(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 111291564 B
(45)授权公告日 2023.10.31

(21)申请号 202010139871.7
(22)申请日 2020.03.03
(65)同一申请的已公布的文献号
申请公布号 CN 111291564 A
(43)申请公布日 2020.06.16
(73)专利权人 腾讯科技(深圳)有限公司
地址 518057 广东省深圳市南山区高新区科技中一路腾讯大厦35层
(72)发明人 缪畅宇
(74)专利代理机构 深圳翼盛智成知识产权事务所(普通合伙) 44300
专利代理师 蔡艾莹
(51)Int.Cl.
G06F 40/289(2020.01)
G06F 18/22(2023.01)
G06F 18/241(2023.01)
G06N 3/08(2023.01)

(56)对比文件
CN 110598213 A,2019.12.20
CN 110717039 A,2020.01.21
CN 110309308 A,2019.10.08
CN 109992773 A,2019.07.09
WO 2019210820 A1,2019.11.07

审查员 李萌

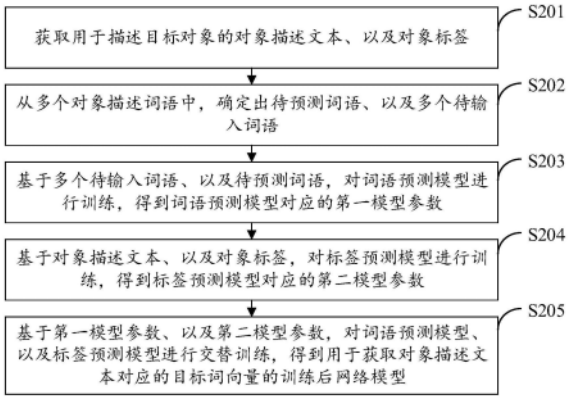
权利要求书3页 说明书18页 附图6页

(54)发明名称

一种用于词向量获取的模型训练方法、装置和存储介质

(57)摘要

本申请实施例公开了一种用于词向量获取的模型训练方法、装置和存储介质，所述方法包括：获取用于描述目标对象的对象描述文本、以及对象标签，从多个对象描述词语中，确定出待预测词语、以及多个待输入词语，基于多个待输入词语、以及待预测词语，对词语预测模型进行训练，得到词语预测模型对应的第一模型参数，基于对象描述文本、以及对象标签，对标签预测模型进行训练，得到标签预测模型对应的第二模型参数，基于第一模型参数、以及第二模型参数，对词语预测模型、以及标签预测模型进行交替训练，得到训练后网络模型。通过该方案得到的训练后网络模型，可以获取到同时包括对象描述文本中信息、以及对象标签中信息的目标词向量。



CN 111291564 B

1.一种用于词向量获取的模型训练方法,其特征在于,包括:

获取用于描述目标对象的对象描述文本、以及对象标签,其中,所述对象描述文本中包括多个对象描述词语;

从所述多个对象描述词语中,确定出待预测词语、以及多个待输入词语;

基于所述多个待输入词语、以及所述待预测词语,对词语预测模型进行训练,得到所述词语预测模型对应的第一模型参数,包括:

将所述多个待输入词语输入至词语预测模型中,并基于所述词语预测模型预测所述多个待输入词语对应的预测后词语;

基于所述预测后词语、以及所述待预测词语,采用第一损失函数确定所述词语预测模型对应的第一模型参数;

基于所述对象描述文本、以及所述对象标签,对标签预测模型进行训练,得到所述标签预测模型对应的第二模型参数,包括:

将所述多个对象描述词语输入至标签预测模型中,并基于所述标签预测模型预测所述多个对象描述词语对应的预测后标签;

基于所述预测后标签、以及所述对象标签,采用第二损失函数确定所述标签预测模型对应的第二模型参数;

基于所述第一模型参数、以及所述第二模型参数,对所述词语预测模型、以及所述标签预测模型进行交替训练,得到用于获取所述对象描述文本对应的目标词向量的训练后网络模型。

2.根据权利要求1所述的用于词向量获取的模型训练方法,其特征在于,所述词语预测模型中包括全连接层、以及第一分类层;

将所述多个待输入词语输入至词语预测模型中,并基于所述词语预测模型预测所述多个待输入词语对应的预测后词语,包括:

将所述多个待输入词语输入至词语预测模型中,并将所述多个待输入词语转换为多个第一待输入词向量;

基于所述全连接层,将所述多个第一待输入词向量映射为第一待分类向量;

基于所述第一分类层、以及所述第一待分类向量,预测所述多个待输入词语对应的预测后词语。

3.根据权利要求1所述的用于词向量获取的模型训练方法,其特征在于,所述标签预测模型中包括全连接层、以及第二分类层;

将所述多个对象描述词语输入至标签预测模型中,并基于所述标签预测模型预测所述多个对象描述词语对应的预测后标签,包括:

将所述多个对象描述词语输入至标签预测模型中,并将所述多个对象描述词语转换为多个第二待输入词向量;

基于所述全连接层,将所述多个第二待输入词向量映射为第二待分类向量;

基于所述第二分类层、以及所述第二待分类向量,预测所述多个对象描述词语对应的预测后标签。

4.根据权利要求1所述的用于词向量获取的模型训练方法,其特征在于,基于所述第一模型参数、以及所述第二模型参数,对所述词语预测模型、以及所述标签预测模型进行交替

训练，得到用于获取所述对象描述文本对应的目标词向量的训练后网络模型，包括：

基于所述多个待输入词语、以及所述待预测词语，对词语预测模型进行训练，得到所述词语预测模型对应的第三模型参数；

基于所述对象描述文本、以及所述对象标签，对标签预测模型进行训练，得到所述标签预测模型对应的第四模型参数；

基于所述第一模型参数、所述第二模型参数、所述第三模型参数、以及所述第四模型参数，对所述词语预测模型、以及所述标签预测模型进行交替训练，得到用于获取所述对象描述文本对应的目标词向量的训练后网络模型。

5.根据权利要求1所述的用于词向量获取的模型训练方法，其特征在于，所述方法还包括：

基于所述训练后网络模型，将所述对象描述文本表示为目标词向量；

基于所述训练后网络模型，将所述对象标签表示为目标标签词向量；

基于所述训练后网络模型的全连接层，对所述目标词向量、以及所述目标标签词向量进行融合，得到所述对象描述文本对应的目标文本表示。

6.根据权利要求1所述的用于词向量获取的模型训练方法，其特征在于，从所述多个对象描述词语中，确定出待预测词语、以及多个待输入词语，包括：

从所述多个对象描述词语中，确定待预测词语；

从所述多个对象描述词语中删除所述待预测词语，得到多个待输入词语。

7.一种用于词向量获取的模型训练装置，其特征在于，包括：

获取模块，用于获取用于描述目标对象的对象描述文本、以及对象标签，其中，所述对象描述文本中包括多个对象描述词语；

词语确定模块，用于从所述多个对象描述词语中，确定出待预测词语、以及多个待输入词语；

第一训练模块，用于基于所述多个待输入词语、以及所述待预测词语，对词语预测模型进行训练，得到所述词语预测模型对应的第一模型参数，包括：

第一预测子模块，用于将所述多个待输入词语输入至词语预测模型中，并基于所述词语预测模型预测所述多个待输入词语对应的预测后词语；

第一确定子模块，用于基于所述预测后词语、以及所述待预测词语，采用第一损失函数确定所述词语预测模型对应的第一模型参数；

第二训练模块，用于基于所述对象描述文本、以及所述对象标签，对标签预测模型进行训练，得到所述标签预测模型对应的第二模型参数，包括：

第二预测子模块，用于将所述多个对象描述词语输入至标签预测模型中，并基于所述标签预测模型预测所述多个对象描述词语对应的预测后标签；

第二确定子模块，用于基于所述预测后标签、以及所述对象标签，采用第二损失函数确定所述标签预测模型对应的第二模型参数；

第三训练模块，用于基于所述第一模型参数、以及所述第二模型参数，对所述词语预测模型、以及所述标签预测模型进行交替训练，得到用于获取所述对象描述文本对应的目标词向量的训练后网络模型。

8.一种计算机存储介质，其上存储有计算机程序，其特征在于，当所述计算机程序在计

算机上运行时，使得所述计算机执行如权利要求1-6任一项所述的用于词向量获取的模型训练方法。

9.一种电子设备，包括存储器，处理器及存储在存储器上并可在处理器上运行的计算机程序，其中，所述处理器执行所述程序时实现如权利要求1至6任一项所述方法的步骤。

一种用于词向量获取的模型训练方法、装置和存储介质

技术领域

[0001] 本申请涉及计算机技术领域，具体涉及一种用于词向量获取的模型训练方法、装置和存储介质。

背景技术

[0002] 在互联网内容领域，短视频标题、新闻标题、商品名称等都可以称为短文本，可以应用短文本对短视频、新闻、文章、商品等对象进行描述，如某个标题为“最畅销的口红，某明星带货!”的短视频，该标题就是该短视频所对应的短文本，它简要描述了该短视频的内容。同时，还可以利用标签对短视频、新闻、文章、商品等对象进行描述，该标签一般是人工运营或者算法标注的结果，表明该对象所属的类别。如，某个标题为“最畅销的口红，某明星带货!”的短视频，还可以对应着“美妆”的标签。

[0003] 在实际应用中，通常需要通过向量的形式对短文本进行表示，由于该向量中包含与短文本相关的信息，因此，该向量可以用于寻找相似内容，或者匹配目标用户等等。

发明内容

[0004] 本申请实施例提供一种用于词向量获取的模型训练方法、装置和存储介质，通过该方案得到的训练后网络模型，可以获取到同时包括对象描述文本中信息、以及对象标签中信息的目标词向量。

[0005] 本申请实施例提供一种用于词向量获取的模型训练方法，包括：

[0006] 获取用于描述目标对象的对象描述文本、以及对象标签，其中，所述对象描述文本中包括多个对象描述词语；

[0007] 从所述多个对象描述词语中，确定出待预测词语、以及多个待输入词语；

[0008] 基于所述多个待输入词语、以及所述待预测词语，对词语预测模型进行训练，得到所述词语预测模型对应的第一模型参数；

[0009] 基于所述对象描述文本、以及所述对象标签，对标签预测模型进行训练，得到所述标签预测模型对应的第二模型参数；

[0010] 基于所述第一模型参数、以及所述第二模型参数，对所述词语预测模型、以及所述标签预测模型进行交替训练，得到用于获取所述对象描述文本对应的目标词向量的训练后网络模型。

