范性的。因此,本领域普通技术人员应当认识到,可以对这里描述的实施例做出各种改变和修改,而不会背离本公开的范围和精神。同样,为了清楚和简明,以下的描述中省略了对公知功能和结构的描述。

[0051] 需要说明的是,事件指称归一是判别两个事件指称是否对应现实世界中的相同事件的任务,其中事件指称是指从新闻标题等多媒体信息来源当中挖掘到的相对简短的事件描述文本。

[0052] 由于事件指称归一任务在本质上仍然是一种文本匹配任务,因此可以直接将事件指称归一任务建模为一个简单的文本匹配任务,但是这种方法严重受限于事件指称的质量,如事件指称对事件描述的完整性等。为解决该问题,本发明提出一种引入结构化信息的事件指称匹配方法、事件指称匹配模型的训练方法、装置、电子设备和存储介质,通过一种结构化信息检测策略完善事件描述,同时利用实体标签等结构化信息辅助匹配模型建模,

提升最终的事件指称归一效果。

[0053] 图1为本公开实施例所提供的一种事件指称匹配方法的流程图。需要说明的是,本公开实施例的事件指称匹配方法可应用于本公开实施例的事件指称匹配装置,该事件指称匹配装置可配置于电子设备上。

[0054] 如图1所示,本公开实施例的事件指称匹配方法可以包括:

[0055] 步骤101,获取待处理的第一事件指称和第二事件指称,并获取第一事件指称的第一辅助文本和第二事件指称的第二辅助文本。

[0056] 作为一种示例,第一事件指称和第二事件指称可以是分别从新闻等多媒体资讯信息来源当中挖掘到的相对简短的事件描述文本。

[0057] 在本公开实施例中,可利用结构化信息为第一事件指称确定有效性辅助文本,即得到第一事件指称的第一辅助文本;同理,可利用结构化信息为第二事件指称确定有效性辅助文本,即得到第二事件指称的第二辅助文本。由此,本公开可通过结构化识别技术为事件指称确定具有辅助性的描述文本,大大降低事件指称描述不全的现象。

[0058] 步骤102,根据第一事件指称、第一辅助文本、第二事件指称和第二辅助文本,生成融合编码特征向量。

[0059] 其中,在本公开实施例中,该融合编码特征向量可以为融合了第一事件指称、第一辅助文本、第二事件指称和第二辅助文本的文本信息、结构化信息的编码特征向量。

[0060] 在一种实现方式中,可将第一事件指称、第一辅助文本、第二事件指称和第二辅助文本进行拼接,并利用深度学习技术,对拼接后的文本进行向量表示,以得到与拼接后的文本对应的融合编码特征向量。

[0061] 作为一种示例,可将第一事件指称、第一辅助文本、第二事件指称和第二辅助文本进行拼接之后作为模型的输入,其中,对于每一个事件,将事件指称+辅助文本作为一个事件的最终描述文本。该模型的输入层包括文本编码、分句编码、触发词编码、实体编码这四种信息,其中文本编码表征当前字的字义信息,分句编码表征当前字所属事件,触发词编码表征当前字是否为触发词的组成部分,实体编码表征当前字所属实体的实体类别。这里通过相加的方式融合文本信息与结构化信息得到融合编码特征向量。

[0062] 步骤103,根据融合编码特征向量,生成文本对编码特征向量;文本对编码特征向量为由第一事件指称和第二事件指称组成的文本对的编码向量。

[0063] 在一种实现方式中,可基于深度学习的编码方式,对融合编码特征向量进行事件信息交互、事件信息进行编码,得到交互之后的文本对编码特征向量。

[0064] 步骤104,根据文本对编码特征向量确定第一事件指称和第二事件指称是否为相同事件。

[0065] 在一种实现方式中,可基于MLP(Multilayer Perceptron,多层线性感知机)等人工神经网络对文本对编码特征向量进行预测,以判断第一事件指称和第二事件指称是否为相同事件。

[0066] 根据本公开实施例的事件指称匹配方法,可利用结构化信息为事件指称确定有效性辅助文本,大大降低了事件指称描述不全的现象;在生成事件对的文本对编码特征向量时,融合了两个事件的事件指称、辅助文本和结构化信息的编码特征向量,进而利用融合了事件指称、辅助文本和结构化信息的编码特征向量预测两个事件是否指代相同事件,可以

有效提升事件指称归一任务的判别效果,即提升事件指称匹配的判别效果,可以推动事件检测、事件库构建等产品的实际落地。

[0067] 需要说明的是,在本公开实施例中,可利用事件句召回、结构化信息识别技术与一定的策略为第一事件指称确定具有辅助性的描述文本。可选地,如图2所示,所述获取第一事件指称的第一辅助文本的实现方式可包括:

[0068] 步骤201,从第一事件指称所属的第一多媒体文本中选取第一候选事件句。

[0069] 可选地,对于给定的第一事件指称,可利用策略从对应的多媒体文本中选取候选的事件句。作为一种示例,可对对应的多媒体文本进行分句,并选取对事件指称覆盖度最高的句子作为该事件指称的候选事件句。

[0070] 举例而言,以多媒体为新闻为例,对于给定的事件指称,可先从新闻等资讯自选文本进行分句,并从分句得到的句子中选取对该事件指称覆盖度最高的句子作为该事件指称的候选事件句。

[0071] 步骤202,获取第一多媒体文本的标题信息。

[0072] 可选地,以多媒体为新闻为例,对于给定的事件指称,可获取该新闻等资讯的标题信息。

[0073] 步骤203,对第一事件指称、第一候选事件句和第一多媒体文本的标题信息分别进行实体识别。

[0074] 可选地,对于第一事件指称、第一候选事件句和第一多媒体文本的标题信息这三类文本,可同时利用通用的实体识别工具识别三类文本当中的时间、地点、人物、机构、触发词等重要的结构化信息,即可得到第一事件指称的实体识别结果,第一候选事件句的实体识别结果和标题信息的实体识别结果,其中,该实体识别结果可包括时间、地点、人物、机构、触发词等中至少一种重要的结构化信息。

[0075] 步骤204,根据实体识别结果,从第一候选事件句和第一多媒体文本的标题信息中选择其中一个作为第一事件指称的第一辅助文本。

[0076] 可选地,根据第一事件指称的实体识别结果、第一候选事件句的实体识别结果和标题信息的实体识别结果,比较标题信息与第一候选事件句中,相对于第一事件指称是否出现新增结构化信息,并从标题信息与第一候选事件句中选取包含最多新增结构化信息的文本作为第一事件指称的第一辅助文本。

[0077] 由此可见,本公开利用事件句召回和结构化信息识别技术,为事件指称确定具有辅助性的描述文本,大大降低了事件指称描述不全的现象。

[0078] 需要说明的是,在本公开实施例中,可利用事件句召回、结构化信息识别技术与一定的策略为第二事件指称确定具有辅助性的描述文本。可选地,如图3所示,所述获取第二事件指称的第二辅助文本的实现方式可包括:

