



(12) 发明专利

(10) 授权公告号 CN 113609863 B

(45) 授权公告日 2024. 05. 07

(21) 申请号 202110155510.6

G06N 3/0455 (2023.01)

(22) 申请日 2021.02.04

(56) 对比文件

(65) 同一申请的已公布的文献号

申请公布号 CN 113609863 A

CN 110532554 A, 2019.12.03

CN 110795945 A, 2020.02.14

CN 110852110 A, 2020.02.28

(43) 申请公布日 2021.11.05

CN 111666766 A, 2020.09.15

(73) 专利权人 腾讯科技(深圳)有限公司

CN 111859919 A, 2020.10.30

地址 518044 广东省深圳市南山区高新区

CN 111859997 A, 2020.10.30

科技中一路腾讯大厦35层

CN 111933129 A, 2020.11.13

(72) 发明人 王龙跃 刘思佑 丁亮

US 2016027333 A1, 2016.01.28

US 2020081979 A1, 2020.03.12

(74) 专利代理机构 北京同达信恒知识产权代理

有限公司 11291

审查员 方萍

专利代理师 朱佳

(51) Int. Cl.

G06F 40/30 (2020.01)

G06F 40/289 (2020.01)

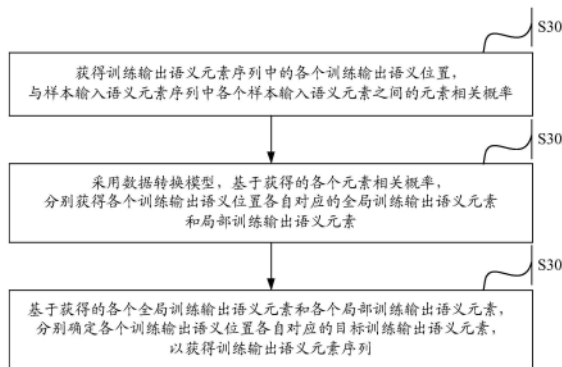
权利要求书4页 说明书27页 附图10页

(54) 发明名称

一种训练、使用数据转换模型的方法、装置及计算机设备

(57) 摘要

本申请提供一种训练、使用数据转换模型的方法、装置及计算机设备,可以应用于人工智能和云安全等领域,用于解决注意力机制的准确性较低的问题。训练数据转换模型的方法至少包括:获得训练输出语义元素序列中的各个训练输出语义位置,与所述样本输入语义元素序列中各个样本输入语义元素之间的元素相关概率;采用所述数据转换模型,基于获得的各个元素相关概率,分别获得所述各个训练输出语义位置各自对应的全局训练输出语义元素和局部训练输出语义元素,并基于获得的各个全局训练输出语义元素和各个局部训练输出语义元素,分别确定所述各个训练输出语义位置各自对应的目标训练输出语义元素,以获得所述训练输出语义元素序列。



1. 一种训练数据转换模型的方法,其特征在于,包括:

采用样本输入语义元素序列集合对数据转换模型进行训练,获得已训练的数据转换模型;其中,在一次训练过程中,针对所述样本输入语义元素序列集合中的样本输入语义元素序列,至少执行以下操作:

获得训练输出语义元素序列中的各个训练输出语义位置,与所述样本输入语义元素序列中各个样本输入语义元素之间的元素相关概率;

采用所述数据转换模型,针对所述各个训练输出语义位置,分别执行以下操作:

基于获得的各个元素相关概率,分别获得所述样本输入语义元素序列中的各个样本输入语义元素,对应于一个训练输出语义位置的第一元素权重;

基于所述各个样本输入语义元素对应的第一元素权重,对所述各个样本输入语义元素进行加权求和处理,获得所述一个训练输出语义位置对应的全局训练输出语义元素;

基于获得的各个元素相关概率,在所述样本输入语义元素序列中,确定所述一个训练输出语义位置对应的目标相关样本输入语义元素,并得到获得的各个目标相关样本输入语义元素,对应于所述一个训练输出语义位置的第二元素权重;

基于所述各个目标相关样本输入语义元素对应的第二元素权重,对所述各个目标相关样本输入语义元素进行加权求和处理,获得所述一个训练输出语义位置对应的局部训练输出语义元素;

基于获得的各个全局训练输出语义元素和各个局部训练输出语义元素,分别确定所述各个训练输出语义位置各自对应的目标训练输出语义元素,以获得所述训练输出语义元素序列,其中,所述全局训练输出语义元素与每个样本输入语义元素相关,所述局部训练输出语义元素与部分样本输入语义元素相关。

2. 根据权利要求1所述的方法,其特征在于,获得训练输出语义元素序列中的各个训练输出语义位置,与样本输入语义元素序列中各个样本输入语义元素之间的元素相关概率,包括:

基于编码模型和解码模型学习的目标子数据的排列规则,分别确定所述各个训练输出语义位置各自对应的目标子数据在目标数据中的关联位置,与初始数据中所述各个样本输入语义位置各自对应的各个初始子数据之间的数据相关概率,其中,所述样本输入语义元素序列是采用所述编码模型对所述初始数据进行编码处理获得的,所述目标数据是采用与所述编码模型对应的所述解码模型对所述训练输出语义元素序列进行解码处理获得的;

基于获得的各个数据相关概率,确定所述训练输出语义元素序列中的各个训练输出语义位置,与所述样本输入语义元素序列中各个样本输入语义元素之间的元素相关概率。

3. 根据权利要求1所述的方法,其特征在于,基于获得的各个元素相关概率,确定所述各个训练输出语义位置中的一个训练输出语义位置对应的目标相关样本输入语义元素,包括:

基于获得的各个元素相关概率,确定所述一个训练输出语义位置对应的最大元素相关概率,获得所述样本输入语义元素序列中,所述一个训练输出语义位置对应的最相关样本输入语义元素;或者,

在所述样本输入语义元素序列中,将与所述最相关样本输入语义元素满足预设相关条件的至少一个样本输入语义元素,确定为所述一个训练输出语义位置对应的目标相关样本

输入语义元素。

4. 根据权利要求3所述的方法,其特征在于,在所述样本输入语义元素序列中,将与所述最相关样本输入语义元素满足预设相关条件的至少一个样本输入语义元素,确定为所述一个训练输出语义位置对应的目标相关样本输入语义元素,包括:

在所述样本输入语义元素序列中,将样本输入语义位置与所述最相关样本输入语义元素的样本输入语义位置之间的距离在预设的距离范围内的至少一个样本输入语义元素,确定为所述一个训练输出语义位置对应的目标相关样本输入语义元素;或者,