[0011] 相应的，本申请实施例还提供一种用于词向量获取的模型训练装置，包括：

[0012] 获取模块，用于获取用于描述目标对象的对象描述文本、以及对象标签，其中，所述对象描述文本中包括多个对象描述词语；

[0013] 词语确定模块，用于从所述多个对象描述词语中，确定出待预测词语、以及多个待输入词语；

[0014] 第一训练模块，用于基于所述多个待输入词语、以及所述待预测词语，对词语预测模型进行训练，得到所述词语预测模型对应的第一模型参数；

[0015] 第二训练模块，用于基于所述对象描述文本、以及所述对象标签，对标签预测模型进行训练，得到所述标签预测模型对应的第二模型参数；

[0016] 第三训练模块，用于基于所述第一模型参数、以及所述第二模型参数，对所述词语预测模型、以及所述标签预测模型进行交替训练，得到用于获取所述对象描述文本对应的目标词向量的训练后网络模型。

[0017] 可选的，在一些实施例中，所述第一训练模块可以包括第一预测子模块和第一确定子模块，如下：

[0018] 第一预测子模块，用于将所述多个待输入词语输入至词语预测模型中，并基于所述词语预测模型预测所述多个待输入词语对应的预测后词语；

[0019] 第一确定子模块，用于基于所述预测后词语、以及所述待预测词语，采用第一损失函数确定所述词语预测模型对应的第一模型参数。

[0020] 则此时，所述第一预测子模块，具体可以用于将所述多个待输入词语输入至词语预测模型中，并将所述多个待输入词语转换为多个第一待输入词向量，基于所述全连接层，将所述多个第一待输入词向量映射为第一待分类向量，基于所述第一分类层、以及所述第一待分类向量，预测所述多个待输入词语对应的预测后词语。

[0021] 可选的，在一些实施例中，所述第二训练模块可以包括第二预测子模块和第二确定子模块，如下：

[0022] 第二预测子模块，用于将所述多个对象描述词语输入至标签预测模型中，并基于所述标签预测模型预测所述多个对象描述词语对应的预测后标签；

[0023] 第二确定子模块，用于基于所述预测后标签、以及所述对象标签，采用第二损失函数确定所述标签预测模型对应的第二模型参数。

[0024] 则此时，所述第二预测子模块，具体可以用于将所述多个对象描述词语输入至标签预测模型中，并将所述多个对象描述词语转换为多个第二待输入词向量，基于所述全连接层，将所述多个第二待输入词向量映射为第二待分类向量，基于所述第二分类层、以及所述第二待分类向量，预测所述多个对象描述词语对应的预测后标签。

[0025] 则此时，所述第三训练模块，具体可以用于基于所述多个待输入词语、以及所述待预测词语，对词语预测模型进行训练，得到所述词语预测模型对应的第三模型参数，基于所述对象描述文本、以及所述对象标签，对标签预测模型进行训练，得到所述标签预测模型对应的第四模型参数，基于所述第一模型参数、所述第二模型参数、所述第三模型参数、以及所述第四模型参数，对所述词语预测模型、以及所述标签预测模型进行交替训练，得到用于获取所述对象描述文本对应的目标词向量的训练后网络模型。

[0026] 可选的，在一些实施例中，所述用于词向量获取的模型训练装置还可以包括第一表示模块、第二表示模块和融合模块，如下：

[0027] 第一表示模块，用于基于所述训练后网络模型，将所述对象描述文本表示为目标词向量；

[0028] 第二表示模块，用于基于所述训练后网络模型，将所述对象标签表示为目标标签词向量；

[0029] 融合模块，用于基于所述训练后网络模型的全连接层，对所述目标词向量、以及所述目标标签词向量进行融合，得到所述对象描述文本对应的目标文本表示。

[0030] 则此时，所述词语确定模块，具体可以用于从所述多个对象描述词语中，确定待预测词语，从所述多个对象描述词语中删除所述待预测词语，得到多个待输入词语。

[0031] 此外，本申请实施例还提供一种计算机存储介质，所述计算机存储介质存储有多条指令，所述指令适于处理器进行加载，以执行本申请实施例提供的任一种用于词向量获取的模型训练方法中的步骤。

[0032] 此外，本申请实施例还提供一种电子设备，包括存储器，处理器及存储在存储器上并可在处理器上运行的计算机程序，所述处理器执行所述程序时实现如本申请实施例提供的任一种用于词向量获取的模型训练方法中的步骤。

[0033] 本申请实施例可以获取用于描述目标对象的对象描述文本、以及对象标签，其中，对象描述文本中包括多个对象描述词语，从多个对象描述词语中，确定出待预测词语、以及多个待输入词语，基于多个待输入词语、以及待预测词语，对词语预测模型进行训练，得到词语预测模型对应的第一模型参数，基于对象描述文本、以及对象标签，对标签预测模型进行训练，得到标签预测模型对应的第二模型参数，基于第一模型参数、以及第二模型参数，对词语预测模型、以及标签预测模型进行交替训练，得到用于获取对象描述文本对应的目标词向量的训练后网络模型。通过该方案得到的训练后网络模型，可以获取到同时包括对象描述文本中信息、以及对象标签中信息的目标词向量。

附图说明

[0034] 为了更清楚地说明本申请实施例中的技术方案，下面将对实施例描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实施例，对于本领域技术人员来讲，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。

[0035] 图1是本申请实施例提供的用于词向量获取的模型训练系统的场景示意图；

[0036] 图2是本申请实施例提供的用于词向量获取的模型训练方法的第一流程图；

[0037] 图3是本申请实施例提供的用于词向量获取的模型训练方法的第二流程图；

[0038] 图4是本申请实施例提供的词向量生成模型的结构示意图；

[0039] 图5是本申请实施例提供的目标文本表示获取的示意图；

[0040] 图6是本申请实施例提供的用于词向量获取的模型训练装置的结构示意图；

[0041] 图7是本申请实施例提供的电子设备的结构示意图。

具体实施方式

[0042] 下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

[0043] 本申请实施例提供一种用于词向量获取的模型训练方法、装置、存储介质和电子设备。具体地，本申请实施例的用于词向量获取的模型训练方法可以由电子设备执行，其中，该电子设备可以为终端或者服务器等设备，该终端可以为手机、平板电脑、笔记本电脑、智能电视、穿戴式智能设备、个人计算机(PC，Personal Computer)等设备。其中，终端可以

包括客户端，该客户端可以是视频客户端或浏览器客户端等，服务器可以是单台服务器，也可以是由多个服务器组成的服务器集群。

[0044] 例如，参见图1，以该用于词向量获取的模型训练方法由电子设备执行为例，该电子设备可以获取用于描述目标对象的对象描述文本、以及对象标签，其中，对象描述文本中包括多个对象描述词语，从多个对象描述词语中，确定出待预测词语、以及多个待输入词语，基于多个待输入词语、以及待预测词语，对词语预测模型进行训练，得到词语预测模型对应的第一模型参数，基于对象描述文本、以及对象标签，对标签预测模型进行训练，得到标签预测模型对应的第二模型参数，基于第一模型参数、以及第二模型参数，对词语预测模型、以及标签预测模型进行交替训练，得到用于获取对象描述文本对应的目标词向量的训练后网络模型。

[0045] 以下分别进行详细说明。需说明的是，以下实施例的描述顺序不作为对实施例优选顺序的限定。

[0046] 本申请实施例提供了一种用于词向量获取的模型训练方法，该方法可以由终端或服务器执行，也可以由终端和服务器共同执行；本申请实施例以用于词向量获取的模型训练方法由服务器执行为例来进行说明，如图2所示，该用于词向量获取的模型训练方法的具体流程可以如下：

[0047] 201、获取用于描述目标对象的对象描述文本、以及对象标签。

[0048] 其中，对象描述文本可以是能够对目标对象进行描述的文本，视频、图像、新闻、文章、商品等都可以称为目标对象，那么能够对这些目标对象进行描述的视频标题、新闻标题、商品名称等都可以作为对象描述文本。比如，针对某个短视频，它的视频标题为“最畅销的口红，某明星带货!”，则该视频标题“最畅销的口红，某明星带货!”可以作为对象描述文本，该对象描述文本简要地描述了该短视频的内容。

[0049] 其中，一个对象描述文本中可以包括多个对象描述词语，而对象描述词语可以为从对象描述文本中划分出的独立单元，如对象描述文本中的单字、词语、词组等，都可以称为对象描述词语。比如，对象描述文本“xxx推荐的口红”，可以对应四个对象描述词语，分别为“xxx”、“推荐”、“的”、以及“口红”。其中，可以根据实际情况的需求，选取合适的文本分词方法，将对象描述文本划分为多个对象描述词语。

[0050] 其中，对象标签可以是能够表明目标对象所属类别的标签信息。该对象标签一般是人工运营或者算法标注的结果，并且也可以利用对象标签对短视频、新闻、文章、商品等目标对象进行描述。比如，针对某个短视频，它的视频标题为“最畅销的口红，某明星带货!”，并且该短视频可以对应着“美妆”、“明星”的对象标签。

[0051] 其中，针对同一个目标对象，既可以应用对象描述文本对其进行描述，还可以应用对象标签对其进行描述。比如，针对某个短视频，它的对象描述文本为“最畅销的口红，某明星带货!”，对象标签为“美妆”、“明星”。

[0052] 在实际应用中，比如，针对短视频A，可以将它的视频标题“xxx推荐的口红”作为对象描述文本，并获取它对应的对象标签“美妆”和“明星”。为方便后续步骤中的模型训练，可以对对象描述文本进行分词，将其分为“xxx”、“推荐”、“的”、以及“口红”四个对象描述词语。其中，可以利用x1～x4表示四个对象描述词语“xxx”、“推荐”、“的”、以及“口红”，利用y1和y2表示两个对象标签“美妆”和“明星”。

[0053] 202、从多个对象描述词语中，确定出待预测词语、以及多个待输入词语。

[0054] 由于在自然语言处理领域中，需要将文本的内容映射为向量，该向量包含了文本中的关键信息，因此可以应用得到的向量代替文本本身，进行寻找相似内容、匹配目标客户等。在利用向量进行匹配等步骤中，自然希望该向量能够尽可能地反应文本中所包含的信息，向量中包含的信息越多，则匹配等步骤所得到的结果也就越准确。

[0055] 其中，由于一个对象描述文本不仅能够划分为多个对象描述词语，还可以与对象标签相对应，因此，为了使得对象描述文本所对应的向量中包含的信息更全面，在利用向量对对象描述文本进行表示的时候，可以使得该向量不仅包括对象描述文本本身所携带的信息，还包括对象描述文本对应的对象标签中所携带的信息。本申请实施例可以利用词向量生成模型达到这样的效果。

[0056] 其中，词向量生成模型可以为能够生成对象描述文本中多个对象描述词语所对应的词向量的网络模型。其中，如图4所示，该词向量生成模型中可以包括词语预测模型、以及标签预测模型，左边部分的网络模型为词语预测模型，右边部分的网络模型为标签预测模型，并且词语预测模型、与标签预测模型为参数共享的两个模型。