[0079] 步骤301,从第二事件指称所属的第二多媒体文本中选取第二候选事件句。

[0080] 可选地,对于给定的第二事件指称,可利用策略从对应的多媒体文本中选取候选的事件句。作为一种示例,可对对应的多媒体文本进行分句,并选取对事件指称覆盖度最高的句子作为该事件指称的候选事件句。

[0081] 举例而言,以多媒体为新闻为例,对于给定的第二事件指称,可先从新闻等资讯自选文本进行分句,并从分句得到的句子中选取对该第二事件指称覆盖度最高的句子作为

该第二事件指称的候选事件句。

[0082] 步骤302,获取第二多媒体文本的标题信息。

[0083] 可选地,以多媒体为新闻为例,对于给定的事件指称,可获取该新闻等资讯的标题信息。

[0084] 步骤303,对第二事件指称、第二候选事件句和第二多媒体文本的标题信息分别进行实体识别。

[0085] 可选地,对于第二事件指称、第二候选事件句和第二多媒体文本的标题信息这三类文本,可同时利用通用的实体识别工具识别三类文本当中的时间、地点、人物、机构、触发词等重要的结构化信息,即可得到第二事件指称的实体识别结果,第二候选事件句的实体识别结果和标题信息的实体识别结果,其中,该实体识别结果可包括时间、地点、人物、机构、触发词等中至少一种重要的结构化信息。

[0086] 步骤304,根据实体识别结果,从第二候选事件句和第二多媒体文本的标题信息中选择其中一个作为第二事件指称的第二辅助文本。

[0087] 可选地,根据第二事件指称的实体识别结果、第二候选事件句的实体识别结果和标题信息的实体识别结果,比较标题信息与第二候选事件句中,相对于第二事件指称是否出现新增结构化信息,并从标题信息与第二候选事件句中选取包含最多新增结构化信息的文本作为第二事件指称的第二辅助文本。

[0088] 由此可见,本公开利用事件句召回和结构化信息识别技术,为事件指称确定具有辅助性的描述文本,大大降低了事件指称描述不全的现象。

[0089] 需要说明的是,本公开实施例可通过预先训练的事件指称匹配模型实现对给定的事件指称对进行匹配处理。可选地,图4为本公开实施例所提供的另一种事件指称匹配方法的流程图。需要说明的是,本公开实施例的事件指称匹配方法可应用于本公开实施例的事件指称匹配装置,该事件指称匹配装置可配置于电子设备上。

[0090] 如图4所示,本公开实施例的事件指称匹配方法可以包括:

[0091] 步骤401,获取待处理的第一事件指称和第二事件指称,并获取第一事件指称的第一辅助文本和第二事件指称的第二辅助文本。

[0092] 在本公开的实施例中,步骤401可以分别采用本公开的各实施例中的任一种方式实现,本公开实施例并不对此作出限定,也不再赘述。

[0093] 步骤402,获取预先经过训练的事件指称匹配模型;事件指称匹配模型包括结构化信息输入层。

[0094] 可选地,通过预先训练一个事件指称匹配模型,通过该事件指称匹配模型来实现对给定的事件指称对的预测,以判断该事件指称对是否指代相同事件。

[0095] 需要说明的是,本公开实施例中的事件指称匹配模型可包括结构化信息输入层、编码层和输出层。举例而言,图5为本公开实施例所提供的事件指称匹配模型的示例图。如图5所示,结构化信息输入层可包含文本编码、分句编码、触发词编码、实体编码这四种信息。其中,文本编码表征当前字的字义信息,分句编码表征当前字所属事件,触发词编码表征当前字是否为触发词的组成部分,实体编码表征当前字所属实体的实体类别。

[0096] 还需要说明的是,本公开实施例中的事件指称匹配模型可以是预先训练而得到的。在一种实现方式中,可通过以下方式预先训练该事件指称匹配模型:

[0097] 步骤1) 获取事件指称对样本和事件指称对样本对应的标签;

[0098] 其中, 该事件指称对样本包括第一事件指称样本和第二事件指称样本; 该标签用于指示第一事件指称样本与第二事件指称样本是否指代相同事件;

[0099] 步骤2) 获取第一事件指称样本的第一辅助文本样本和第二事件指称样本的第二辅助文本样本;

[0100] 其中, 第一辅助文本样本的获取方式与上述第一辅助文本的获取方式类似, 第二辅助文本样本的获取方式与上述第二辅助文本的获取方式类似, 在此不再赘述。

[0101] 步骤3) 构建事件指称匹配模型; 该事件指称匹配模型包括结构化信息输入层、编码层和输出层;

[0102] 步骤4) 利用第一事件指称样本、第一辅助文本样本、第二事件指称样本、第二辅助文本样本和标签, 对事件指称匹配模型进行训练。

[0103] 可选地, 将第一事件指称样本、第一辅助文本样本、第二事件指称样本和第二辅助文本样本进行拼接, 并将拼接后的文本和标签作为事件指称匹配模型的输入; 通过结构化信息输入层对输入文本分别进行文本编码、分句编码、触发词编码和实体编码, 并将文本编码结果和分句编码结果、触发词编码结果和实体编码结果进行融合处理, 获得融合编码特征向量; 通过编码层采用融合编码特征向量进行事件信息交互, 并对事件信息进行编码, 获得文本对编码特征向量; 通过输出层对文本对编码特征向量进行预测, 获取输出层输出的预测结果; 计算预测结果和标签的损失值, 并根据损失值训练事件指称匹配模型。

[0104] 由此可见, 本公开在建模过程中利用结构化信息增强模型对结构化新的感知, 有效提升了事件指称归一任务的判别效果。

[0105] 步骤403, 将第一事件指称、第一辅助文本、第二事件指称和第二辅助文本进行拼接, 将拼接后的文本作为事件指称匹配模型的输入。

[0106] 其中, 对于每一个事件, 将事件指称和对应的辅助文本作为一个事件的最终描述文本。

[0107] 步骤404, 通过结构化信息输入层对输入文本分别进行文本编码、分句编码、触发词编码和实体编码, 并将文本编码结果和分句编码结果、触发词编码结果和实体编码结果进行融合处理, 获得融合编码特征向量。

[0108] 可选地, 结构化信息输入层可包含文本编码、分句编码、触发词编码、实体编码这四种信息。其中, 文本编码表征当前字的字义信息, 分句编码表征当前字所属事件, 触发词编码表征当前字是否为触发词的组成部分, 实体编码表征当前字所属实体的实体类别。在本实施例中, 可利用结构化信息输入层对分别对每个事件的最终描述文本分别进行文本编码、分句编码、触发词编码和实体编码, 并将每个事件的文本编码结果和分句编码结果、触发词编码结果和实体编码结果进行融合处理, 获得每个事件的融合编码特征向量。