在所述样本输入语义元素序列中,将元素相关概率与所述最大元素相关概率之间的误差在预设的误差范围内的至少一个样本输入语义元素,确定为所述一个训练输出语义位置对应的目标相关样本输入语义元素。

5. 根据权利要求1~4任一项所述的方法,其特征在于,基于获得的各个全局训练输出语义元素和各个局部训练输出语义元素,分别确定所述各个训练输出语义位置各自对应的目标训练输出语义元素,包括:

针对所述各个训练输出语义位置,分别执行以下操作:

获得所述各个训练输出语义位置中的一个训练输出语义位置对应的全局训练输出语义元素的全局权重和局部训练输出语义元素的局部权重;

基于获得的全局权重和局部权重,对所述一个训练输出语义位置对应的全局训练输出语义元素和局部训练输出语义元素进行加权求和处理,获得所述一个训练输出语义位置对应的目标训练输出语义元素。

6. 根据权利要求5所述的方法,其特征在于,获得所述各个训练输出语义位置中的一个训练输出语义位置对应的全局训练输出语义元素的全局权重和局部训练输出语义元素的局部权重,包括:

对所述一个训练输出语义位置对应的位置标识进行线性变换处理,获得所述一个训练输出语义位置对应的线性变换后的所述位置标识;

将经过线性变换后的所述位置标识,确定为所述一个训练输出语义位置对应的全局训练输出语义元素的全局权重;

将与获得的全局权重满足预设权重关系的权重系数,确定为所述一个训练输出语义位置对应的局部训练输出语义元素的局部权重。

7. 根据权利要求2所述的方法,其特征在于,在获得所述训练输出语义元素序列之后,还包括:

基于所述训练输出语义元素序列对应的目标数据与所述样本输入语义元素序列对应的初始数据之间的相似度,确定所述编码模型和所述解码模型对应的编解码训练损失;

将所述编解码训练损失确定为所述数据转换模型的训练损失,并基于所述训练损失对所述数据转换模型的模型参数进行调整。

8. 一种使用数据转换模型的方法,其特征在于,包括:

获得待处理输入语义元素序列,以及转换输出语义元素序列中的各个转换输出语义位置,与所述待处理输入语义元素序列中各个待处理输入语义元素之间的元素相关概率;

采用所述数据转换模型,针对所述各个转换输出语义位置,分别执行以下操作:

基于获得的各个元素相关概率,分别获得所述待处理输入语义元素序列中的各个待处

理输入语义元素,对应于一个转换输出语义位置的第一元素权重;

基于所述各个待处理输入语义元素对应的第一元素权重,对所述各个待处理输入语义元素进行加权求和处理,获得所述一个转换输出语义位置对应的全局转换输出语义元素;

基于获得的各个元素相关概率,在所述待处理输入语义元素序列中,确定所述一个转换输出语义位置对应的目标相关待处理输入语义元素,并得到获得的各个目标相关待处理输入语义元素,对应于所述一个转换输出语义位置的第二元素权重;

基于所述各个目标相关待处理输入语义元素对应的第二元素权重,对所述各个目标相关待处理输入语义元素进行加权求和处理,获得所述一个转换输出语义位置对应的局部转换输出语义元素;

基于获得的各个全局转换输出语义元素和各个局部转换输出语义元素,分别确定所述各个转换输出语义位置各自对应的目标转换输出语义元素,以获得所述转换输出语义元素序列,其中,所述全局转换输出语义元素与每个待处理输入语义元素相关,所述局部转换输出语义元素与部分待处理输入语义元素相关。

9.根据权利要求8所述的方法,其特征在于,在获得所述转换输出语义元素序列之后,还包括:

基于获得的各个元素相关概率,获得所述转换输出语义元素序列对应的局部熵;

基于获得的局部熵与预设的局部熵阈值的比对结果,获得所述数据转换模型的准确性分析结果,其中,所述准确性分析结果用于表征在确定各个转换输出语义位置各自对应的目标转换输出语义元素时,所关注的待处理输入语义元素是否准确。

10.一种训练数据转换模型的装置,其特征在于,包括:

训练模块:用于采用样本输入语义元素序列集合对数据转换模型进行训练,获得已训练的数据转换模型;其中,在一次训练过程中,针对所述样本输入语义元素序列集合中的样本输入语义元素序列,至少执行以下操作:

获取模块:用于获得训练输出语义元素序列中的各个训练输出语义位置,与所述样本输入语义元素序列中各个样本输入语义元素之间的元素相关概率;

所述训练模块还用于:采用所述数据转换模型,针对所述各个训练输出语义位置,分别执行以下操作:

基于获得的各个元素相关概率,分别获得所述样本输入语义元素序列中的各个样本输入语义元素,对应于一个训练输出语义位置的第一元素权重;

基于所述各个样本输入语义元素对应的第一元素权重,对所述各个样本输入语义元素进行加权求和处理,获得所述一个训练输出语义位置对应的全局训练输出语义元素;

基于获得的各个元素相关概率,在所述样本输入语义元素序列中,确定所述一个训练输出语义位置对应的目标相关样本输入语义元素,并得到获得的各个目标相关样本输入语义元素,对应于所述一个训练输出语义位置的第二元素权重;

基于所述各个目标相关样本输入语义元素对应的第二元素权重,对所述各个目标相关样本输入语义元素进行加权求和处理,获得所述一个训练输出语义位置对应的局部训练输出语义元素;

基于获得的各个全局训练输出语义元素和各个局部训练输出语义元素,分别确定所述各个训练输出语义位置各自对应的目标训练输出语义元素,以获得所述训练输出语义元素

序列,其中,所述全局训练输出语义元素与每个样本输入语义元素相关,所述局部训练输出语义元素与部分样本输入语义元素相关。

11. 一种使用数据转换模型的装置,其特征在于,包括:

获取模块:用于获得待处理输入语义元素序列,以及转换输出语义元素序列中的各个转换输出语义位置,与所述待处理输入语义元素序列中各个待处理输入语义元素之间的元素相关概率;

转换模块:用于采用所述数据转换模型,针对所述各个转换输出语义位置,分别执行以下操作:

基于获得的各个元素相关概率,分别获得所述待处理输入语义元素序列中的各个待处理输入语义元素,对应于一个转换输出语义位置的第一元素权重;

基于所述各个待处理输入语义元素对应的第一元素权重,对所述各个待处理输入语义元素进行加权求和处理,获得所述一个转换输出语义位置对应的全局转换输出语义元素;

基于获得的各个元素相关概率,在所述待处理输入语义元素序列中,确定所述一个转换输出语义位置对应的目标相关待处理输入语义元素,并得到获得的各个目标相关待处理输入语义元素,对应于所述一个转换输出语义位置的第二元素权重;