[0057] 其中，词语预测模型可以为能够预测出文本中缺失的词语的网络模型，比如，如图4所示，文本A“x1x2x3x4”，由“x1”、“x2”、“x3”、以及“x4”四个词语按顺序排列组成，现将“x2”从文本A中抠除，并将“x1”、“x2”、以及“x4”输入至词语预测模型中，则基于该词语预测模型可以预测出缺失的词语为“x2”。该词语预测模型可以采用word2vec中的cbow结构。

[0058] 其中，标签预测模型可以为能够预测出文本所对应对象标签的网络模型。比如，如图4所示，文本A“x1x2x3x4”，对应着对象标签“y1”，那么将“x1”、“x2”、“x3”、以及“x4”输入至标签预测模型中，就可以预测得到文本A所对应的对象标签“y1”。

[0059] 由于本申请实施例提供的词向量生成模型中，包括两个处理不同任务的网络模型：词语预测模型、以及标签预测模型，因此，本申请实施例属于采用同一个网络模型，同时完成多个任务的情况，这种任务可以称为多任务学习。在多任务学习中，可以利用网络模型之间共享的参数对整体模型进行优化，从而在根本上减少了学习新概念所需的数据量。因此，该词向量生成模型中需要包括参数共享的部分，也即，词语预测模型和标签预测模型中包括参数共享的部分，但是词语预测模型和标签预测模型中的其他部分，由于各自任务的不同而不同。比如，如图4所示，词语预测模型中的全连接层，可以和标签预测模型中的全连接层参数共享，而词语预测模型中的分类层、与标签预测模型中的分类层，会由于各自任务的不同而不同。

[0060] 在实际应用中，由于词语预测模型是能够预测出文本中缺失的词语的网络模型，因此，在对词语预测模型进行训练时，需要提供相应的训练样本，比如，可以从多个对象描述词语中，确定出待预测词语、以及多个待输入词语，其中，该待预测词语可以作为从对象描述文本中抠除的词语，词语预测模型需要根据多个待输入词语，预测出该待预测词语。

[0061] 在一实施例中，具体地，步骤“从所述多个对象描述词语中，确定出待预测词语、以及多个待输入词语”，可以包括：

[0062] 从所述多个对象描述词语中，确定待预测词语；

[0063] 从所述多个对象描述词语中删除所述待预测词语，得到多个待输入词语。

[0064] 在实际应用中，比如，对象描述文本A“x1x2x3x4”中包括“x1”、“x2”、“x3”、以及

“x4”四个对象描述词语，为了保证模型训练的准确性，因此，可以对象描述文本中随机抠除一个词语，如抠除“x2”，此时该对象描述词语“x2”就可以确定为待预测词语，并且“x1”、“x3”、以及“x4”可以确定为多个待输入词语。

[0065] 203、基于多个待输入词语、以及待预测词语，对词语预测模型进行训练，得到词语预测模型对应的第一模型参数。

[0066] 其中，由于词向量生成模型中包括词语预测模型、以及标签预测模型，因此在对词向量生成模型进行训练的过程中，包括了对词语预测模型的训练过程、以及对标签预测模型的训练过程。比如，可以采用对词语预测模型、以及标签预测模型进行交替训练的方法，完成对词向量生成模型的训练。

[0067] 其中，交替训练就是交替地对词语预测模型、以及标签预测模型进行训练，比如，可以先训练词语预测模型，再训练标签预测模型，再训练词语预测模型，再训练标签预测模型，直至两个模型都训练完毕，此时可以得到需要的训练后网络模型；或者可以先训练标签预测模型，再训练词语预测模型，再训练标签预测模型，再训练词语预测模型，直至两个模型都训练完毕，并得到需要的训练后网络模型，等等。

[0068] 在实际应用中，比如，在采用对词语预测模型、以及标签预测模型进行交替训练的方法，完成对词向量生成模型训练的过程中，可以先训练词语预测模型，再训练标签预测模型，再训练词语预测模型，再训练标签预测模型，直至两个模型都训练完毕，得到训练后网络模型。那么在对词语预测模型进行训练的过程中，可以基于多个待输入词语、以及待预测词语，对词语预测模型进行训练，并得到词语预测模型对应的第一模型参数。

[0069] 在一实施例中，比如，在进行模型训练的过程中，可以首先将标签预测模型的模型参数进行固定，以保证在词语预测模型训练的过程中，不会对标签预测模型的模型参数产生影响。然后将多个待输入词语，输入至词语预测模型中，并根据待预测词语、与预测得到的结果对词语预测模型进行训练，得到词语预测模型对应的第一模型参数。

[0070] 在一实施例中，由于损失函数可以用于估计模型预测值与真实值之间的不一致程度，因此，可以利用损失函数对词语预测模型进行训练，并得到词语预测模型对应的第一模型参数。具体地，步骤“基于所述多个待输入词语、以及所述待预测词语，对词语预测模型进行训练，得到所述词语预测模型对应的第一模型参数”，可以包括：

[0071] 将所述多个待输入词语输入至词语预测模型中，并基于所述词语预测模型预测所述多个待输入词语对应的预测后词语；

[0072] 基于所述预测后词语、以及所述待预测词语，采用第一损失函数确定所述词语预测模型对应的第一模型参数。

[0073] 由于词语预测模型是能够预测出文本中缺失的词语的网络模型，因此，针对词语预测模型的训练样本的构成方式可以和word2vec相同，也即将删除待预测词语后的多个待输入词语作为输入，输入至词语预测模型中进行预测，并预测出多个对象描述词语对应的预测后标签。在获取到多个待输入词语对应的预测后词语之后，此时模型预测值即为该预测后词语，真实值即为待预测词语，可以采用第一损失函数确定该词语预测模型对应的第一模型参数。

[0074] 其中，监督学习主要用于解决两类问题，分别为回归问题、以及分类问题，回归对应于定量输出，分类对应于定性输出。简单地说，由已知数据通过计算得到一个明确的值就

是回归，比如，y＝f(x)就是典型的回归关系。由已知数据或者已标注的数据，通过计算得到一个类别就是分类。由于该词语预测模型的任务是根据多个待输入词语，预测缺失的待预测词语，因此，可以在训练该词语预测模型的过程中应用分类损失函数。

[0075] 在一实施例中，该用于词向量获取的模型训练方法，可以采用第一分类损失函数确定词语预测模型所对应的第一模型参数，其中，该第一分类损失函数即为第一损失函数。比如，可以首先定义第一损失函数，该第一损失函数就是给定输入上下文的输出词语的条件概率，然后基于网络模型获取模型预测值，观察模型预测值和真实值之间的误差，并计算这些误差的梯度，在梯度方向纠正权重矩阵，通过这种方法训练得到的训练后词语预测模型，可以更准确地预测出文本中缺失的词语。

[0076] 在一实施例中，具体地，步骤“将所述多个待输入词语输入至词语预测模型中，并基于所述词语预测模型预测所述多个待输入词语对应的预测后词语”，可以包括：

[0077] 将所述多个待输入词语输入至词语预测模型中，并将所述多个待输入词语转换为多个第一待输入词向量；

[0078] 基于所述全连接层，将所述多个第一待输入词向量映射为第一待分类向量；

[0079] 基于所述第一分类层、以及所述第一待分类向量，预测所述多个待输入词语对应的预测后词语。

[0080] 其中，词语预测模型中可以包括全连接层和第一分类层，其中，词语预测模型中的全连接层、与标签预测模型中的全连接层参数共享，即二者的全连接层是同样的结构。由于词语预测模型是能够预测出文本中缺失的词语的网络模型，因此，若词表大小为N，那么第一分类层是N分类。

[0081] 在实际应用中，比如，可以首先将多个待输入词语转换为多个第一待输入词向量，也即将词语映射为词向量，然后将多个第一待输入词向量输入至全连接层中，并利用全连接层，将多个第一待输入词向量映射为第一待分类向量，然后将第一待分类向量输入至第一分类层中，并利用第一分类层预测多个待输入词语中缺失的预测后词语。

[0082] 204、基于对象描述文本、以及对象标签，对标签预测模型进行训练，得到标签预测模型对应的第二模型参数。

[0083] 在实际应用中，比如，在采用对词语预测模型、以及标签预测模型进行交替训练的方法，完成对词向量生成模型训练的过程中，可以先训练词语预测模型，再训练标签预测模型，再训练词语预测模型，再训练标签预测模型，直至两个模型都训练完毕，得到训练后网络模型。那么在对标签预测模型进行训练的过程中，可以基于对象描述文本、以及对象标签，对标签预测模型进行训练，得到标签预测模型对应的第二模型参数。

[0084] 在一实施例中，在模型训练得到词语预测模型对应的第一模型参数之后，基于模型的交替训练，可以对该词语预测模型对应的第一模型参数进行固定，以保证在标签预测模型训练的过程中，不会对词语预测模型的模型参数产生影响。然后将多个对象描述词语，输入至标签预测模型中，并根据对象标签、与预测得到的结果对标签预测模型进行训练，得到标签预测模型对应的第二模型参数。

[0085] 在一实施例中，由于损失函数可以用于估计模型预测值与真实值之间的不一致程度，因此，可以利用损失函数对标签预测模型进行训练，并得到标签预测模型对应的第二模型参数。具体地，步骤“基于所述对象描述文本、以及所述对象标签，对标签预测模型进行训

练，得到所述标签预测模型对应的第二模型参数”，可以包括：

[0086] 将所述多个对象描述词语输入至标签预测模型中，并基于所述标签预测模型预测所述多个对象描述词语对应的预测后标签；

[0087] 基于所述预测后标签、以及所述对象标签，采用第二损失函数确定所述标签预测模型对应的第二模型参数。

[0088] 由于标签预测模型是能够预测出文本所对应对象标签的网络模型，因此，针对标签预测模型的训练样本的构成方式可以和文本分类相同，也即将对象描述文本分词后得到的多个对象描述词语作为输入，输入至标签预测模型中进行预测，并预测出多个待输入词语对应的预测后词语。其中，对于一个对象描述文本而言，可能对应不止一个对象标签，此时，针对标签预测模型而言，有几个对象标签就可以对应几个训练样本。在获取到多个对象描述词语对应的预测后标签之后，此时模型预测值即为该预测后标签，真实值即为对象标签，可以采用第二损失函数确定该标签预测模型对应的第二模型参数。

[0089] 其中，监督学习主要用于解决两类问题，分别为回归问题、以及分类问题，回归对应于定量输出，分类对应于定性输出。简单地说，由已知数据通过计算得到一个明确的值就是回归，比如，y＝f(x)就是典型的回归关系。由已知数据或者已标注的数据，通过计算得到一个类别就是分类。由于该标签预测模型的任务是根据多个对象描述词语，预测出对象描述文本对应的对象标签，因此，可以在训练该标签预测模型的过程中应用分类损失函数。