[0109] 步骤405, 将融合编码特征向量作为编码层的输入, 通过编码层采用融合编码特征向量进行事件信息交互, 并对事件信息进行编码, 获得文本对编码特征向量。

[0110] 作为一种示例, 该编码层可以利用但不限于Transformer等常用的基于深度学习的编码方式进行事件信息交互、对事件信息进行编码, 得到交互之后的文本对编码特征向量。

[0111] 步骤406, 将文本对编码特征向量作为输出层的输入, 通过输出层对文本对编码特

征向量进行预测。

[0112] 可选地,本公开实施例可以利用但不限于多层线性感知机 (MLP) 等人工神经网络进行输出层建模。通过输出层对该文本对编码特征向量进行预测。其中,该输出层的输出结果可以为预测两个事件指称是否为相同事件的概率得分。

[0113] 步骤407,获取输出层输出的预测结果,并根据预测结果确定第一事件指称和第二事件指称是否为相同事件。

[0114] 其中,可以理解,该测结果可以为预测两个事件指称是否为相同事件的概率得分。在本步骤中,可根据该预测结果确定两个事件是否匹配的输出结果,即确定第一事件指称和第二事件指称是否为相同事件,例如,该预测结果大于或等于一定阈值,确定第一事件指称和第二事件指称为相同事件;该预测结果小于该阈值,则确定第一事件指称和第二事件指称为不同事件。

[0115] 根据本公开实施例的事件指称匹配方法,利用事件句召回、结构化信息识别技术与一定的策略为事件指称确定具有辅助性的描述文本,大大降低了事件指称描述不全的现象,通过接过话信息检测策略完善事件描述。在建模过程中引入触发词、实体类型等结构化信息,增强模型对结构化新的感知,即使模型更好感知事件的结构化信息,可以有效提升事件指称归一任务的判别效果。

[0116] 为了实现上述实施例,本公开还提供了一种事件指称匹配模型的训练方法。

[0117] 图6为本公开实施例所提供的一种事件指称匹配模型的训练方法的流程图。需要说明的是,本公开实施例的事件指称匹配模型的训练方法可应用于本公开实施例的事件指称匹配模型的训练装置,该事件指称匹配模型的训练装置可配置于电子设备上。

[0118] 如图6所示,本公开实施例的事件指称匹配模型的训练方法可以包括:

[0119] 步骤601,获取事件指称对样本和事件指称对样本对应的标签。

[0120] 其中,在本公开实施例中,事件指称对样本包括第一事件指称样本和第二事件指称样本;标签用于指示第一事件指称样本与第二事件指称样本是否指代相同事件。

[0121] 步骤602,获取第一事件指称样本的第一辅助文本样本和第二事件指称样本的第二辅助文本样本。

[0122] 在本公开一些实施例中,可从第一事件指称样本所属的第一多媒体文本样本中选取第一候选事件句样本;获取第一多媒体文本样本的标题信息;对第一事件指称样本、第一候选事件句样本和第一多媒体文本样本的标题信息分别进行实体识别;根据实体识别结果,从第一候选事件句样本和第一多媒体文本样本的标题信息中选择其中一个作为第一事件指称样本的第一辅助文本样本。

[0123] 在本公开实施例中,可从第二事件指称样本所属的第二多媒体文本样本中选取第二候选事件句样本;获取第二多媒体文本样本的标题信息;对第二事件指称样本、第二候选事件句样本和第二多媒体文本样本的标题信息分别进行实体识别;根据实体识别结果,从第二候选事件句样本和第二多媒体文本样本的标题信息中选择其中一个作为第二事件指称样本的第二辅助文本样本。

[0124] 步骤603,构建事件指称匹配模型;事件指称匹配模型包括结构化信息输入层、编码层和输出层。

[0125] 举例而言,图5为本公开实施例所提供的事件指称匹配模型的示例图。如图5所示,

结构化信息输入层可包含文本编码、分句编码、触发词编码、实体编码这四种信息。其中,文本编码表征当前字的字义信息,分句编码表征当前字所属事件,触发词编码表征当前字是否为触发词的组成部分,实体编码表征当前字所属实体的实体类别。

[0126] 其中,在本公开实施例中,该编码层可以利用但不限于Transformer等常用的基于深度学习的编码方式进行事件信息交互、对事件信息进行编码,得到交互之后的文本对编码特征向量。

[0127] 可选地,本公开实施例可以利用但不限于多层线性感知机 (MLP) 等人工神经网络进行输出层建模。

[0128] 步骤604,利用第一事件指称样本、第一辅助文本样本、第二事件指称样本、第二辅助文本样本和标签,对事件指称匹配模型进行训练。

[0129] 在本公开实施例中,将第一事件指称样本、第一辅助文本样本、第二事件指称样本和第二辅助文本样本进行拼接,并将拼接后的文本和标签作为事件指称匹配模型的输入;通过结构化信息输入层对输入文本分别进行文本编码、分句编码、触发词编码和实体编码,并将文本编码结果和分句编码结果、触发词编码结果和实体编码结果进行融合处理,获得融合编码特征向量;通过编码层采用融合编码特征向量进行事件信息交互,并对事件信息进行编码,获得文本对编码特征向量;通过输出层对文本对编码特征向量进行预测,获取输出层输出的预测结果;计算预测结果和标签的损失值,并根据损失值训练事件指称匹配模型。

[0130] 根据本公开实施例的事件指称匹配模型的训练方法,对于每一个事件,将事件指称和其对应的辅助文本作为模型的输入文本,在建模过程中,将识别出的实体标签作为一种特征融入到事件指称匹配模型的建模过程中,增强模型对于结构化信息的感知。由此可见,在建模过程中引入触发词、实体类型等结构化信息,使模型更好感知事件的结构化信息,从而可以有效提升事件指称归一任务的判别效果。

[0131] 为了实现上述实施例,本公开还提出了一种事件指称匹配装置。

[0132] 图7为本公开实施例所提供的一种事件指称匹配装置的结构示意图。如图7所示,该事件指称匹配装置可以包括:获取模块701、第一生成模块702、第二生成模块703和确定模块704。

[0133] 具体地,获取模块701用于获取待处理的第一事件指称和第二事件指称,并获取第一事件指称的第一辅助文本和第二事件指称的第二辅助文本。

[0134] 在本公开一些实施例中,获取模块701从第一事件指称所属的第一多媒体文本中选取第一候选事件句;获取第一多媒体文本的标题信息;对第一事件指称、第一候选事件句和第一多媒体文本的标题信息分别进行实体识别;根据实体识别结果,从第一候选事件句和第一多媒体文本的标题信息中选择其中一个作为第一事件指称的第一辅助文本。