基于所述各个目标相关待处理输入语义元素对应的第二元素权重,对所述各个目标相关待处理输入语义元素进行加权求和处理,获得所述一个转换输出语义位置对应的局部转换输出语义元素;

基于获得的各个全局转换输出语义元素和各个局部转换输出语义元素,分别确定所述各个转换输出语义位置各自对应的目标转换输出语义元素,以获得所述转换输出语义元素序列,其中,所述全局转换输出语义元素与每个待处理输入语义元素相关,所述局部转换输出语义元素与部分待处理输入语义元素相关。

12. 一种计算机设备,其特征在于,包括:

存储器,用于存储程序指令;

处理器,用于调用所述存储器中存储的程序指令,按照获得的程序指令执行如权利要求1~9中任一项所述的方法。

13. 一种存储介质,其特征在于,所述存储介质存储有计算机可执行指令,所述计算机可执行指令用于使计算机执行如权利要求1~9中任一项所述的方法。

一种训练、使用数据转换模型的方法、装置及计算机设备

技术领域

[0001] 本申请涉及计算机技术领域,尤其涉及一种训练、使用数据转换模型的方法、装置及计算机设备。

背景技术

[0002] 随着科技的不断发展,在很多领域中,机器可以基于已训练的模型对初始数据进行一些智能化的处理,并基于模型的输出获得目标数据,从而机器可以像人一样执行一些智能化行为。例如,自然语言处理领域中,机器可以基于已训练的模型将第一语言形式的初始数据,翻译成第二语言形式的目标数据。又例如,计算机视觉领域中,机器可以基于已训练的模型将图像形式的初始数据,翻译成文字形式的目标数据等。

[0003] 已训练的模型中的数据转换模型通常是将注意力集中在初始数据的全局特征上,然而,初始数据的全局特征仅涉及初始数据宏观角度的特征,使得基于已训练的模型获得的目标数据也仅能够表征初始数据的宏观角度的特征,可见,数据转换模型的注意力机制的准确性较低。

发明内容

[0004] 本申请实施例提供一种训练、使用数据转换模型的方法、装置及计算机设备,用于解决注意力机制的准确性较低的问题。

[0005] 第一方面,提供一种训练数据转换模型的方法,包括:

[0006] 采用样本输入语义元素序列集合对数据转换模型进行训练,获得已训练的数据转换模型;其中,在一次训练过程中,针对所述样本输入语义元素序列集合中的样本输入语义元素序列,至少执行以下操作:

[0007] 获得训练输出语义元素序列中的各个训练输出语义位置,与所述样本输入语义元素序列中各个样本输入语义元素之间的元素相关概率;

[0008] 采用所述数据转换模型,基于获得的各个元素相关概率,分别获得所述各个训练输出语义位置各自对应的全局训练输出语义元素和局部训练输出语义元素,并基于获得的各个全局训练输出语义元素和各个局部训练输出语义元素,分别确定所述各个训练输出语义位置各自对应的目标训练输出语义元素,以获得所述训练输出语义元素序列,其中,所述全局训练输出语义元素与每个样本输入语义元素相关,所述局部训练输出语义元素与部分样本输入语义元素相关。

[0009] 第二方面,提供一种使用数据转换模型的方法,包括:

[0010] 获得待处理输入语义元素序列,以及转换输出语义元素序列中的各个转换输出语义位置,与所述待处理输入语义元素序列中各个待处理输入语义元素之间的元素相关概率;

[0011] 采用所述数据转换模型,基于获得的各个元素相关概率,分别获得所述各个转换输出语义位置各自对应的全局转换输出语义元素和局部转换输出语义元素,并基于获得的

各个全局转换输出语义元素和各个局部转换输出语义元素,分别确定所述各个转换输出语义位置各自对应的目标转换输出语义元素,以获得所述转换输出语义元素序列,其中,所述全局转换输出语义元素与每个待处理输入语义元素相关,所述局部转换输出语义元素与部分待处理输入语义元素相关。

[0012] 第三方面,提供一种训练数据转换模型的装置,包括:

[0013] 训练模块:用于采用样本输入语义元素序列集合对数据转换模型进行训练,获得已训练的数据转换模型;其中,在一次训练过程中,针对所述样本输入语义元素序列集合中的样本输入语义元素序列,至少执行以下操作:

[0014] 获取模块:用于获得训练输出语义元素序列中的各个训练输出语义位置,与所述样本输入语义元素序列中各个样本输入语义元素之间的元素相关概率;

[0015] 所述训练模块还用于:采用所述数据转换模型,基于获得的各个元素相关概率,分别获得所述各个训练输出语义位置各自对应的全局训练输出语义元素和局部训练输出语义元素,并基于获得的各个全局训练输出语义元素和各个局部训练输出语义元素,分别确定所述各个训练输出语义位置各自对应的目标训练输出语义元素,以获得所述训练输出语义元素序列,其中,所述全局训练输出语义元素与每个样本输入语义元素相关,所述局部训练输出语义元素与部分样本输入语义元素相关。

[0016] 第四方面,提供一种使用数据转换模型的装置,包括:

[0017] 获取模块:用于获得待处理输入语义元素序列,以及转换输出语义元素序列中的各个转换输出语义位置,与所述待处理输入语义元素序列中各个待处理输入语义元素之间的元素相关概率;

[0018] 转换模块:用于采用所述数据转换模型,基于获得的各个元素相关概率,分别获得所述各个转换输出语义位置各自对应的全局转换输出语义元素和局部转换输出语义元素,并基于获得的各个全局转换输出语义元素和各个局部转换输出语义元素,分别确定所述各个转换输出语义位置各自对应的目标转换输出语义元素,以获得所述转换输出语义元素序列,其中,所述全局转换输出语义元素与每个待处理输入语义元素相关,所述局部转换输出语义元素与部分待处理输入语义元素相关。

[0019] 第五方面,提供一种计算机设备,包括:

[0020] 存储器,用于存储程序指令;

[0021] 处理器,用于调用所述存储器中存储的程序指令,按照获得的程序指令执行如第一方面或第二方面所述的方法。

[0022] 第六方面,提供一种存储介质,所述存储介质存储有计算机可执行指令,所述计算机可执行指令用于使计算机执行如第一方面或第二方面所述的方法。

[0023] 本申请实施例中,基于训练输出语义元素序列中的各个训练输出语义位置,与样本输入语义元素序列中各个样本输入语义元素之间的元素相关概率,分别获得各个训练输出语义位置各自对应的全局训练输出语义元素和局部训练输出语义元素。不仅将注意力集中在样本输入语义元素序列的全局特征上,还将注意力集中在训练输出语义位置对应的部分样本输入语义元素的局部特征上,从多个角度出发,对样本输入语义元素序列进行描述,数据转换模型可以更加准确地对样本输入语义元素序列进行数据转换处理。