[0090] 在一实施例中，该用于词向量获取的模型训练方法，可以采用第二分类损失函数确定标签预测模型所对应的第二模型参数，其中，该第二分类损失函数即为第二损失函数。比如，可以首先根据多个对象描述词语，确定在各个标签所对应的预测概率值，然后根据各个标签所对应的预测概率值，确定第二损失函数的分类损失，当第二损失函数的分类损失为最大值时，可以确定标签预测模型对应的第二模型参数。

[0091] 在一实施例中，具体地，步骤“将所述多个对象描述词语输入至标签预测模型中，并基于所述标签预测模型预测所述多个对象描述词语对应的预测后标签”，可以包括：

[0092] 将所述多个对象描述词语输入至标签预测模型中，并将所述多个对象描述词语转换为多个第二待输入词向量；

[0093] 基于所述全连接层，将所述多个第二待输入词向量映射为第二待分类向量；

[0094] 基于所述第二分类层、以及所述第二待分类向量，预测所述多个对象描述词语对应的预测后标签。

[0095] 其中，标签预测模型中可以包括全连接层和第二分类层，其中，词语预测模型中的全连接层、与标签预测模型中的全连接层参数共享，即二者的全连接层是同样的结构。由于标签预测模型是能够预测出文本对应的标签的网络模型，因此，若标签数量为M，那么第二分类层是M分类，与词语预测模型中的第一分类层不同。

[0096] 在实际应用中，比如，可以首先将多个对象描述词语转换为多个第二待输入词向量，也即将词语映射为词向量，然后将多个第二待输入词向量输入至全连接层中，并利用全连接层，将多个第二待输入词向量映射为第二待分类向量，然后将第二待分类向量输入至第二分类层中，并利用第二分类层预测对象描述文本对应的预测后对象标签。

[0097] 205、基于第一模型参数、以及第二模型参数，对词语预测模型、以及标签预测模型进行交替训练，得到用于获取对象描述文本对应的目标词向量的训练后网络模型。

[0098] 在实际应用中，比如，可以基于第一模型参数、以及第二模型参数，对词语预测模型、以及标签预测模型进行交替训练，得到用于获取对象描述文本对应的目标词向量的训练后网络模型，其中，训练后网络模型中包括训练后标签预测模型、以及训练后词语预测模型。根据该训练后网络模型可以获取到同时包括对象描述文本中信息、以及对象标签中信息的目标词向量，从而使得对象描述文本本身的上下文信息、与对象描述文本对应的对象标签信息可以充分地融合，并体现在词向量中。

[0099] 在一实施例中，比如，对词语预测模型进行训练，得到词语预测模型对应的第一模型参数，并对标签预测模型进行训练，得到标签预测模型对应的第二模型参数之后，可以根据模型训练得到的模型参数(包括第一模型参数、以及第二模型参数)，对词向量生成模型进行训练，得到训练后网络模型。可以理解为，在获取到第一模型参数、以及第二模型参数之后，再重复上述交替训练的步骤，不断地进行模型训练，直至整个模型收敛，得到训练后网络模型。

[0100] 在一实施例中，在对词语预测模型、以及标签预测模型进行交替训练的过程中，可以先固定标签预测模型的模型参数，并对词语预测模型进行训练，得到词语预测模型对应的模型参数。也可以先固定词语预测模型的模型参数，并对标签预测模型进行训练，得到标签预测模型对应的模型参数。本申请以先固定标签预测模型的模型参数，并对词语预测模型进行训练为例进行介绍，然而，这并不应理解为对本申请的限定。

[0101] 在一实施例中，由于交替训练是需要不断地对多个模型进行交替地训练，因此针对每个网络模型，都要经历不止一次的训练过程。具体地，步骤“基于所述第一模型参数、以及所述第二模型参数，对所述词语预测模型、以及所述标签预测模型进行交替训练，得到用于获取所述对象描述文本对应的目标词向量的训练后网络模型”，可以包括：

[0102] 基于所述多个待输入词语、以及所述待预测词语，对词语预测模型进行训练，得到所述词语预测模型对应的第三模型参数；

[0103] 基于所述对象描述文本、以及所述对象标签，对标签预测模型进行训练，得到所述标签预测模型对应的第四模型参数；

[0104] 基于所述第一模型参数、所述第二模型参数、所述第三模型参数、以及所述第四模型参数，对所述词语预测模型、以及所述标签预测模型进行交替训练，得到用于获取所述对象描述文本对应的目标词向量的训练后网络模型。

[0105] 在实际应用中，比如，在经历了一轮模型交替训练，得到词语预测模型对应的第一模型参数、以及标签预测模型对应的第二模型参数之后，既可开始下一轮的模型训练。在训练词语预测模型的时候，需要先固定标签预测模型的权重值，即固定标签预测模型的第二模型参数，然后将多个待输入词语输入至词语预测模型中，由该词语预测模型预测出缺失的预测后词语。并且采用分类损失函数对词语预测模型进行训练，得到词语预测模型对应的第三模型参数。

[0106] 在训练得到词语预测模型对应的第三模型参数之后，基于模型的交替训练，需要固定词语预测模型的权重值，即固定词语预测模型对应的第三模型参数，然后将对象描述文本输入至标签预测模型中，由该标签预测模型预测出对象描述文本所对应的预测后标签。并且采用分类损失函数对标签预测模型进行训练，得到标签预测模型对应的第四模型参数。在经过多轮交替训练之后，可以根据每轮训练得到的模型参数，包括第一模型参数、

第二模型参数、第三模型参数、第四模型参数等等，对词向量生成模型进行训练，并得到训练后网络模型。

[0107] 通过上述方式，固定一个网络模型的模型参数，训练另一个网络模型，能够使得该词向量生成网络模型不断地进行对抗学习，一方面使得词语预测模型能够预测得到更为准确的词语，另一方面还能使得标签预测模型能够预测得到更为准确的标签。

[0108] 在一实施例中，还可以利用该训练后网络模型进行目标词向量的获取，该目标词向量中同时包括对象描述文本中信息、以及对象标签中信息的目标词向量。具体地，该用于词向量获取的模型训练方法，还可以包括：

[0109] 基于所述训练后网络模型，将所述对象描述文本表示为目标词向量；

[0110] 基于所述训练后网络模型，将所述对象标签表示为目标标签词向量；

[0111] 基于所述训练后网络模型的全连接层，对所述目标词向量、以及所述目标标签词向量进行融合，得到所述对象描述文本对应的目标文本表示。

[0112] 在实际应用中，比如，模型训练完毕得到训练后网络模型之后，可以利用该训练后网络模型，将多个对象描述词语表示为目标词向量，并获取到对象标签对应的目标标签词向量。此时可以固定住全连接层，并将每个对象描述词语对应的目标词向量、以及对象标签对应的目标标签词向量，输入至全连接层中，得到一个输出的向量，该向量便是对象描述文本对应的目标文本表示。由于本申请的用于词向量获取的模型训练方法，不仅使用到了对象描述文本本身的信息，还用到了对象标签的信息，因此，训练出来的目标词向量中，不仅包含了对象描述文本本身的信息，还包含了对象标签的信息。

[0113] 在一实施例中，还可以输入对象描述文本对应的多个对象描述词语、以及对象描述文本对应的对象标签，并分别查表得到他们对应的词向量，输入至全连接层中，得到一个输出的向量，该向量便是对象描述文本对应的目标文本表示。

[0114] 在一实施例中，获取到每个对象描述词语对应的目标词向量后，可以将该融合了标签信息的词向量，存储至基础词向量库中，并服务于更高层的自然语言处理任务，如词性标注、命名实体识别、文本摘要等等。

[0115] 在一实施例中，具体地，该用于词向量获取的模型训练方法还可以包括：

[0116] 获取用于描述目标视频的视频描述文本、以及视频标签，其中，所述视频描述文本中包括多个视频描述词语；

[0117] 从所述多个视频描述词语中，确定出待预测词语、以及多个待输入词语；

[0118] 基于标签预测模型预测所述多个对象描述词语对应的预测后标签，并基于词语预测模型预测所述多个待输入词语对应的预测后词语；

[0119] 基于所述预测后标签、所述对象标签、所述预测后词语、以及所述待预测词语，对所述词语预测模型、以及所述标签预测模型进行交替训练，得到训练后网络模型，其中，所述训练后网络模型中包括训练后标签预测模型、以及训练后词语预测模型；

[0120] 基于所述训练后网络模型，将所述多个视频描述词语表示为目标词向量，并将所述视频标签表示为目标标签词向量；

[0121] 基于所述训练后网络模型的全连接层，对所述目标词向量、以及所述目标标签词向量进行融合，得到所述视频描述文本对应的目标文本表示；

[0122] 获取视频集合中每个视频对应的视频文本、以及所述视频文本对应的视频文本表

示；

[0123] 基于所述目标文本表示、以及所述视频文本表示，获取所述视频描述文本与所述视频文本之间的相似度；

[0124] 基于所述相似度，从所述视频集合的多个视频中，确定与所述目标视频相似的相似视频。

[0125] 在实际应用中，比如，本申请可以应用于相似视频推荐场景中，可以首先获取需要推荐相似视频的目标视频、以及目标视频对应的视频描述文本和视频标签，并根据本申请的用于词向量获取的模型训练方法，获取视频描述文本所对应的目标文本表示。并获取视频集合，该视频集合中包括多个视频、以及每个视频对应的视频文本表示，此时，可以将目标文本表示和视频文本表示进行相似度计算，当计算得到的相似度数值超过相似度阈值时，就可以将该视频文本表示所对应的视频，确定为与目标视频相似的相似视频。

[0126] 本申请可以应用于许多场景，如公众号文章检索、公众号文章推荐、商品检索、商品推荐、短视频检索、短视频推荐、音乐检索、音乐推荐等。并且可以利用本申请获取到的目标对象对应的目标文本表示、和其他目标对象的向量之间计算相似度，来寻找相似目标对象；或者利用本申请获取到的目标对象对应的目标文本表示、和用户的向量之间计算相似度，来寻找匹配的用户；或者利用本申请获取到的目标对象对应的目标文本表示，作为内容的特征服务于更大的算法系统，等等。

[0127] 其中，在某些情况下，针对同一个目标对象而言，该目标对象对应对象描述文本的描述更注重细节，而该目标对象对应对象标签的描述更加宏观，比如，针对某个短视频，它的对象描述文本为“最畅销的口红，某明星带货!”，该短视频可以对应着“美妆”、“明星”的对象标签，也即该对象描述文本已经精确到了“口红”和“某明星”，而对象标签则是更宽泛的“美妆”和“明星”，也即，对象描述文本和对象标签之间是粗细粒度关系。