[0135] 在本公开实施例中,获取模块701从第二事件指称所属的第二多媒体文本中选取第二候选事件句;获取第二多媒体文本的标题信息;对第二事件指称、第二候选事件句和第二多媒体文本的标题信息分别进行实体识别;根据实体识别结果,从第二候选事件句和第二多媒体文本的标题信息中选择其中一个作为第二事件指称的第二辅助文本。

[0136] 第一生成模块702用于根据第一事件指称、第一辅助文本、第二事件指称和第二辅助文本,生成融合编码特征向量;融合编码特征向量为融合了第一事件指称、第一辅助文

本、第二事件指称和第二辅助文本的文本信息、结构化信息的编码特征向量。

[0137] 作为一种示例,第一生成模块702具体用于:获取预先经过训练的事件指称匹配模型;事件指称匹配模型包括结构化信息输入层;将第一事件指称、第一辅助文本、第二事件指称和第二辅助文本进行拼接,将拼接后的文本作为事件指称匹配模型的输入;通过结构化信息输入层对输入文本分别进行文本编码、分句编码、触发词编码和实体编码,并将文本编码结果和分句编码结果、触发词编码结果和实体编码结果进行融合处理,获得融合编码特征向量。

[0138] 第二生成模块703用于根据融合编码特征向量,生成文本对编码特征向量;文本对编码特征向量为由第一事件指称和第二事件指称组成的文本对的编码向量。在本公开一些实施例中,事件指称匹配模型还包括编码层;第二生成模块703将融合编码特征向量作为编码层的输入,通过编码层采用融合编码特征向量进行事件信息交互,并对事件信息进行编码,获得文本对编码特征向量。

[0139] 确定模块704用于根据文本对编码特征向量确定第一事件指称和第二事件指称是否为相同事件。

[0140] 在本公开实施例中,事件指称匹配模型还包括输出层。其中,确定模块704将文本对编码特征向量作为输出层的输入,通过输出层对文本对编码特征向量进行预测;获取输出层输出的预测结果,并根据预测结果确定第一事件指称和第二事件指称是否为相同事件。

[0141] 在本公开一些实施例中,如图8所示,该事件指称匹配装置还可包括:模型训练模块805。其中,模型训练模块805具体用于:获取事件指称对样本和事件指称对样本对应的标签;事件指称对样本包括第一事件指称样本和第二事件指称样本;标签用于指示第一事件指称样本与第二事件指称样本是否指代相同事件;获取第一事件指称样本的第一辅助文本样本和第二事件指称样本的第二辅助文本样本;构建事件指称匹配模型;事件指称匹配模型包括结构化信息输入层、编码层和输出层;利用第一事件指称样本、第一辅助文本样本、第二事件指称样本、第二辅助文本样本和标签,对事件指称匹配模型进行训练。

[0142] 其中,图8中801-804和图7中701-704具有相同功能和结构。

[0143] 关于上述实施例中的装置,其中各个模块执行操作的具体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。

[0144] 根据本公开实施例的事件指称匹配装置,利用事件句召回、结构化信息识别技术与一定的策略为事件指称确定具有辅助性的描述文本,大大降低了事件指称描述不全的现象,通过接过话信息检测策略完善事件描述。在建模过程中引入触发词、实体类型等结构化信息,增强模型对结构化新的感知,即使模型更好感知事件的结构化信息,可以有效提升事件指称归一任务的判别效果。

[0145] 为了实现上述实施例,本公开还提出了一种事件指称匹配模型的训练装置。

[0146] 图9为本公开实施例所提供的一种事件指称匹配模型的训练装置的结构示意图。如图9所示,该事件指称匹配模型的训练装置可以包括:第一获取模块901、第二获取模块902、构建模块903和训练模块904。

[0147] 具体地,第一获取模块901用于获取事件指称对样本和事件指称对样本对应的标签;事件指称对样本包括第一事件指称样本和第二事件指称样本;标签用于指示第一事件

指称样本与第二事件指称样本是否指代相同事件;

[0148] 第二获取模块902用于获取第一事件指称样本的第一辅助文本样本和第二事件指称样本的第二辅助文本样本;在本公开一些实施例中,第二获取模块902从第一事件指称样本所属的第一多媒体文本样本中选取第一候选事件句样本;获取第一多媒体文本样本的标题信息;对第一事件指称样本、第一候选事件句样本和第一多媒体文本样本的标题信息分别进行实体识别;根据实体识别结果,从第一候选事件句样本和第一多媒体文本样本的标题信息中选择其中一个作为第一事件指称样本的第一辅助文本样本。

[0149] 构建模块903用于构建事件指称匹配模型;事件指称匹配模型包括结构化信息输入层、编码层和输出层;

[0150] 训练模块904用于利用第一事件指称样本、第一辅助文本样本、第二事件指称样本、第二辅助文本样本和标签,对事件指称匹配模型进行训练。

[0151] 可选地,在本公开实施例中,训练模块904具体用于:将第一事件指称样本、第一辅助文本样本、第二事件指称样本和第二辅助文本样本进行拼接,并将拼接后的文本和标签作为事件指称匹配模型的输入;通过结构化信息输入层对输入文本分别进行文本编码、分句编码、触发词编码和实体编码,并将文本编码结果和分句编码结果、触发词编码结果和实体编码结果进行融合处理,获得融合编码特征向量;通过编码层采用融合编码特征向量进行事件信息交互,并对事件信息进行编码,获得文本对编码特征向量;通过输出层对文本对编码特征向量进行预测,获取输出层输出的预测结果;计算预测结果和标签的损失值,并根据损失值训练事件指称匹配模型。

[0152] 关于上述实施例中的装置,其中各个模块执行操作的具体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。

[0153] 根据本公开实施例的事件指称匹配模型的训练装置,对于每一个事件,将事件指称和其对应的辅助文本作为模型的输入文本,在建模过程中,将识别出的实体标签作为一种特征融入到事件指称匹配模型的建模过程中,增强模型对于结构化信息的感知。由此可见,在建模过程中引入触发词、实体类型等结构化信息,使模型更好感知事件的结构化信息,从而可以有效提升事件指称归一任务的判别效果。

[0154] 根据本公开的实施例,本公开还提供了一种电子设备和一种可读存储介质。

[0155] 如图10所示,是本公开实施例所提供的一种电子设备的框图。电子设备旨在表示各种形式的数字计算机,诸如,膝上型计算机、台式计算机、工作台、个人数字助理、服务器、刀片式服务器、大型计算机、和其它适合的计算机。电子设备还可以表示各种形式的移动装置,诸如,个人数字处理、蜂窝电话、智能电话、可穿戴设备和其它类似的计算装置。本文所示的部件、它们的连接和关系、以及它们的功能仅作为示例,并且不意在限制本文中描述的和/或者要求的本公开的实现。