[0024] 并且,采用数据转换模型,基于各个全局训练输出语义元素和各个局部训练输出

语义元素,获得各个训练输出语义位置各自对应的目标训练输出语义元素,以获得训练输出语义元素序列。将对样本输入语义元素序列多维度的描述一同作为获得训练输出语义元素序列的依据,使得获得的训练输出语义元素序列能够更加准确地表征样本输入语义元素序列的语义,提高了数据转换模型的准确性。

[0025] 本申请实施例中,基于转换输出语义元素序列中的各个转换输出语义位置,与待处理输入语义元素序列中各个待处理输入语义元素之间的元素相关概率,分别获得各个转换输出语义位置各自对应的全局转换输出语义元素和局部转换输出语义元素。不仅将注意力集中在待处理输入语义元素序列的全局特征上,还将注意力集中在转换输出语义位置对应的部分待处理输入语义元素的局部特征上,从多个角度出发,对待处理输入语义元素序列进行描述,数据转换模型可以更加准确地对待处理输入语义元素序列进行数据转换处理。

[0026] 并且,采用数据转换模型,基于各个全局转换输出语义元素和各个局部转换输出语义元素,获得各个转换输出语义位置各自对应的目标转换输出语义元素,以获得转换输出语义元素序列。将对待处理输入语义元素序列多维度的描述一同作为获得转换输出语义元素序列的依据,使得获得的转换输出语义元素序列能够更加准确地表征待处理输入语义元素序列的语义,提高了数据转换模型的准确性。

附图说明

- [0027] 图1为本申请实施例提供的训练、使用数据转换模型的方法的一种应用场景;
- [0028] 图2a为本申请实施例提供的训练、使用数据转换模型的方法的一种原理示意图一;
- [0029] 图2b为本申请实施例提供的训练、使用数据转换模型的方法的一种流程示意图一;
- [0030] 图3为本申请实施例提供的训练、使用数据转换模型的方法的一种流程示意图二;
- [0031] 图4为本申请实施例提供的训练、使用数据转换模型的方法的一种流程示意图三;
- [0032] 图5a为本申请实施例提供的训练、使用数据转换模型的方法的一种原理示意图二;
- [0033] 图5b为本申请实施例提供的训练、使用数据转换模型的方法的一种流程示意图四;
- [0034] 图6为本申请实施例提供的训练、使用数据转换模型的方法的一种流程示意图五;
- [0035] 图7a为本申请实施例提供的训练、使用数据转换模型的方法的一种结构示意图一;
- [0036] 图7b为本申请实施例提供的训练、使用数据转换模型的方法的一种结构示意图二;
- [0037] 图7c为本申请实施例提供的训练、使用数据转换模型的方法的一种结构示意图三;
- [0038] 图8为本申请实施例提供的训练、使用数据转换模型的方法的一种折线示意图一;
- [0039] 图9为本申请实施例提供的训练、使用数据转换模型的方法的一种折线示意图二;
- [0040] 图10为本申请实施例提供的训练、使用数据转换模型的装置的结构示意图一;

[0041] 图11为本申请实施例提供的训练、使用数据转换模型的装置的结构示意图二；

[0042] 图12为本申请实施例提供的训练、使用数据转换模型的装置的结构示意图二。

具体实施方式

[0043] 为了使本申请实施例的目的、技术方案和优点更加清楚,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。

[0044] 下面对本申请实施例中的部分用语进行解释说明,以便于本领域技术人员理解。

[0045] (1) 深度学习(Deep Learning,DL)和神经网络(Neural Network,NN):

[0046] 深度学习是机器学习的分支,深度学习是一种试图使用包含复杂结构的多个处理层,或由多重非线性变换构成的多个处理层,对数据进行高层抽象的方法。

[0047] 神经网络是在机器学习和认知科学领域的一种模仿生物神经网络结构和功能的深度学习模型。

[0048] (2) 机器翻译(Machine Translation,MT)和神经网络机器翻译(Neural Machine Translation,NMT):

[0049] 机器翻译是利用电子计算机等装置自动地把一种语言文字译成另一种语言文字的方法。

[0050] 神经网络机器翻译是目前最新一代基于神经网络的机器翻译技术。

[0051] (3) 交叉注意力机制(Cross-Attention Mechanism或Encoder-decoder Attention):

[0052] 交叉注意力机制是一种建立神经网络中编码模型与解码模型的隐藏状态之间的依赖关系的方法,也就是说,交叉注意力机制是从解码模型出发关注编码模型的神经网络结构。

[0053] (4) 自关注神经网络(Self-Attention Network,SAN)和Transformer:

[0054] 自关注神经网络是一种基于自关注注意力机制的神经网络结构模型。

[0055] Transformer是基于SAN的编码模型-解码模型框架。Transformer是目前最主流的序列到序列生成(sequence-to-sequence generation)的模型结构。

[0056] (5) 机器翻译评价指标(Bilingual Evaluation Understudy,BLEU):

[0057] 机器翻译评价指标是用于评价及其翻译模型的,机器翻译评价指标的值越高表示翻译模型的翻译效果越好。

[0058] 本申请实施例涉及云技术(Cloud technology)和人工智能(Artificial Intelligence,AI)。基于云技术中的云计算(cloud computing)和云存储(cloud storage)而设计。基于人工智能中的计算机视觉技术(Computer Vision,CV)、语音技术(Speech Technology)、自然语言处理(natural language processing,NLP)技术和机器学习(Machine Learning,ML)而设计。

[0059] 云技术是指在广域网或局域网内将硬件、软件、网络等系列资源统一起来,实现数据的计算、储存、处理和共享的一种托管技术。云技术基于云计算商业模式应用的网络技术、信息技术、整合技术、管理平台技术、应用技术等的总称,可以组成资源池,按需所用,灵活便利。云计算技术将变成重要支撑。技术网络系统的后台服务需要大量的计算、存储资源,如视频网站、图片类网站和更多的门户网站。伴随着互联网行业的高度发展和应用,将

来每个物品都有可能存在自己的识别标志,都需要传输到后台系统进行逻辑处理,不同程度级别的数据将会分开处理,各类行业数据皆需要强大的系统后盾支撑,只能通过云计算来实现。

[0060] 云计算是一种计算模式,它将计算任务分布在大量计算机构成的资源池上,使各种应用系统能够根据需要获取计算力、存储空间和信息服务。提供资源的网络被称为“云”。“云”中的资源在使用者看来是可以无限扩展的,并且可以随时获取,按需使用,随时扩展,按使用付费。