[0128] 其中，在某些情况下，对象描述文本和对象标签之间还可能存在互补关系，比如，针对某篇文章，其对应的对象描述文本是“震惊!活到40岁才明白”，从对象描述文本看，只能了解到这篇文章和人生、年龄相关，但是通过获取它对应的标签“养生”、“健康”，才能进一步地了解到该文章所描述的大概内容。

[0129] 无论是针对哪种情况，对对象描述文本进行表示时，如果只针对对象描述文本本身进行文本表示，就会忽略掉对象标签信息，无法表达对象描述文本和对象标签之间的“粗细粒度关系”或者“互补关系”。而本申请的用于词向量获取的模型训练方法，则可以通过训练词向量生成模型，得到训练后网络模型，使得利用该训练后网络模型获取到的目标词向量中，不仅包括对象描述文本本身的信息，还包括对象标签的信息。并且最终获取到的对象描述文本对应的目标文本表示，既能表征不同的粗细粒度内容，也能更全面地包含对象标签信息、以及对象描述文本信息。利用这种方法获取到的目标词向量或者目标文本表示，由于其包括了更丰富完整的信息，因此可以使得后续的匹配、相似度计算等工作，得到的结果准确性更高，更可靠。

[0130] 由上可知，本申请实施例可以获取用于描述目标对象的对象描述文本、以及对象标签，其中，对象描述文本中包括多个对象描述词语，从多个对象描述词语中，确定出待预测词语、以及多个待输入词语，基于多个待输入词语、以及待预测词语，对词语预测模型进行训练，得到词语预测模型对应的第一模型参数，基于对象描述文本、以及对象标签，对标

签预测模型进行训练，得到标签预测模型对应的第二模型参数，基于第一模型参数、以及第二模型参数，对词语预测模型、以及标签预测模型进行交替训练，得到用于获取对象描述文本对应的目标词向量的训练后网络模型。该方案可以通过模型训练得到训练后网络模型，使得利用该训练后网络模型获取到的目标词向量中，不仅包括对象描述文本本身的信息，还包括对象标签的信息。并且最终获取到的对象描述文本对应的目标文本表示，既能表征不同的粗细粒度内容，也能更全面地包含对象标签信息、以及对象描述文本信息。利用这种方法获取到的目标词向量或者目标文本表示，由于其包括了更丰富完整的信息，因此可以使得后续的匹配、相似度计算等工作，得到的结果准确性更高，更可靠。

[0131] 根据前面实施例所描述的方法，以下将以该用于词向量获取的模型训练装置具体集成在电子设备中举例作进一步详细说明。

[0132] 参考图3，本申请实施例的用于词向量获取的模型训练方法的具体流程可以如下：

[0133] 301、获取短文本分词后得到的多个词语、以及短文本对应的标签。

[0134] 在实际应用中，比如，可以获取短文本A，并对该短文本A进行分词，将该短文本A划分为四个词语：x1、x2、x3、以及x4。同时获取该短文本A所对应的标签y1、以及y2。

[0135] 302、确定词向量生成模型。

[0136] 在实际应用中，比如，可以确定如图4所示的词向量生成模型，该词向量生成模型中包括词语预测模型、以及标签预测模型，其中，词语预测模型中包括全连接层和第一分类层，可以用于预测文本中缺失的词语；标签预测模型中包括全连接层和第二分类层，可以用于预测短文本所对应的标签。并且，词语预测模型中的全连接层、与标签预测模型中的全连接层参数共享，但是由于他们各自的任务不同，因此，词语预测模型中的第一分类层、与标签预测模型中的第二分类层不同，若词表大小为N，标签数量为M，则第一分类层为N分类，第二分类层为M分类。

[0137] 303、随机从多个词语中抠除待预测词语，并将剩余词语确定为待输入词语。

[0138] 在实际应用中，比如，由于词语预测模型是典型的word2vec里的cbow结构，也即输入是在短文本分词后得到的多个词语中抠除某个词语后得到的文本，输出是被抠除的词语。因此，对于词语预测模型的训练样本的构成方式可以与word2vec相同，也即随机从多个词语中抠除一个词语，该被抠除的词语即为待预测词语，并将剩余词语确定为待输入词语，将待输入词语输入至词语预测模型中以完成预测。

[0139] 在一实施例中，比如，可以从多个词语x1、x2、x3、以及x4中，随机抠除x2，此时，x2即为待预测词语，那么剩余的词语x1、x3、以及x4，即为多个待输入词语。

[0140] 304、将多个待输入词语输入至词语预测模型中，并将多个词语输入至标签预测模型中。

[0141] 在实际应用中，比如，如图4所示，可以将多个待输入词语x1、x3、以及x4，输入至词语预测模型中，将多个词语x1、x2、x3、以及x4，输入至标签预测模型中。

[0142] 305、对词语预测模型进行训练，得到词语预测模型对应的第一模型参数。

[0143] 在实际应用中，比如，需要对词语预测模型、以及标签预测模型进行交替训练时，可以选择首先对词语预测模型进行训练，在对词语预测模型进行训练的过程中，需要首先将标签预测模型的模型参数进行固定，以保证在词语预测模型训练的过程中，不会对标签预测模型的模型参数产生影响。然后将多个待输入词语x1、x3、以及x4，输入至词语预测模

型中，基于词语预测模型将多个待输入词语转换为多个第一待输入词向量。然后将多个第一待输入词向量输入至全连接层中，并利用全连接层，将多个第一待输入词向量映射为第一待分类向量，然后将第一待分类向量输入至第一分类层中，并利用第一分类层预测多个待输入词语中缺失的预测后词语x2'。然后采用损失函数对词语预测模型进行训练，并得到词语预测模型对应的第一模型参数。

[0144] 306、对标签预测模型进行训练，得到标签预测模型对应的第二模型参数。

[0145] 在实际应用中，比如，在模型训练得到词语预测模型对应的第一模型参数之后，基于模型的交替训练，可以对该词语预测模型对应的第一模型参数进行固定，以保证在标签预测模型训练的过程中，不会对词语预测模型的模型参数产生影响。然后将多个词语x1、x2、x3、以及x4，输入至标签预测模型中，基于标签预测模型将多个词语转换为多个第二待输入词向量。然后将多个第二待输入词向量输入至全连接层中，并利用全连接层，将多个第二待输入词向量映射为第二待分类向量，然后将第二待分类向量输入至第二分类层中，并利用第二分类层预测短文本A对应的预测后标签y1'。然后采用损失函数对标签预测模型进行训练，并得到标签预测模型对应的第二模型参数。

[0146] 在一实施例中，在对词语预测模型、以及标签预测模型进行交替训练的过程中，可以先固定标签预测模型的模型参数，并对词语预测模型进行训练，得到词语预测模型对应的模型参数。也可以先固定词语预测模型的模型参数，并对标签预测模型进行训练，得到标签预测模型对应的模型参数。本申请以先固定标签预测模型的模型参数，并对词语预测模型进行训练为例进行介绍，然而，这并不应理解为对本申请的限定。

[0147] 其中，对于标签预测模型而言，它的训练样本的构成方式可以与文本分类一样，也即将短文本分词后的多个词语作为输入，标签作为输出，由于一个短文本可能对应不止一个标签，因此，有几个标签就对应几个训练样本。

[0148] 307、基于第一模型参数、以及第二模型参数，对词向量生成模型进行训练，得到训练后网络模型。

[0149] 在实际应用中，比如，在经过至少一轮交替训练之后，可以根据模型训练得到的模型参数(包括第一模型参数、以及第二模型参数)，对词向量生成模型进行训练，得到训练后网络模型。可以理解为，在获取到第一模型参数、以及第二模型参数之后，再重复上述交替训练的步骤，不断地进行模型训练，直至整个模型收敛，得到训练后网络模型。

[0150] 308、基于训练后网络模型，将多个词语表示为目标词向量，并将标签表示为目标标签词向量。

[0151] 在实际应用中，比如，得到训练后网络模型之后，就可以利用该训练后网络模型，分别获取词语x1、x2、x3、以及x4对应的目标词向量、以及标签y1、y2对应的目标标签词向量。此时获取到的词向量中就不仅包括短文本A本身的信息，还包括短文本A所对应标签的信息，大大丰富了词向量所携带的内容。

[0152] 在一实施例中，获取到每个对象描述词语对应的目标词向量后，可以将该融合了标签信息的词向量，存储至基础词向量库中，并服务于更高层的自然语言处理任务，如词性标注、命名实体识别、文本摘要等等。

[0153] 309、基于训练后网络模型的全连接层，对目标词向量、以及目标标签词向量进行融合，得到短文本对应的目标文本表示。

[0154] 在实际应用中，比如，如图5所示，获取到每个词语对应的目标词向量、以及标签对应的目标标签词向量后，可以固定住全连接层，并将目标词向量、以及目标标签词向量输入至全连接层中，得到一个输出的向量，该向量便是短文本A对应的目标文本表示。

[0155] 在一实施例中，可以利用本申请获取到的目标对象对应的目标文本表示、和其他目标对象的向量之间计算相似度，来寻找相似目标对象；或者利用本申请获取到的目标对象对应的目标文本表示、和用户的向量之间计算相似度，来寻找匹配的用户；或者利用本申请获取到的目标对象对应的目标文本表示，作为内容的特征服务于更大的算法系统，等等。

[0156] 由上可知，本申请实施例可以通过电子设备获取短文本分词后得到的多个词语、以及短文本对应的标签，确定词向量生成模型，随机从多个词语中抠除待预测词语，并将剩余词语确定为待输入词语，将多个待输入词语输入至词语预测模型中，并将多个词语输入至标签预测模型中，对词语预测模型进行训练，得到词语预测模型对应的第一模型参数，对标签预测模型进行训练，得到标签预测模型对应的第二模型参数，基于第一模型参数、以及第二模型参数，对词向量生成模型进行训练，得到训练后网络模型，基于训练后网络模型，将多个词语表示为目标词向量，并将标签表示为目标标签词向量，基于训练后网络模型的全连接层，对目标词向量、以及目标标签词向量进行融合，得到短文本对应的目标文本表示。该方案可以通过训练词向量生成模型，得到训练后网络模型，使得利用该训练后网络模型获取到的目标词向量中，不仅包括对象描述文本本身的信息，还包括对象标签的信息。并且最终获取到的对象描述文本对应的目标文本表示，既能表征不同的粗细粒度内容，也能更全面地包含对象标签信息、以及对象描述文本信息。利用这种方法获取到的目标词向量或者目标文本表示，由于其包括了更丰富完整的信息，因此可以使得后续的匹配、相似度计算等工作，得到的结果准确性更高，更可靠。