[0156] 如图10所示,该电子设备包括:一个或多个处理器1001、存储器1002,以及用于连接各部件的接口,包括高速接口和低速接口。各个部件利用不同的总线互相连接,并且可以被安装在公共主板上或者根据需要以其它方式安装。处理器可以对在电子设备内执行的指令进行处理,包括存储在存储器中或者存储器上以在外部输入/输出装置(诸如,耦合至接口的显示设备)上显示GUI的图形信息的指令。在其它实施方式中,若需要,可以将多个处理器和/或多条总线与多个存储器和多个存储器一起使用。同样,可以连接多个电子设备,各

个设备提供部分必要的操作(例如,作为服务器阵列、一组刀片式服务器、或者多处理器系统)。图10中以一个处理器1001为例。

[0157] 存储器1002即为本公开所提供的非瞬时计算机可读存储介质。其中,所述存储器存储有可由至少一个处理器执行的指令,以使所述至少一个处理器执行本公开所提供的事件指称匹配方法和/或事件指称匹配模型的训练方法。本公开的非瞬时计算机可读存储介质存储计算机指令,该计算机指令用于使计算机执行本公开所提供的事件指称匹配方法和/或事件指称匹配模型的训练方法。

[0158] 存储器1002作为一种非瞬时计算机可读存储介质,可用于存储非瞬时软件程序、非瞬时计算机可执行程序以及模块,如本公开实施例中的事件指称匹配方法和/或事件指称匹配模型的训练方法对应的程序指令/模块。处理器1001通过运行存储在存储器1002中的非瞬时软件程序、指令以及模块,从而执行服务器的各种功能应用以及数据处理,即实现上述方法实施例中的事件指称匹配方法和/或事件指称匹配模型的训练方法。

[0159] 存储器1002可以包括存储程序区和存储数据区,其中,存储程序区可存储操作系统、至少一个功能所需要的应用程序;存储数据区可存储根据事件指称匹配的电子设备的使用所创建的数据等。此外,存储器1002可以包括高速随机存取存储器,还可以包括非瞬时存储器,例如至少一个磁盘存储器件、闪存器件、或其他非瞬时固态存储器件。在一些实施例中,存储器1002可选包括相对于处理器1001远程设置的存储器,这些远程存储器可以通过网络连接至事件指称匹配的电子设备。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

[0160] 本公开实施例的电子设备还可以包括:输入装置1003和输出装置1004。处理器1001、存储器1002、输入装置1003和输出装置1004可以通过总线或者其他方式连接,图10中以通过总线连接为例。

[0161] 输入装置1003可接收输入的数字或字符信息,以及产生与事件指称匹配的电子设备的用户设置以及功能控制有关的键信号输入,例如触摸屏、小键盘、鼠标、轨迹板、触摸板、指示杆、一个或者多个鼠标按钮、轨迹球、操纵杆等输入装置。输出装置1004可以包括显示设备、辅助照明装置(例如,LED)和触觉反馈装置(例如,振动电机)等。该显示设备可以包括但不限于,液晶显示器(LCD)、发光二极管(LED)显示器和等离子体显示器。在一些实施方式中,显示设备可以是触摸屏。

[0162] 此处描述的系统和技术各种实施方式可以在数字电子电路系统、集成电路系统、专用ASIC(专用集成电路)、计算机硬件、固件、软件、和/或它们的组合中实现。这些各种实施方式可以包括:实施在一个或者多个计算机程序中,该一个或者多个计算机程序可在包括至少一个可编程处理器的可编程系统上执行和/或解释,该可编程处理器可以是专用或者通用可编程处理器,可以从存储系统、至少一个输入装置、和至少一个输出装置接收数据和指令,并且将数据和指令传输至该存储系统、该至少一个输入装置、和该至少一个输出装置。

[0163] 这些计算程序(也称作程序、软件、软件应用、或者代码)包括可编程处理器的机器指令,并且可以利用高级过程和/或面向对象的编程语言、和/或汇编/机器语言来实施这些计算程序。如本文使用的,术语“机器可读介质”和“计算机可读介质”指的是用于将机器指令和/或数据提供给可编程处理器的任何计算机程序产品、设备、和/或装置(例如,磁盘、光

盘、存储器、可编程逻辑装置(PLD)),包括,接收作为机器可读信号的机器指令的机器可读介质。术语“机器可读信号”指的是用于将机器指令和/或数据提供给可编程处理器的任何信号。

[0164] 为了提供与用户的交互,可以在计算机上实施此处描述的系统和技术,该计算机具有:用于向用户显示信息的显示装置(例如,CRT(阴极射线管)或者LCD(液晶显示器)监视器);以及键盘和指向装置(例如,鼠标或者轨迹球),用户可以通过该键盘和该指向装置来将输入提供给计算机。其它种类的装置还可以用于提供与用户的交互;例如,提供给用户的反馈可以是任何形式的传感反馈(例如,视觉反馈、听觉反馈、或者触觉反馈);并且可以用任何形式(包括声输入、语音输入或者、触觉输入)来接收来自用户的输入。

[0165] 可以将此处描述的系统和技术实施在包括后台部件的计算系统(例如,作为数据服务器)、或者包括中间件部件的计算系统(例如,应用服务器)、或者包括前端部件的计算系统(例如,具有图形用户界面或者网络浏览器的用户计算机,用户可以通过该图形用户界面或者该网络浏览器来与此处描述的系统和技术实施方式交互)、或者包括这种后台部件、中间件部件、或者前端部件的任何组合的计算系统中。可以通过任何形式或者介质的数字数据通信(例如,通信网络)来将系统的部件相互连接。通信网络的示例包括:局域网(LAN)、广域网(WAN)、互联网和区块链网络。

[0166] 计算机系统可以包括客户端和服务端。客户端和服务端一般远离彼此并且通常通过通信网络进行交互。通过在相应的计算机上运行并且彼此具有客户端-服务器关系的计算机程序来产生客户端和服务端的关系。服务器可以是云服务器,又称为云计算服务器或云主机,是云计算服务体系中的一项主机产品,以解决了传统物理主机与VPS服务(“Virtual Private Server”,或简称“VPS”)中,存在的管理难度大,业务扩展性弱的缺陷。服务器也可以为分布式系统的服务器,或者是结合了区块链的服务器。

[0167] 应该理解,可以使用上面所示的各种形式的流程,重新排序、增加或删除步骤。例如,本发申请中记载的各步骤可以并行地执行也可以顺序地执行也可以不同的次序执行,只要能够实现本公开公开的技术方案所期望的结果,本文在此不进行限制。

[0168] 上述具体实施方式,并不构成对本公开保护范围的限制。本领域技术人员应该明白的是,根据设计要求和因素,可以进行各种修改、组合、子组合和替代。任何在本公开的精神和原则之内所作的修改、等同替换和改进等,均应包含在本公开保护范围之内。