[0061] 作为云计算的基础能力提供商,会建立云计算资源池(简称云平台,一般称为IaaS(Infrastructure as a Service,基础设施即服务)平台,在资源池中部署多种类型的虚拟资源,供外部客户选择使用。云计算资源池中主要包括:计算设备(为虚拟化机器,包含操作系统)、存储设备、网络设备。

[0062] 按照逻辑功能划分,在IaaS(Infrastructure as a Service,基础设施即服务)层上可以部署PaaS(Platform as a Service,平台即服务)层,PaaS层之上再部署SaaS(Software as a Service,软件即服务)层,也可以直接将SaaS部署在IaaS上。PaaS为软件运行的平台,如数据库、web容器等。SaaS为各式各样的业务软件,如web门户网站、短信群发器等。一般来说,SaaS和PaaS相对于IaaS是上层。

[0063] 云存储是在云计算概念上延伸和发展出来的一个新的概念,分布式云存储系统(以下简称存储系统)是指通过集群应用、网格技术以及分布存储文件系统等功能,将网络中大量各种不同类型的存储设备(存储设备也称之为存储节点)通过应用软件或应用接口集合起来协同工作,共同对外提供数据存储和业务访问功能的一个存储系统。

[0064] 目前,存储系统的存储方法为:创建逻辑卷,在创建逻辑卷时,就为每个逻辑卷分配物理存储空间,该物理存储空间可能是某个存储设备或者某几个存储设备的磁盘组成。客户端在某一逻辑卷上存储数据,也就是将数据存储在文件系统中,文件系统将数据分成许多部分,每一部分是一个对象,对象不仅包含数据而且还包含数据标识(ID, ID entity)等额外的信息,文件系统将每个对象分别写入该逻辑卷的物理存储空间,且文件系统会记录每个对象的存储位置信息,从而当客户端请求访问数据时,文件系统能够根据每个对象的存储位置信息让客户端对数据进行访问。

[0065] 存储系统为逻辑卷分配物理存储空间的过程,具体为:按照对存储于逻辑卷的对象的容量估量(该估量往往相对于实际要存储的对象的容量有很大余量)和独立冗余磁盘阵列(RAID, Redundant Array of Independent Disk)的组别,预先将物理存储空间划分成分条,一个逻辑卷可以理解为一个分条,从而为逻辑卷分配了物理存储空间。

[0066] 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。人工智能技术主要包括计算机视觉技术、自然语言处理技术、机器学习以及深度学习等几大方向。

[0067] 随着人工智能技术研究和进步,人工智能在多个领域展开研究和应用,例如常见的智能家居、智能推荐系统、虚拟助理、智能音箱、智能营销、智能翻译、自动驾驶、机器人、

智能医疗等,相信随着技术的发展,人工智能将在更多的领域得到应用,并发挥越来越重要的价值。

[0068] 计算机视觉技术计算机视觉是一门研究如何使机器“看”的科学,更进一步的说,就是指用摄影机和电脑代替人眼对目标进行识别、跟踪和测量等机器视觉,并进一步做图形处理,使电脑处理成为更适合人眼观察或传送给仪器检测的图像。作为一个科学学科,计算机视觉研究相关的理论和技术,试图建立能够从图像或者多维数据中获取信息的人工智能系统。计算机视觉技术通常包括图像处理、图像识别、图像语义理解、图像检索、OCR、视频处理、视频语义理解、视频内容、行为识别、三维物体重建、3D技术、虚拟现实、增强现实、同步定位与地图构建等技术,还包括常见的人脸识别、指纹识别等生物特征识别技术。

[0069] 语音技术的关键技术有自动语音识别技术(ASR)和语音合成技术(TTS)以及声纹识别技术。让计算机能听、能看、能说、能感觉,是未来人机交互的发展方向,其中语音成为未来最被看好的人机交互方式之一。

[0070] 自然语言处理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。

[0071] 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,其应用遍及人工智能的各个领域。机器学习和深度学习通常包括人工神经网络、置信网络、强化学习、迁移学习、归纳学习、式教学习等技术。

[0072] 下面对本申请实施例提供的训练、使用数据转换模型的应用领域进行简单介绍。

[0073] 随着科技的不断发展,在很多领域中,机器可以基于已训练的模型对初始数据进行一些智能化的处理,并基于模型的输出获得目标数据,从而机器可以像人一样执行一些智能化行为。然而,已训练的模型中的数据转换模型通常是将注意力集中在初始数据的全局特征上,然而,初始数据的全局特征仅涉及初始数据宏观角度的特征,使得基于已训练的模型获得的目标数据也仅能够表征初始数据的宏观角度的特征,从而使得得到的目标数据的准确性较低。

[0074] 例如,在翻译领域中,机器可以基于已训练的模型将第一语言形式的初始文本或初始语音,翻译成第二语言形式的目标文本或目标语音。又例如,在视频审核领域中,机器可以基于已训练的模型将图像形式的初始视频帧,翻译成文字形式的目标文本等,下面进行示例介绍。

[0075] 在翻译领域中,例如由于目标账号使用的是中文,而需要用英文与其他账号交流时,机器在接收目标账号输入的中文文本或中文语音之后,需要将中文文本或中文语音转换为英文文本向其他账号展示,或转换为英文语音向其他账号播放等,从而目标账号不需要使用英文就可以和其他账号交流。然而,对于一个中文词语来说,可以有多种英文词语与之对应,选择不同的英文词语来翻译该中文词语,可以得到不同的语义,如果通过中文语句中的所有词语的宏观特征,选择一个英文词语来翻译该中文词语,容易出现翻译出的英文

语句与原中文语句的语义不同的情况,造成其他账号无法理解目标账号的意思,或误解目标账号的意思等问题,如中文语句,A在和一个小女孩交往,可能会翻译成英文语句,A is socializing with a girl,然而实际上应该翻译为A is dating a girl。

[0076] 在视频审核领域中,例如目标账号在视频平台上传视频时,机器需要识别视频中的视频帧所表达的内容,才可以判断视频是否可以通过审核。针对每个视频帧,需要将视频帧中的每个图像区域转换为文本,确定视频帧所表达的内容。然而,对于一个图像区域来说,可以有多种本文与之对应,选择不同的文本来表达该图像区域,可以得到不同的视频帧内容,如果通过视频帧的宏观特征,选择一个文本来表达该图像区域,容易出现得到的视频帧内容与原视频帧实际展示的内容不同的情况,造成审核错误的情况,如图像区域中包括一本课本,机器可能会用读物来表达该图像区域,读物和课本实际上体现的是完全不同的场景。