[0157] 为了更好地实施以上方法，相应的，本申请实施例还提供一种用于词向量获取的模型训练装置，该用于词向量获取的模型训练装置可以集成在电子设备中，参考图6，该用于词向量获取的模型训练装置包括获取模块61、词语确定模块62、第一训练模块63、第二训练模块64和第三训练模块65，如下：

[0158] 获取模块61，用于获取用于描述目标对象的对象描述文本、以及对象标签，其中，所述对象描述文本中包括多个对象描述词语；

[0159] 词语确定模块62，用于从所述多个对象描述词语中，确定出待预测词语、以及多个待输入词语；

[0160] 第一训练模块63，用于基于所述多个待输入词语、以及所述待预测词语，对词语预测模型进行训练，得到所述词语预测模型对应的第一模型参数；

[0161] 第二训练模块64，用于基于所述对象描述文本、以及所述对象标签，对标签预测模型进行训练，得到所述标签预测模型对应的第二模型参数；

[0162] 第三训练模块65，用于基于所述第一模型参数、以及所述第二模型参数，对所述词语预测模型、以及所述标签预测模型进行交替训练，得到用于获取所述对象描述文本对应的目标词向量的训练后网络模型。

[0163] 在一实施例中，所述第一训练模块63可以包括第一预测子模块631和第一确定子模块632，如下：

[0164] 第一预测子模块631，用于将所述多个待输入词语输入至词语预测模型中，并基于

所述词语预测模型预测所述多个待输入词语对应的预测后词语；

[0165] 第一确定子模块632，用于基于所述预测后词语、以及所述待预测词语，采用第一损失函数确定所述词语预测模型对应的第一模型参数。

[0166] 在一实施例中，所述第一预测子模块631可以具体用于：

[0167] 将所述多个待输入词语输入至词语预测模型中，并将所述多个待输入词语转换为多个第一待输入词向量；

[0168] 基于所述全连接层，将所述多个第一待输入词向量映射为第一待分类向量；

[0169] 基于所述第一分类层、以及所述第一待分类向量，预测所述多个待输入词语对应的预测后词语。

[0170] 在一实施例中，所述第二训练模块64可以包括第二预测子模块641和第二确定子模块642，如下：

[0171] 第二预测子模块641，用于将所述多个对象描述词语输入至标签预测模型中，并基于所述标签预测模型预测所述多个对象描述词语对应的预测后标签；

[0172] 第二确定子模块642，用于基于所述预测后标签、以及所述对象标签，采用第二损失函数确定所述标签预测模型对应的第二模型参数。

[0173] 在一实施例中，所述第二预测子模块641可以具体用于：

[0174] 将所述多个对象描述词语输入至标签预测模型中，并将所述多个对象描述词语转换为多个第二待输入词向量；

[0175] 基于所述全连接层，将所述多个第二待输入词向量映射为第二待分类向量；

[0176] 基于所述第二分类层、以及所述第二待分类向量，预测所述多个对象描述词语对应的预测后标签。

[0177] 在一实施例中，所述第三训练模块65可以具体用于：

[0178] 基于所述多个待输入词语、以及所述待预测词语，对词语预测模型进行训练，得到所述词语预测模型对应的第三模型参数；

[0179] 基于所述对象描述文本、以及所述对象标签，对标签预测模型进行训练，得到所述标签预测模型对应的第四模型参数；

[0180] 基于所述第一模型参数、所述第二模型参数、所述第三模型参数、以及所述第四模型参数，对所述词语预测模型、以及所述标签预测模型进行交替训练，得到用于获取所述对象描述文本对应的目标词向量的训练后网络模型。

[0181] 在一实施例中，所述用于词向量获取的模型训练装置还可以包括标签表示模块66、第二表示模块67和融合模块68，如下：

[0182] 第一表示模块66，用于基于所述训练后网络模型，将所述对象描述文本表示为目标词向量；

[0183] 第二表示模块67，用于基于所述训练后网络模型，将所述对象标签表示为目标标签词向量；

[0184] 融合模块68，用于基于所述训练后网络模型的全连接层，对所述目标词向量、以及所述目标标签词向量进行融合，得到所述对象描述文本对应的目标文本表示。

[0185] 在一实施例中，所述词语确定模块62可以具体用于：

[0186] 从所述多个对象描述词语中，确定待预测词语；

[0187] 从所述多个对象描述词语中删除所述待预测词语，得到多个待输入词语。

[0188] 具体实施时，以上各个单元可以作为独立的实体来实现，也可以进行任意组合，作为同一或若干个实体来实现，以上各个单元的具体实施可参见前面的方法实施例，在此不再赘述。

[0189] 由上可知，本申请实施例可以通过获取模块61获取用于描述目标对象的对象描述文本、以及对象标签，其中，对象描述文本中包括多个对象描述词语，通过词语确定模块62从多个对象描述词语中，确定出待预测词语、以及多个待输入词语，通过第一训练模块63基于多个待输入词语、以及待预测词语，对词语预测模型进行训练，得到词语预测模型对应的第一模型参数，通过第二训练模块64基于对象描述文本、以及对象标签，对标签预测模型进行训练，得到标签预测模型对应的第二模型参数，通过第三训练模块65基于第一模型参数、以及第二模型参数，对词语预测模型、以及标签预测模型进行交替训练，得到用于获取对象描述文本对应的目标词向量的训练后网络模型。该方案可以通过模型训练得到训练后网络模型，使得利用该训练后网络模型获取到的目标词向量中，不仅包括对象描述文本本身的信息，还包括对象标签的信息。并且最终获取到的对象描述文本对应的目标文本表示，既能表征不同的粗细粒度内容，也能更全面地包含对象标签信息、以及对象描述文本信息。利用这种方法获取到的目标词向量或者目标文本表示，由于其包括了更丰富完整的信息，因此可以使得后续的匹配、相似度计算等工作，得到的结果准确性更高，更可靠。

[0190] 本申请实施例还提供一种电子设备，该电子设备可以集成本申请实施例所提供的任一种用于词向量获取的模型训练装置。

[0191] 例如，如图7所示，其示出了本申请实施例所涉及的电子设备的结构示意图，具体来讲：

[0192] 该电子设备可以包括一个或者一个以上处理核心的处理器71、一个或一个以上计算机可读存储介质的存储器72、电源73和输入单元74等部件。本领域技术人员可以理解，图7中示出的电子设备结构并不构成对电子设备的限定，可以包括比图示更多或更少的部件，或者组合某些部件，或者不同的部件布置。

[0193] 其中：

[0194] 处理器71是该电子设备的控制中心，利用各种接口和线路连接整个电子设备的各个部分，通过运行或执行存储在存储器72内的软件程序和/或模块，以及调用存储在存储器72内的数据，执行电子设备的各种功能和处理数据，从而对电子设备进行整体检测。可选的，处理器71可包括一个或多个处理核心；优选的，处理器71可集成应用处理器和调制解调处理器，其中，应用处理器主要处理操作系统、用户界面和应用程序等，调制解调处理器主要处理无线通信。可以理解的是，上述调制解调处理器也可以不集成到处理器71中。

[0195] 存储器72可用于存储软件程序以及模块，处理器71通过运行存储在存储器72的软件程序以及模块，从而执行各种功能应用以及数据处理。存储器72可主要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、至少一个功能所需的应用程序(比如声音播放功能、图像播放功能等)等；存储数据区可存储根据电子设备的使用所创建的数据等。此外，存储器72可以包括高速随机存取存储器，还可以包括非易失性存储器，例如至少一个磁盘存储器件、闪存器件、或其他易失性固态存储器件。相应地，存储器72还可以包括存储器控制器，以提供处理器71对存储器72的访问。

[0196] 电子设备还包括给各个部件供电的电源73，优选的，电源73可以通过电源管理系统与处理器71逻辑相连，从而通过电源管理系统实现管理充电、放电、以及功耗管理等功能。电源73还可以包括一个或一个以上的直流或交流电源、再充电系统、电源故障检测电路、电源转换器或者逆变器、电源状态指示器等任意组件。

[0197] 该电子设备还可包括输入单元74，该输入单元74可用于接收输入的数字或字符信息，以及产生与用户设置以及功能控制有关的键盘、鼠标、操作杆、光学或者轨迹球信号输入。

[0198] 尽管未示出，电子设备还可以包括显示单元等，在此不再赘述。具体在本实施例中，电子设备中的处理器71会按照如下的指令，将一个或一个以上的应用程序的进程对应的可执行文件加载到存储器72中，并由处理器71来运行存储在存储器72中的应用程序，从而实现各种功能，如下：

[0199] 获取用于描述目标对象的对象描述文本、以及对象标签，其中，对象描述文本中包括多个对象描述词语，从多个对象描述词语中，确定出待预测词语、以及多个待输入词语，基于多个待输入词语、以及待预测词语，对词语预测模型进行训练，得到词语预测模型对应的第一模型参数，基于对象描述文本、以及对象标签，对标签预测模型进行训练，得到标签预测模型对应的第二模型参数，基于第一模型参数、以及第二模型参数，对词语预测模型、以及标签预测模型进行交替训练，得到用于获取对象描述文本对应的目标词向量的训练后网络模型。

[0200] 以上各个操作的具体实施可参见前面的实施例，在此不再赘述。

[0201] 由上可知，本申请实施例可以获取用于描述目标对象的对象描述文本、以及对象标签，其中，对象描述文本中包括多个对象描述词语，从多个对象描述词语中，确定出待预测词语、以及多个待输入词语，基于多个待输入词语、以及待预测词语，对词语预测模型进行训练，得到词语预测模型对应的第一模型参数，基于对象描述文本、以及对象标签，对标签预测模型进行训练，得到标签预测模型对应的第二模型参数，基于第一模型参数、以及第二模型参数，对词语预测模型、以及标签预测模型进行交替训练，得到用于获取对象描述文本对应的目标词向量的训练后网络模型。该方案可以通过模型训练得到训练后网络模型，使得利用该训练后网络模型获取到的目标词向量中，不仅包括对象描述文本本身的信息，还包括对象标签的信息。并且最终获取到的对象描述文本对应的目标文本表示，既能表征不同的粗细粒度内容，也能更全面地包含对象标签信息、以及对象描述文本信息。利用这种方法获取到的目标词向量或者目标文本表示，由于其包括了更丰富完整的信息，因此可以使得后续的匹配、相似度计算等工作，得到的结果准确性更高，更可靠。

[0202] 本领域普通技术人员可以理解，上述实施例的各种方法中的全部或部分步骤可以通过指令来完成，或通过指令控制相关的硬件来完成，该指令可以存储于一计算机可读存储介质中，并由处理器进行加载和执行。

[0203] 为此，本申请实施例提供一种电子设备，其中存储有多条指令，该指令能够被处理器进行加载，以执行本申请实施例所提供的任一种用于词向量获取的模型训练方法中的步骤。例如，该指令可以执行如下步骤：