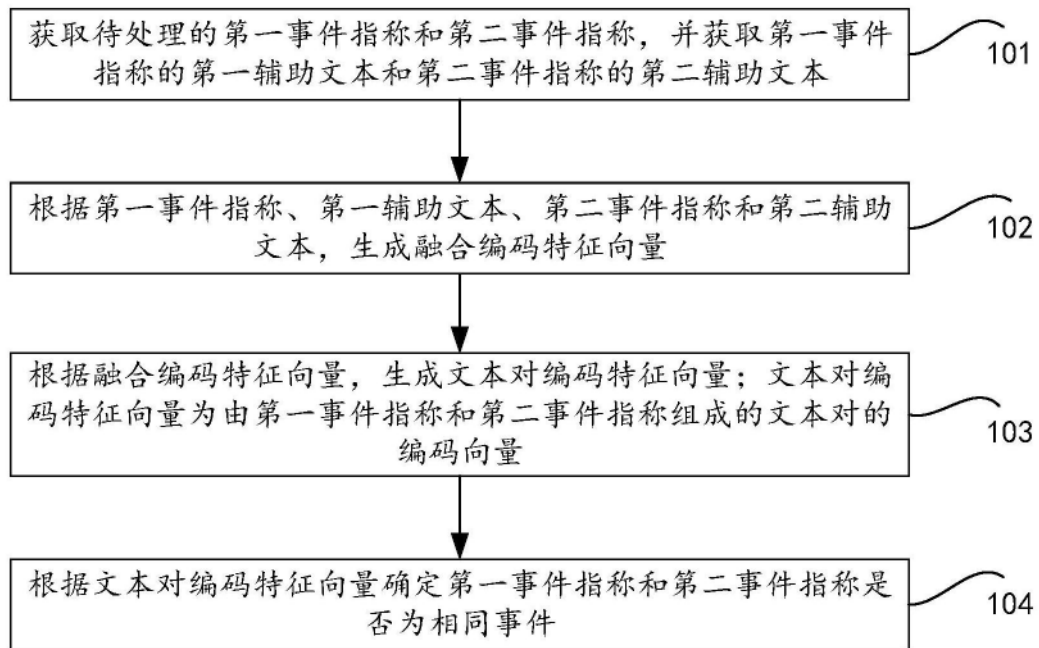


图1

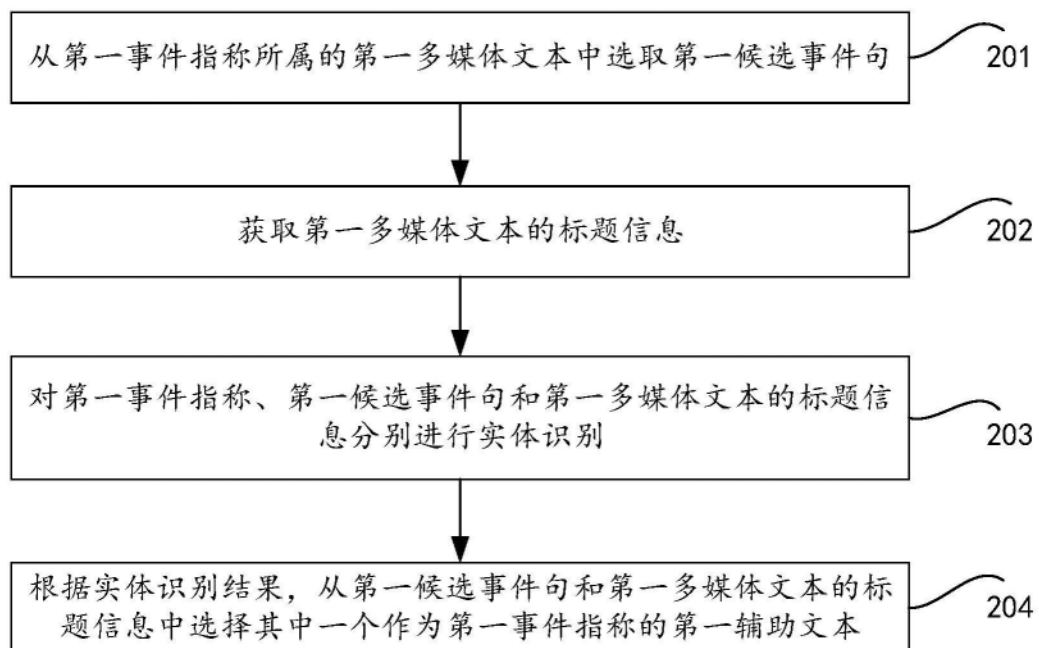


图2

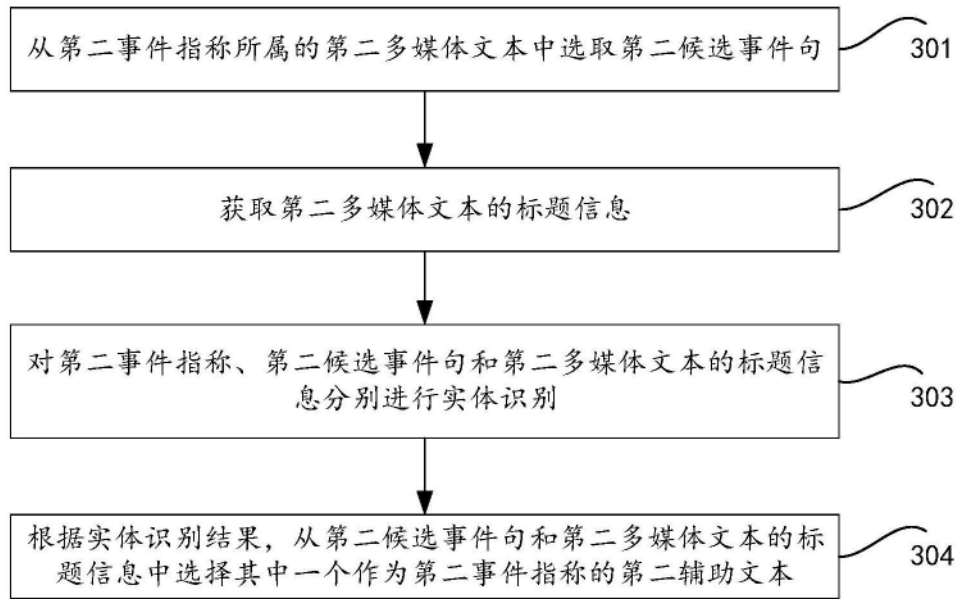


图3