[0077] 可见,已训练的模型中的数据转换模型的注意力机制的准确性较低。

[0078] 为了解决数据转换模型的注意力机制的准确性较低的问题,本申请提供一种训练数据转换模型的方法和一种使用数据转换模型的方法。下面先对训练数据转换模型的方法进行介绍。

[0079] 训练数据转换模型的方法采用样本输入语义元素序列集合对数据转换模型进行训练,直到数据转换模型的训练损失满足预设的收敛条件为止,获得已训练的数据转换模型。其中,在一次训练过程中,针对样本输入语义元素序列集合中的样本输入语义元素序列,执行以下操作:

[0080] 获得训练输出语义元素序列中的各个训练输出语义位置,与样本输入语义元素序列中各个样本输入语义元素之间的元素相关概率。在获得各个元素相关概率之后,采用数据转换模型,基于获得的各个元素相关概率,分别获得各个训练输出语义位置各自对应的全局训练输出语义元素和局部训练输出语义元素。在获得各个全局训练输出语义元素和各个局部训练输出语义元素之后,基于获得的各个全局训练输出语义元素和各个局部训练输出语义元素,分别确定各个训练输出语义位置各自对应的目标训练输出语义元素,以获得训练输出语义元素序列。在获得训练输出语义元素序列之后,基于样本输入语义元素序列和训练输出语义元素序列确定训练损失。在获得训练损失之后,基于训练损失对数据转换模型的模型参数进行调整。

[0081] 作为一种实施例,每个训练输出语义位置对应的全局训练输出语义元素用于表征样本输入语义元素序列中全部样本输入语义元素针对相应的输出语义位置的语义特征,每个训练输出语义位置对应的局部训练输出语义元素用于表征样本输入语义元素序列中部分样本输入语义元素针对相应的训练输出语义位置的语义特征。

[0082] 本申请实施例中,采用数据转换模型,基于各个全局训练输出语义元素和各个局部训练输出语义元素,获得各个训练输出语义位置各自对应的目标训练输出语义元素,以获得训练输出语义元素序列。将对样本输入语义元素序列多维度的描述一同作为获得训练输出语义元素序列的依据,使得获得的训练输出语义元素序列能够更加准确地表征样本输入语义元素序列的语义,提高了数据转换模型的准确性。

[0083] 下面对本申请提供的训练、使用数据转换模型的方法的应用场景进行说明。

[0084] 请参考图1,为本申请提供的训练数据转换模型的方法的一种应用场景。该应用场

景中包括客户端101和服务端102。客户端101包括第一客户端1011和第二客户端1012。客户端101和服务端102之间可以通信,例如,第一客户端1011和服务端102之间可以通信,第二客户端1012和服务端102之间可以通信。第一客户端1011和第二客户端1012之间可以通信。通信方式可以是采用有线通信技术进行通信,例如通过连接网线或串口线进行通信;也可以是采用无线通信技术进行通信,例如通过蓝牙或无线保真(wireless fidelity,WIFI)等技术进行通信,具体不做限制。

[0085] 第一客户端1011泛指可以为服务端102提供样本输入语义元素序列或待处理输入语义元素序列的设备,第二客户端1012泛指可以展示基于训练输出语义元素序列或转换输出语义元素序列所产生的输出数据的设备。客户端101例如是,终端设备、终端设备可以访问的第三方应用程序或终端设备可以访问的网页等。终端设备例如是手机、平板电脑或个人计算机等。服务端102泛指可以采用数据转换模型进行数据处理的设备,例如,终端设备或服务器等。服务器例如是云服务器或本地服务器等。客户端101和服务端102均可以采用云计算,以减少本地计算资源的占用;同样也可以采用云存储,以减少本地存储资源的占用。

[0086] 作为一种实施例,第一客户端1011和服务端102可以是同一个设备;或者,第二客户端1012和服务端102可以是同一个设备;或者,第一客户端1011和第二客户端1012可以是同一个设备;或者,第一客户端1011、第二客户端1012和服务端102可以是同一个设备等,具体不做限制。本申请实施例中,以第一客户端1011、第二客户端1012和服务端102分别为不同的设备为例进行介绍。

[0087] 下面对本申请实施例提供的训练数据转换模型的方法进行介绍。

[0088] 由于数据转换模型的输入数据一般是用于表征初始数据特征的向量或者矩阵等,数据转换模型的输出数据一般是用于表征目标数据特征的向量或者矩阵等,因此数据转换模型的输入数据不能完整的代替初始数据,输出数据不能完整的代替目标数据。同时,在真实的场景中,需要获得准确的目标数据,因此,服务端102在对数据转换模型进行训练时,可以结合基于初始数据为数据转换模型提供输入数据的编码模型,和基于数据转换模型的输出数据获得目标数据的解码模型对数据转换模型进行训练;或者,服务端102可以根据数据转换模型的真实使用场景,结合其他与数据转换模型结合使用的模型对数据转换模型进行训练等,具体不做限制。从而,可以基于目标数据和初始数据来训练数据转换模型,提高了训练数据转换模型的准确性。本申请实施例中,以编码模型、数据转换模型和解码模型结合对数据转换模型进行训练的方式为例进行介绍。

[0089] 请参考图2a,为编码模型、数据转换模型和解码模型结合后的原理示意图。在对数据转换模型进行训练的过程中,针对一个输入数据的训练过程,服务端102将初始数据集合中的一初始数据输入编码模型,编码模型对该初始数据进行编码处理,获得该初始数据对应的样本输入语义元素序列。服务端102将样本输入语义元素序列输入数据转换模型,数据转换模型对样本输入语义元素序列进行数据转换处理,获得训练输出语义元素序列。服务端102将训练输出语义元素序列输入解码模型,解码模型对训练输出语义元素序列进行解码处理,获得训练输出语义元素序列对应的目标数据。

[0090] 服务端102中的训练模型基于初始数据和目标数据确定编解码训练损失,编解码训练损失可以作为数据转换模型的训练损失。服务端102基于编解码训练损失对编码模型

的模型参数、数据转换模型的模型参数和解码模型的模型参数进行调整。对编码模型的模型参数、数据转换模型的模型参数和解码模型的模型参数的一次调整过程,相当于编码模型、数据转换模型和解码模型基于初始数据的一次学习过程。通过基于初始数据集中的各个初始数据进行不断地学习,直到编解码训练损失满足预设的收敛条件,编解码训练损失满足预设的收敛条件可以表示数据转换模型的训练损失满足预设的收敛条件。在数据转换模型的训练损失满足预设的收敛条件时,基于数据转换模型当前的模型参数,可以获得已训练的数据转换模型。