[0204] 获取用于描述目标对象的对象描述文本、以及对象标签，其中，对象描述文本中包括多个对象描述词语，从多个对象描述词语中，确定出待预测词语、以及多个待输入词语，

基于多个待输入词语、以及待预测词语，对词语预测模型进行训练，得到词语预测模型对应的第一模型参数，基于对象描述文本、以及对象标签，对标签预测模型进行训练，得到标签预测模型对应的第二模型参数，基于第一模型参数、以及第二模型参数，对词语预测模型、以及标签预测模型进行交替训练，得到用于获取对象描述文本对应的目标词向量的训练后网络模型。

[0205] 以上各个操作的具体实施可参见前面的实施例，在此不再赘述。

[0206] 其中，该存储介质可以包括：只读存储器(ROM，Read Only Memory)、随机存取记忆体(RAM，Random Access Memory)、磁盘或光盘等。

[0207] 由于该存储介质中所存储的指令，可以执行本申请实施例所提供的任一种用于词向量获取的模型训练方法中的步骤，因此，可以实现本申请实施例所提供的任一种用于词向量获取的模型训练方法所能实现的有益效果，详见前面的实施例，在此不再赘述。

[0208] 以上对本申请实施例所提供的一种用于词向量获取的模型训练方法、装置和存储介质进行了详细介绍，本文中应用了具体个例对本申请的原理及实施方式进行了阐述，以上实施例的说明只是用于帮助理解本申请的方法及其核心思想；同时，对于本领域的技术人员，依据本申请的思想，在具体实施方式及应用范围上均会有改变之处，综上所述，本说明书内容不应理解为对本申请的限制。

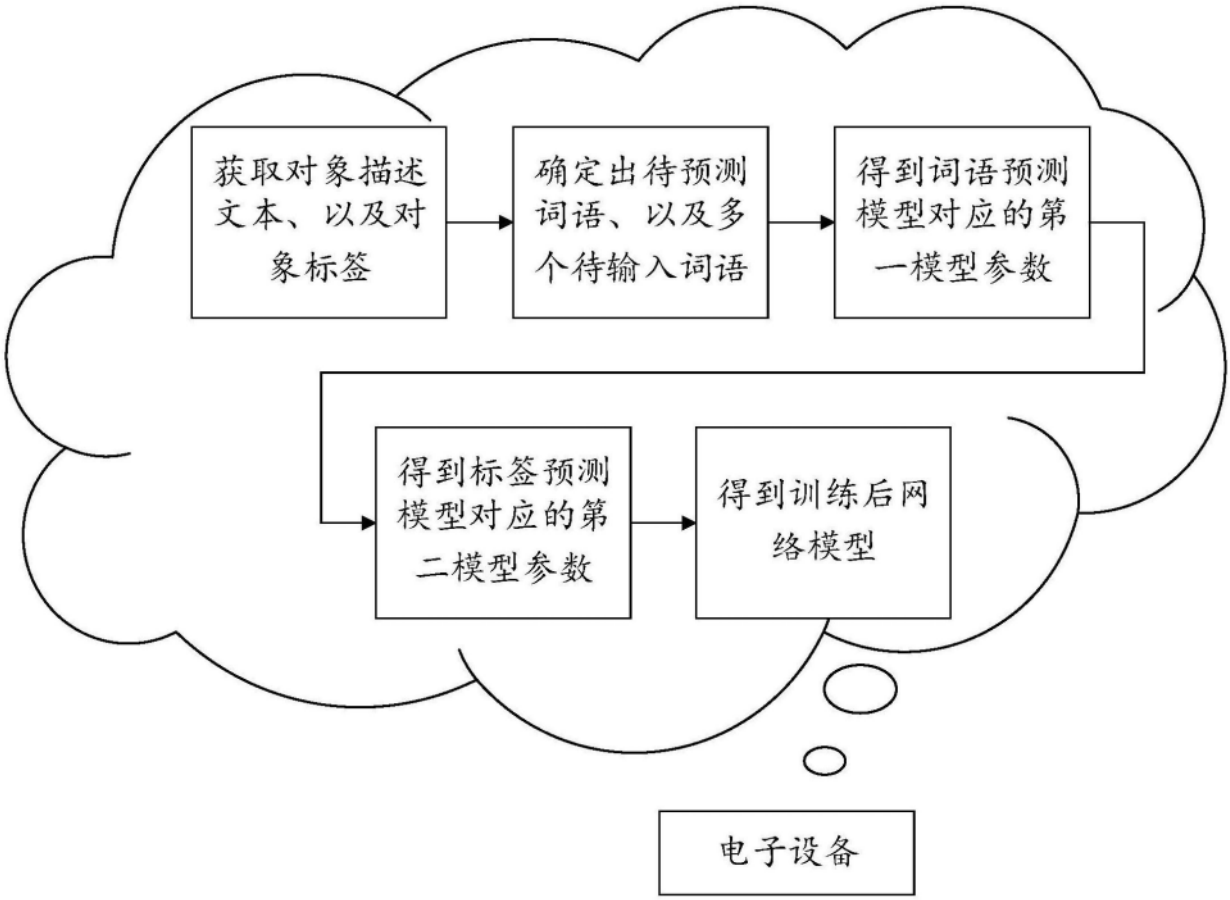

获取对象描述文本、以及对象标签
确定出待预测词语、以及多个待输入词语
得到词语预测模型对应的第一模型参数
得到标签预测模型对应的第二模型参数
得到训练后网络模型
电子设备

图1

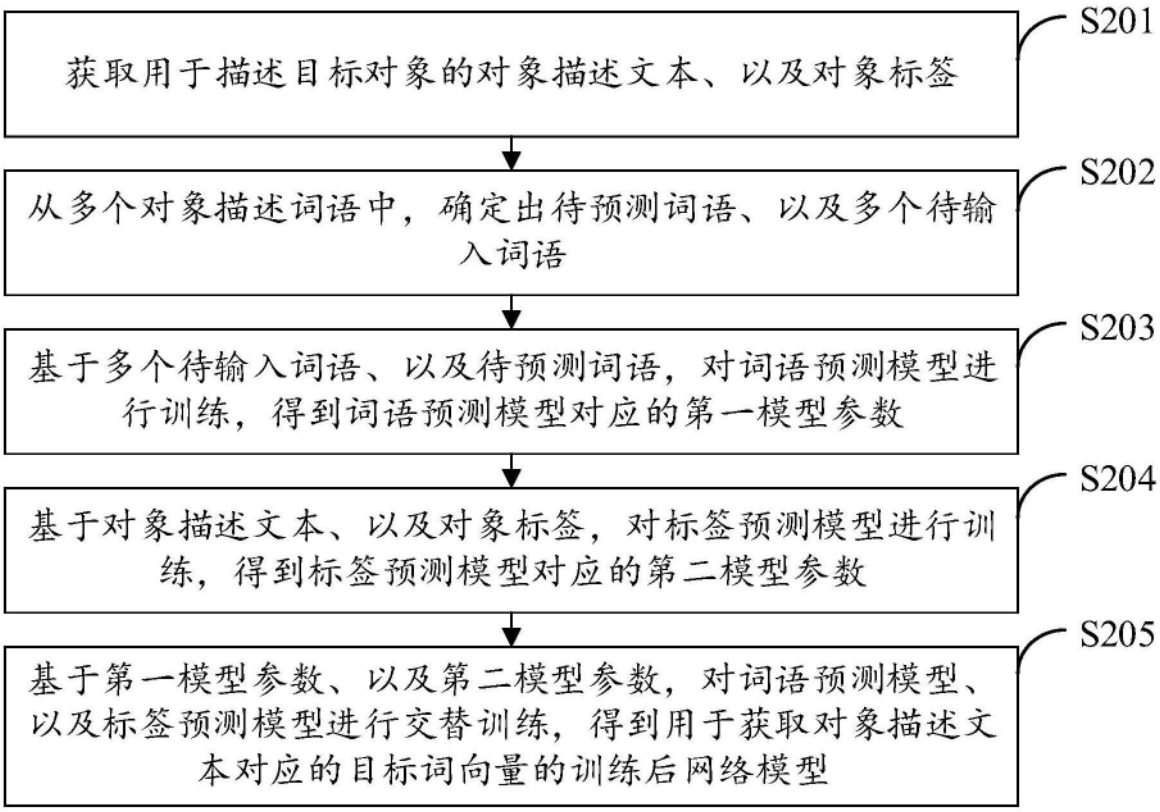

S201
获取用于描述目标对象的对象描述文本、以及对象标签
S202
从多个对象描述词语中，确定出待预测词语、以及多个待输入词语
S203
基于多个待输入词语、以及待预测词语，对词语预测模型进行训练，得到词语预测模型对应的第一模型参数
S204
基于对象描述文本、以及对象标签，对标签预测模型进行训练，得到标签预测模型对应的第二模型参数
S205
基于第一模型参数、以及第二模型参数，对词语预测模型、以及标签预测模型进行交替训练，得到用于获取对象描述文本对应的目标词向量的训练后网络模型