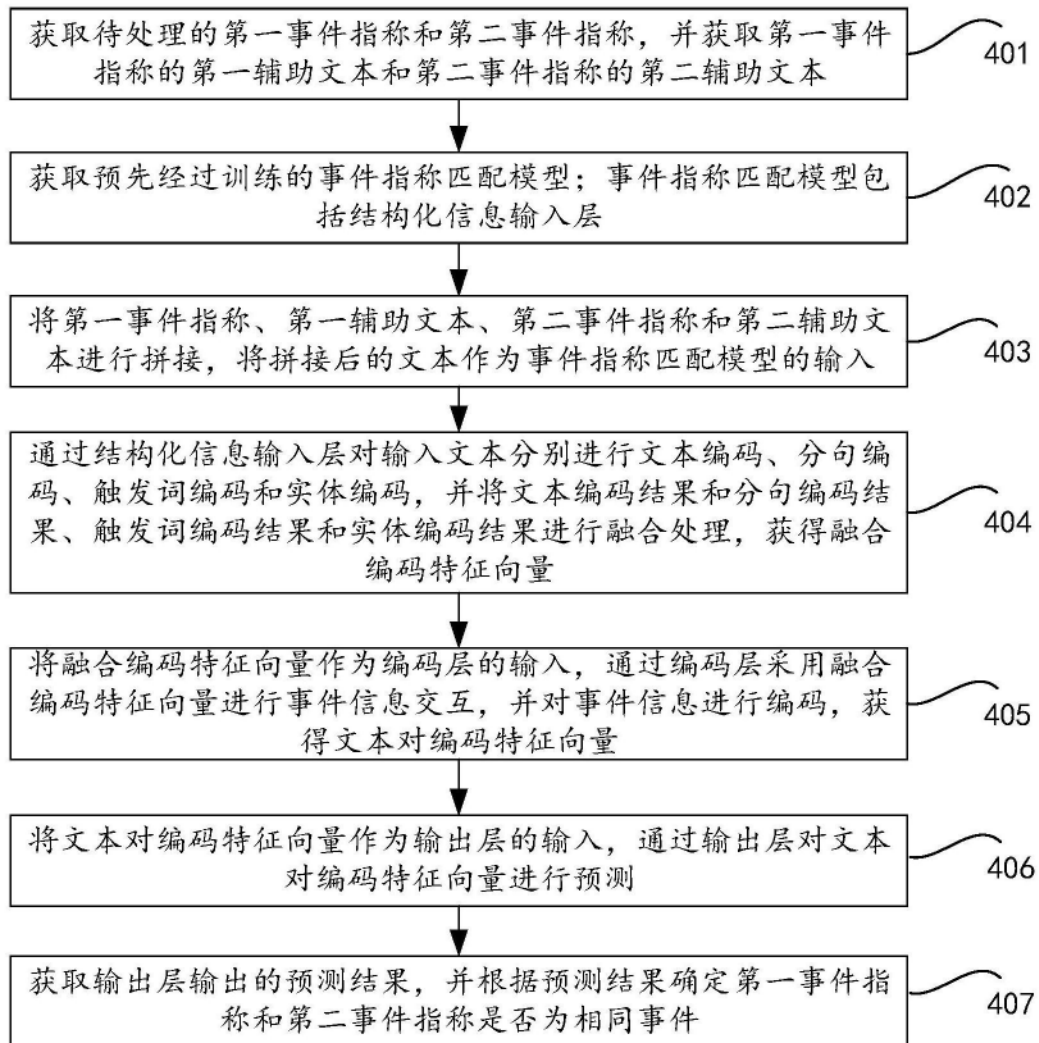


图4

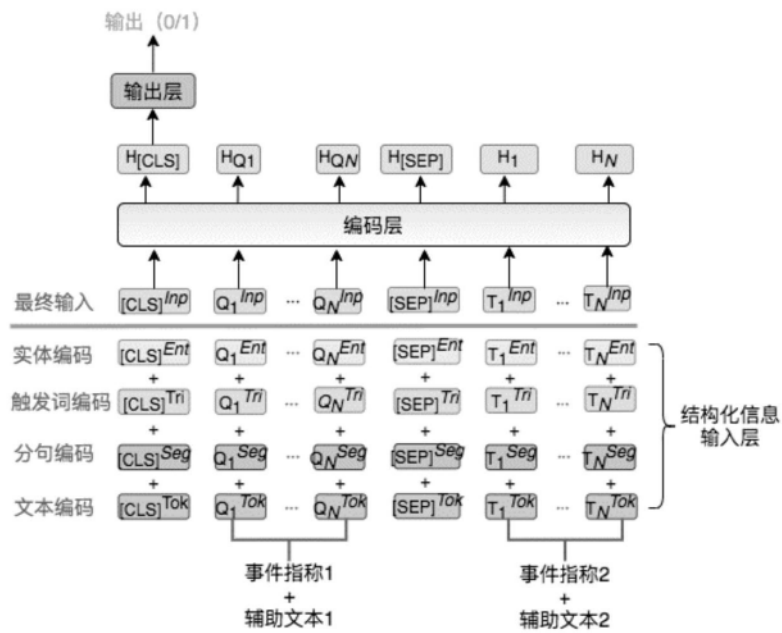


图5

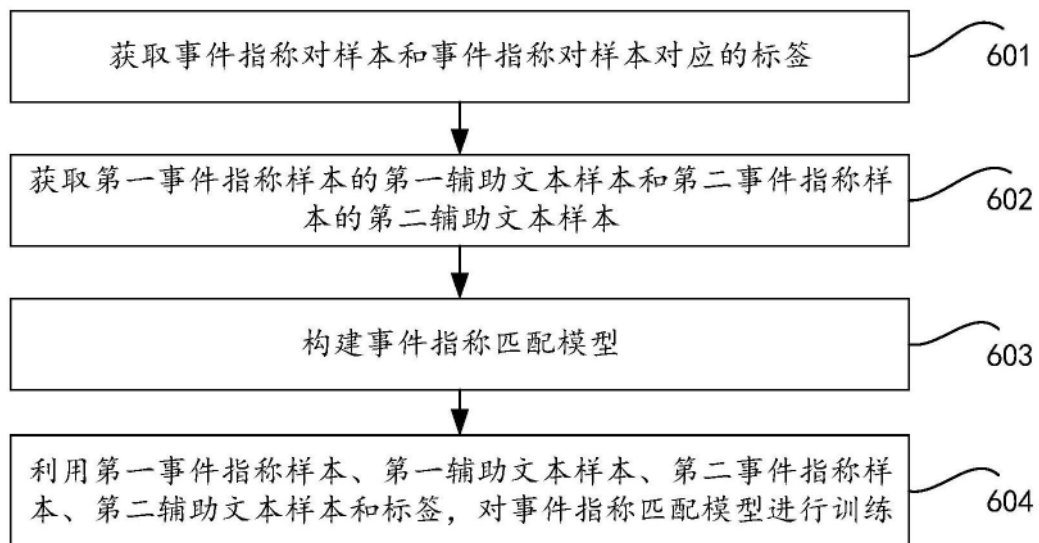


图6

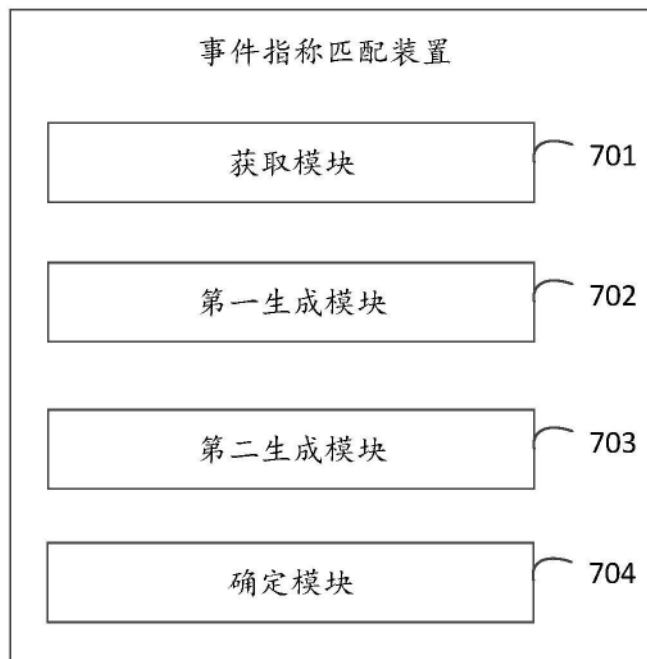