[0091] 下面基于图2a的原理示意图,对编码模型、数据转换模型和解码模型结合的过程进行示例介绍。请参考图2b,为编码模型、数据转换模型和解码模型结合的流程示意图。

[0092] S201,基于编码模型,获得样本输入语义元素序列。

[0093] 服务端102获得初始数据集合,分别对初始数据集合中的各个初始数据进行处理。服务端102获得初始数据集合的方式有多种,例如可以接收第一客户端1011发送的初始数据集合;或者,可以在关联的存储设备中获取初始数据集合;或者,可以在指定链接中下载初始数据集合等,具体不做限制。

[0094] 针对初始数据中的一初始数据,服务端102将初始数据输入编码模型,编码模型分别对初始数据中的各个初始子数据进行特征提取,获得每个初始子数据对应的样本输入语义元素,每个样本输入语义元素可以表征对应的初始子数据的语义。服务端102在获得各个样本输入语义元素之后,可以确定各个样本输入语义元素在样本输入语义元素序列中的排列顺序。服务端102在确定各个样本输入语义元素的排列顺序之后,可以获得样本输入语义元素序列。

[0095] 服务端102对各个样本输入语义元素进行排序的方式有多种,下面以其中的两种为例进行介绍。

[0096] 排序方法一:

[0097] 根据初始子数据在初始数据中的排列顺序,排列各个样本输入语义元素。

[0098] 由于一个初始子数据对应一个样本输入语义元素,因此,样本输入语义元素的排列顺序可以按照对应的初始子数据在初始数据中的位置进行排列。

[0099] 例如,初始数据为中文语句,A在跟一个女孩交往。各个初始子数据可以包括“A”、“在”、“跟”、“一个”、“女孩”和“交往”这六个初始子数据;或者,可以包括“A”、“在跟”、“一个女孩”和“交往”这四个初始子数据等,初始子数据的划分方式可以根据实际使用情况对编码模型进行训练,具体不做限制。以初始子数据包括“A”、“在”、“跟”、“一个”、“女孩”和“交往”为例,那么样本输入语义元素序列中的各个样本输入语义元素可以按照“A”、“在”、“跟”、“一个”、“女孩”和“交往”的顺序排列。

[0100] 排序方法二:

[0101] 根据编码模型和解码模型学习的目标子数据的排列规则,排列各个样本输入语义元素。

[0102] 由于解码模型输出的目标数据中各个目标子数据的排列顺序,与各个目标子数据各自对应的初始子数据在初始数据中的排列顺序可能不同,因此在获得各个初始子数据对应的样本输入语义元素之后,可以根据编码模型和解码模型学习的目标子数据的排列规则,确定样本输入语义元素在样本输入语义元素序列中的排列顺序,从而,数据转换模型输

出的训练输出语义元素序列中各个训练输出语义元素是按照编码模型和解码模型学习的目标子数据的排列规则进行排列的,解码模型输出的目标子数据也是按照编码模型和解码模型学习的目标子数据的排列规则进行排列的,数据转换模型和解码模型不需要再执行数据排列的过程。

[0103] 例如,初始数据为中文语句,A在跟一个女孩交往,各个初始子数据包括“A”、“在”、“跟”、“一个”、“女孩”和“交往”这六个初始子数据。目标数据为英文语句,A is dating a girl.各个目标子数据包括“A”、“is”、“dating”、“a”和“girl”。编码模型和解码模型学习的目标子数据的排列规则为“A”、“is”、“dating”、“a”和“girl”的顺序。因此,样本输入语义元素序列中的各个样本输入语义元素可以按照“A”、“在”、“跟”、“交往”、“一个”和“女孩”的顺序排列。

[0104] S202,基于数据转换模型,获得训练输出语义元素序列。

[0105] 服务端102在基于编码模型获得样本输入语义元素序列之后,可以将样本输入语义元素序列输入数据转换模型。服务端102采用数据转换模型对样本输入语义元素序列中的各个样本输入语义元素进行数据转换处理,获得各个训练输出语义元素。服务端102基于编码模型和解码模型学习的排列规则,对各个训练输出语义元素进行排序,获得训练输出语义元素序列。

[0106] 例如,初始数据为中文语句,A在跟一个女孩交往,各个初始子数据包括“A”、“在”、“跟”、“一个”、“女孩”和“交往”这六个初始子数据。目标数据为英文语句,A is dating a girl.各个目标子数据包括“A”、“is”、“dating”、“a”和“girl”。各个样本输入语义元素的排列顺序为“A”、“在”、“跟”、“一个”、“女孩”和“交往”。服务端102确定编码模型和解码模型学习的目标子数据的排列规则为“A”、“is”、“dating”、“a”和“girl”的顺序。因此,服务端102可以将各个训练输出语义元素按照“A”、“is”、“dating”、“a”和“girl”的顺序排列,获得训练输出语义元素序列。

[0107] S203,基于解码模型,获得目标数据。

[0108] 在服务端102在获得训练输出语义元素序列之后,将训练输出语义元素序列输入解码模型,获得训练输出语义元素序列中各个训练输出语义元素各自对应的目标子数据。如果训练输出语义元素序列中各个训练输出语义元素按照编码模型和解码模型学习的目标子数据的排列规则进行排列,那么获得的各个目标子数据可以按照各个训练输出语义元素在训练输出语义元素序列中的排列顺序进行排列,获得初始数据对应的目标数据。如果训练输出语义元素序列中各个训练输出语义元素按照初始子数据在初始数据中的排列顺序或其他排列顺序进行排列,那么各个目标子数据按照编码模型和解码模型学习的目标子数据的排列规则进行排列,获得初始数据对应的目标数据。例如,按照排列规则进行排列的各个目标子数据为“A”、“is”、“dating”、“a”和“girl”,那么获得的目标数据为“A is dating a girl”。

[0109] 下面基于图2a,对数据转换模型的一次训练过程进行具体介绍。请参考图3,为本申请实施例提供的训练数据转换模型的方法的一种流程示意图。

[0110] S301,获得训练输出语义元素序列中的各个训练输出语义位置,与样本输入语义元素序列中各个样本输入语义元素之间的元素相关概率。

[0111] 由于数据转换模型是结合编码模型和解码模型进行训练的,服务端102可以基于

初始子数据与目标子数据之间的数据相关概率,确定对应的样本输入语义元素与训练输出语义元素之间的元素相关概率。

[0112] 在训练过程中,通过不断的基于目标数据和初始数据,调整编码模型的模型参数和解码模型的模型参数,编码模型和解码模型可以学习到在确定各个目标子数据时,各个目标子数据在目标数据中的关联位置,分别与各个初始子数据之间的数据相关概率。一关联位置与一初始子数据之间的数据相关概率越大表示在该关联位置上的目标子数据与该初始子数据越相关;一关联位置与一初始子数据之间的数据相关概率越小表示在该关联位置上的目标子数据与该初始子数据越不相关。