图2

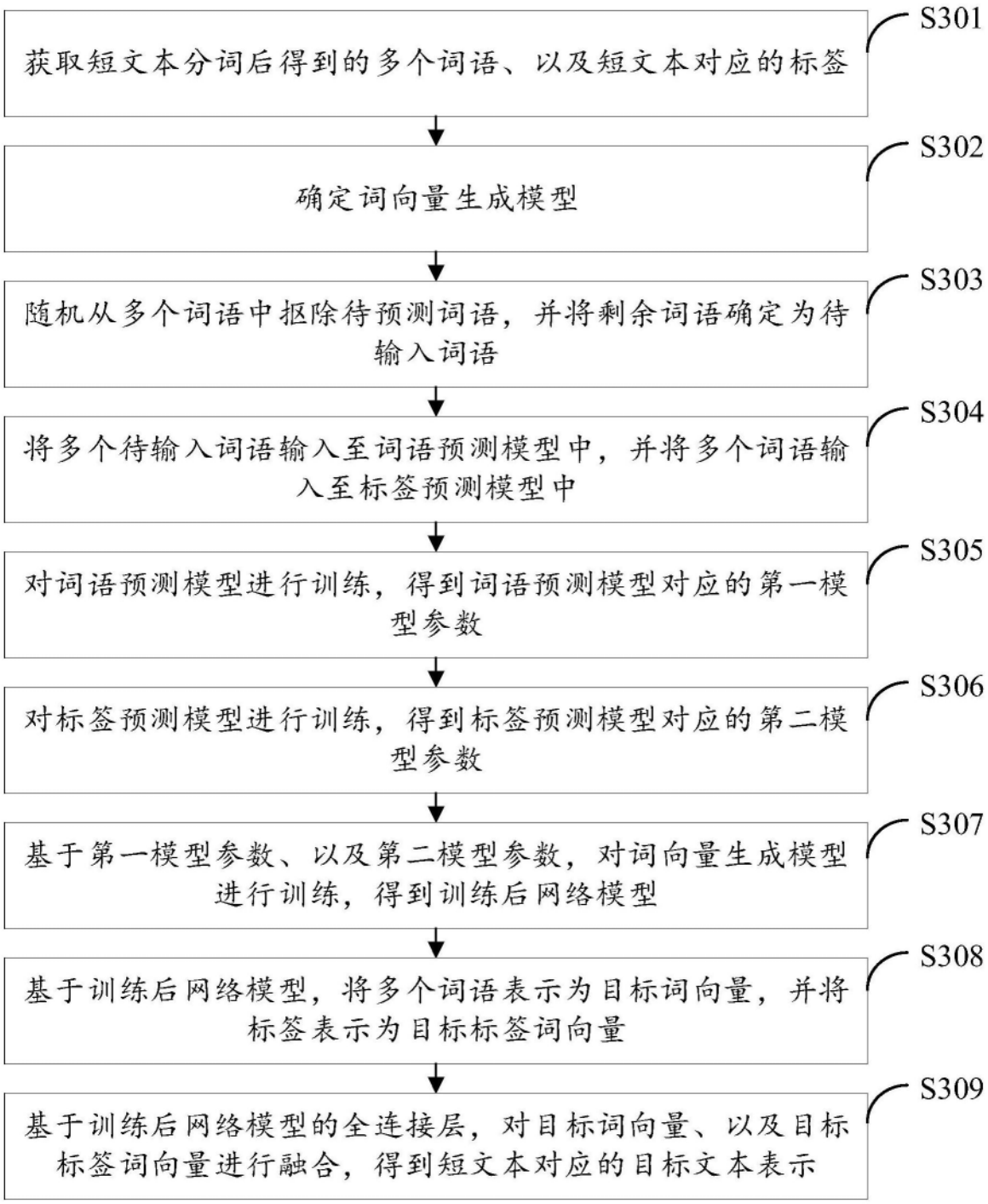

S301
获取短文本分词后得到的多个词语、以及短文本对应的标签
S302
确定词向量生成模型
S303
随机从多个词语中抠除待预测词语，并将剩余词语确定为待输入词语
S304
将多个待输入词语输入至词语预测模型中，并将多个词语输入至标签预测模型中
S305
对词语预测模型进行训练，得到词语预测模型对应的第一模型参数
S306
对标签预测模型进行训练，得到标签预测模型对应的第二模型参数
S307
基于第一模型参数、以及第二模型参数，对词向量生成模型进行训练，得到训练后网络模型
S308
基于训练后网络模型，将多个词语表示为目标词向量，并将标签表示为目标标签词向量
S309
基于训练后网络模型的全连接层，对目标词向量、以及目标标签词向量进行融合，得到短文本对应的目标文本表示

图3

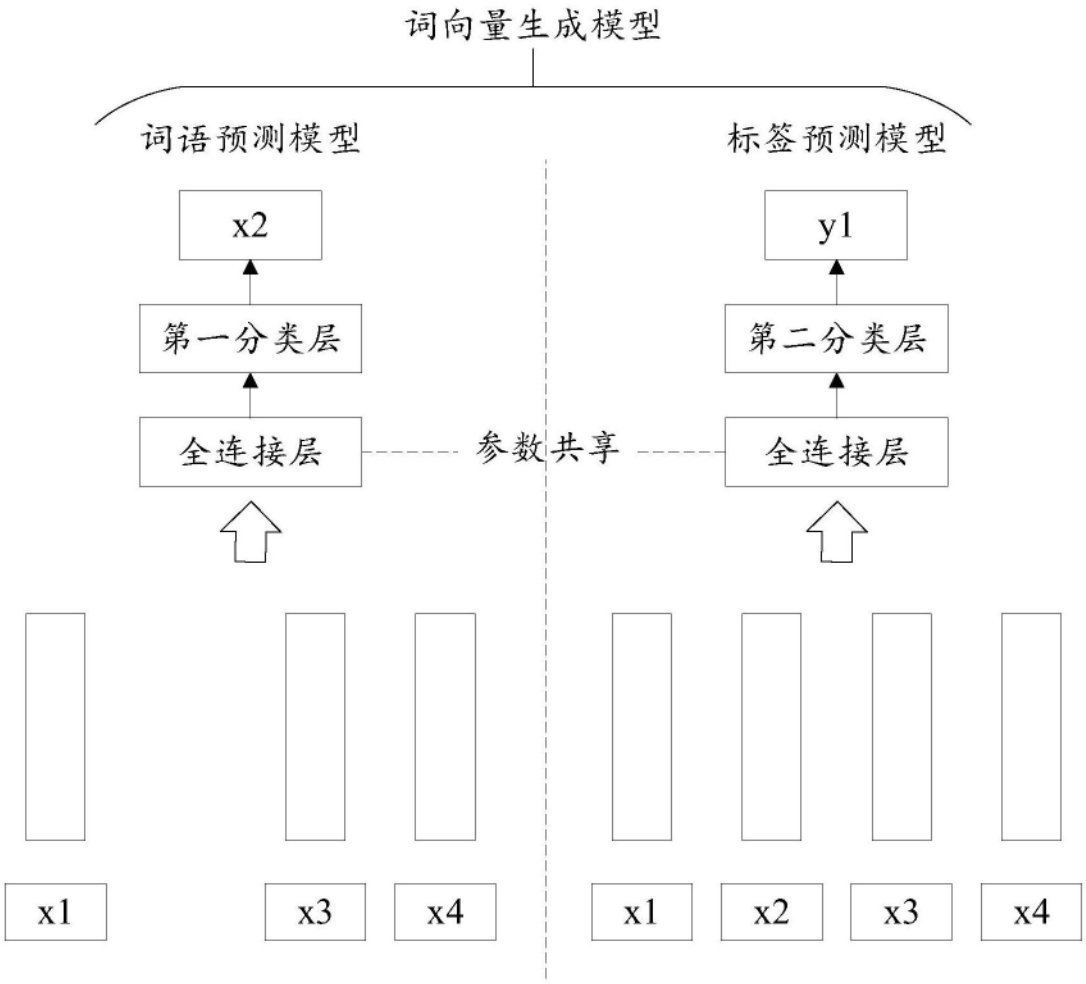

词向量生成模型
词语预测模型
标签预测模型
x2
y1
第一分类层
第二分类层
全连接层
参数共享
全连接层
x1
x3
x4
x1
x2
x3
x4

图4

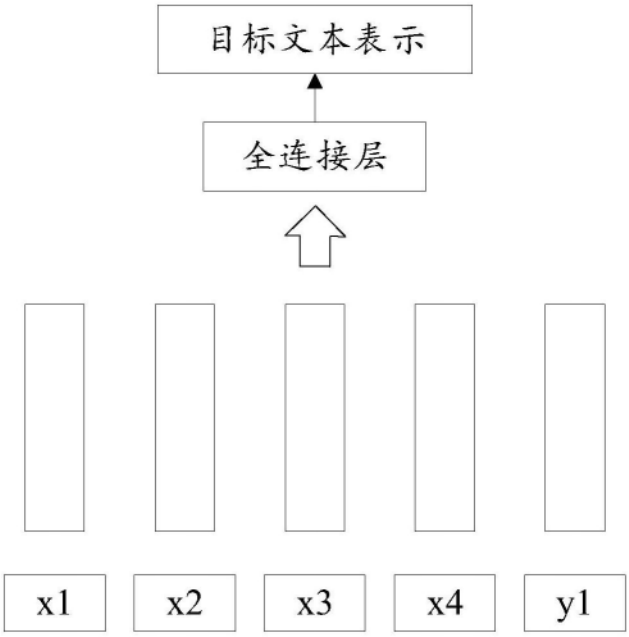

目标文本表示
全连接层
x1
x2
x3
x4
y1

图5

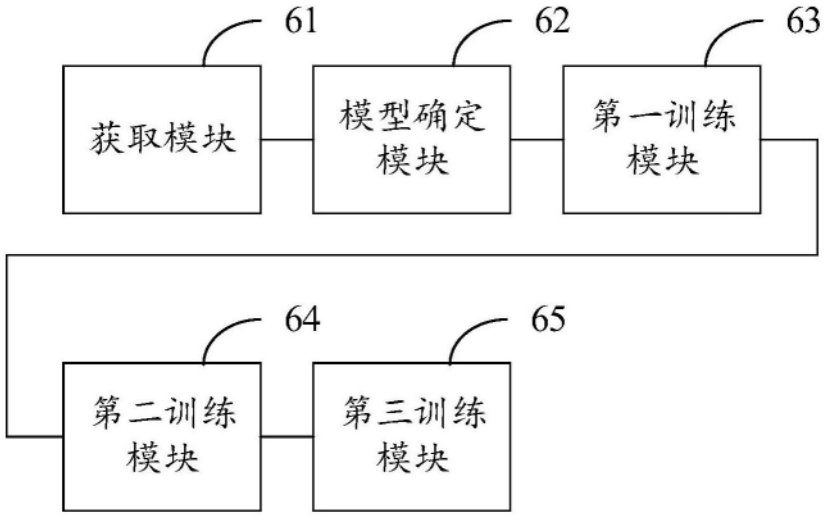

61
62
63
获取模块
模型确定模块
第一训练模块
64
65
第二训练模块
第三训练模块

图6

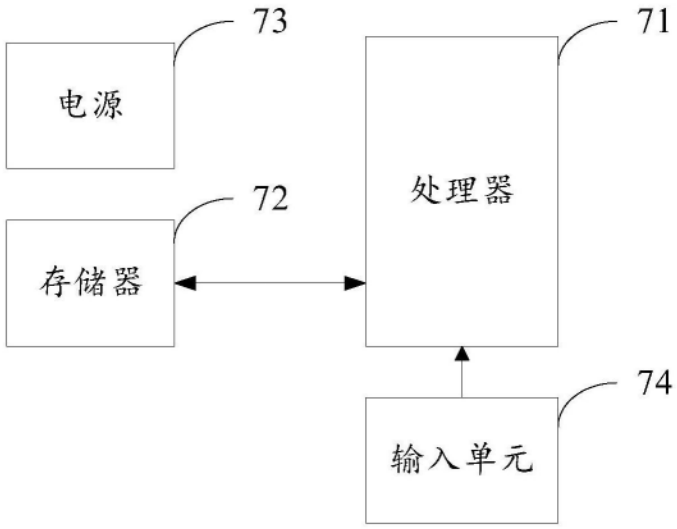


图7