图7

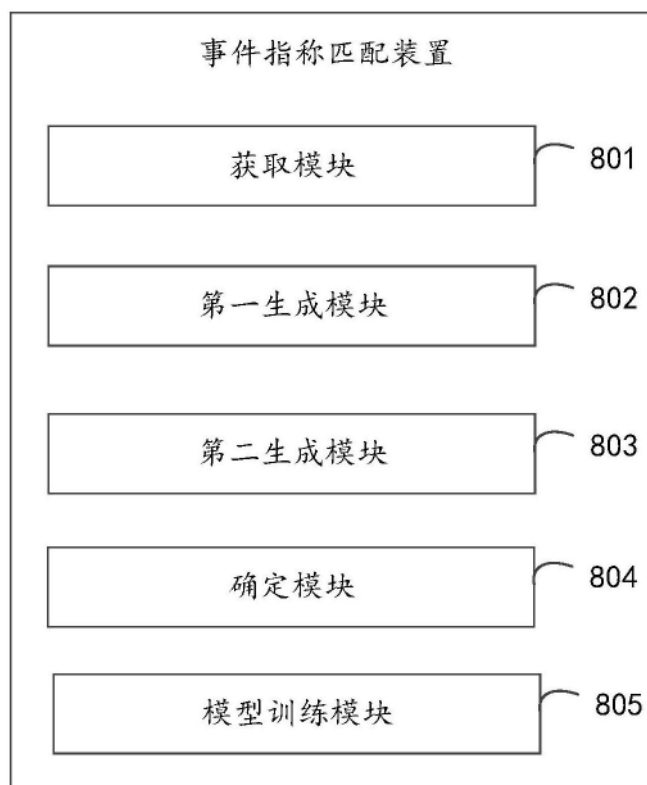


图8

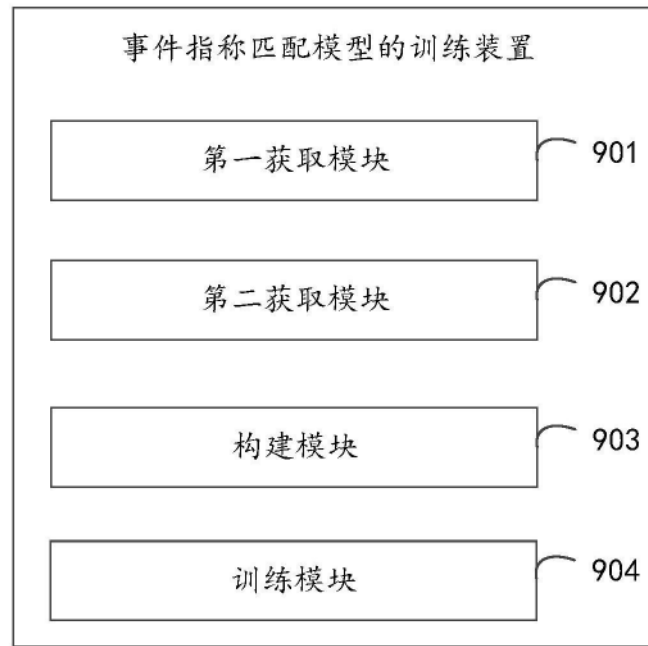


图9

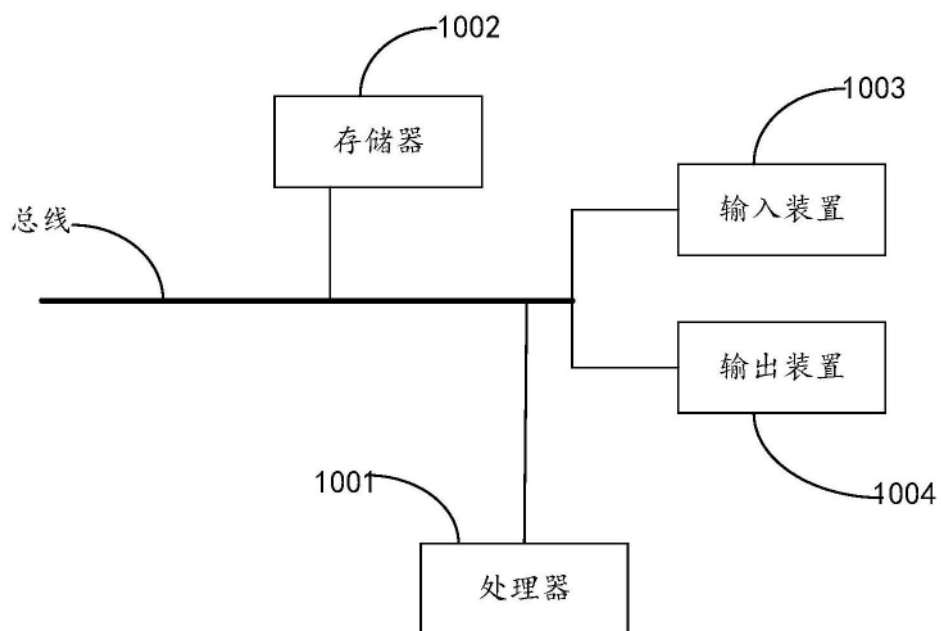


图10