[0113] 例如,在目标数据中,每个关联位置上的目标子数据为“A”、“is”、“dating”、“a”和“girl”。各个初始子数据为“A”、“在”、“跟”、“一个”、“女孩”和“交往”。其中,目标子数据“dating”在目标数据中的关联位置,与初始数据中各个初始子数据之间的数据相关概率包括,“dating”与“A”之间的数据相关概率为0.18,“dating”与“在”之间的数据相关概率为0.02,“dating”与“跟”之间的数据相关概率为0.05,“dating”与“一个”之间的数据相关概率为0.03,“dating”与“女孩”之间的数据相关概率为0.04,“dating”与“交往”之间的数据相关概率为0.68。“dating”在目标数据中的关联位置与“交往”是最相关的,“dating”在目标数据中的关联位置与“在”是最不相关的。因此,在基于初始数据“A在跟一个女孩交往”,获得目标子数据“dating”时,应该更多的注意“交往”,更少的注意“在”。

[0114] 在获得各个目标子数据在目标数据中的关联位置,分别与各个初始子数据之间的数据相关概率之后,服务端102将各个数据相关概率确定为对应的训练输出语义元素序列中的各个训练输出语义位置,与所述样本输入语义元素序列中各个样本输入语义元素之间的元素相关概率。

[0115] 例如,“dating”在目标数据中的关联位置与“A”之间的数据相关概率,确定为“dating”对应的训练输出语义元素在训练输出语义元素序列中的训练输出语义位置,与“A”对应的样本输入语义元素之间的元素相关概率。

[0116] S302,采用数据转换模型,基于获得的各个元素相关概率,分别获得各个训练输出语义位置各自对应的全局训练输出语义元素和局部训练输出语义元素。

[0117] 服务端102在获得样本输入语义元素序列之后,服务端102可以针对训练输出语义元素序列中的各个训练输出语义位置,获得每个训练输出语义位置上的训练输出语义元素,以获得训练输出语义元素序列。

[0118] 针对一个训练输出语义位置,服务端102可以根据该训练输出语义位置对应的全局训练输出语义元素和局部训练输出语义元素,来确定该训练输出语义位置对应的训练输出语义元素。全局训练输出语义元素用于表征样本输入语义元素序列中全部样本输入语义元素针对该输出语义位置的语义特征,局部训练输出语义元素用于表征样本输入语义元素序列中部分样本输入语义元素针对该训练输出语义位置的语义特征。

[0119] 例如,初始数据为中文语句,“A在跟一个女孩交往”,样本输入语义元素序列包括各个初始子数据“A”、“在”、“跟”、“一个”、“女孩”和“交往”对应的各个样本输入语义元素。目标数据为英文语句,“A is dating a girl”。训练输出语义元素序列包括各个目标子数据“A”、“is”、“dating”、“a”和“girl”对应的训练输出语义元素。那么针对训练输出语义元素序列中第三个关联位置,即针对目标子数据“dating”对应的训练输出语义元素,服务端

102可以通过注意“A”、“在”、“跟”、“一个”、“女孩”和“交往”每个初始子数据对应的样本输入语义元素,来从宏观的角度确定出目标子数据“dating”对应的训练输出语义元素。也就是说,利用全局训练输出语义元素可以从宏观的角度表征每个样本输入语义元素的语义特征。

[0120] 进一步的,在训练过程中,模型容易将注意力集中在与训练输出语义元素最相关的样本输入语义元素上,产生过拟合的情况。为了减少过拟合对模型准确性的影响,可以将注意力分散到其他与训练输出语义元素次相关的样本输入语义元素上。

[0121] 例如,针对训练输出语义元素序列中第三个关联位置,即针对目标子数据“dating”对应的训练输出语义元素,服务端102如果将过多的注意力集中在初始子数据“交往”对应的样本输入语义元素上,而极少的注意力集中在其他初始子数据对应的样本输入语义元素上,那么容易将“交往”翻译成“socializing with”。服务端102可以通过将“交往”上的注意力分散到“A”、“女孩”和“交往”上,来从微观的角度确定出目标子数据“dating”对应的训练输出语义元素,减少了确定出目标子数据“socializing with”对应的训练输出语义元素的可能性。也就是说,利用局部训练输出语义元素可以从微观的角度表征部分样本输入语义元素的语义特征。

[0122] 下面对获得一个训练输出语义位置对应的全局训练输出语义元素和局部训练输出语义元素的过程进行具体介绍。请参考图4,为获得一个训练输出语义位置对应的全局训练输出语义元素和局部训练输出语义元素的流程示意图。

[0123] S401,基于获得的各个元素相关概率,获得一个训练输出语义位置对应的全局训练输出语义元素。

[0124] 基于获得的各个元素相关概率,分别确定样本输入语义元素序列中的各个样本输入语义元素,对应于该训练输出语义位置的第一元素权重。服务端102将该训练输出语义位置,与样本输入语义元素序列中第一个样本输入语义元素之间的元素相关概率,作为第一个样本输入语义元素对应于该训练输出语义位置的第一元素权重。服务端102将该训练输出语义位置,与样本输入语义元素序列中第二个样本输入语义元素之间的元素相关概率,作为第二个样本输入语义元素对应于该训练输出语义位置的第一元素权重。从而获得样本输入语义元素序列中的各个样本输入语义元素,对应于该训练输出语义位置的第一元素权重。

[0125] 在获得各个样本输入语义元素对应于该训练输出语义位置的第一元素权重之后,服务端102基于各个第一元素权重,对各个样本输入语义元素进行加权求和处理,确定该训练输出语义位置对应的全局训练输出语义元素。

[0126] S402,基于获得的各个元素相关概率,确定该训练输出语义位置对应的目标相关样本输入语义元素。

[0127] 基于获得的各个元素相关概率,服务端102确定该训练输出语义位置对应的元素相关概率中的最大值,即最大元素相关概率。在获得该训练输出语义位置对应的最大元素相关概率之后,服务端102在样本输入语义元素序列中,确定最大元素相关概率对应的样本输入语义元素,获得该训练输出语义位置对应的最相关样本输入语义元素。

[0128] 在获得该训练输出语义位置对应的最相关样本输入语义元素之后,服务端102在样本输入语义元素序列中,确定与最相关样本输入语义元素满足预设相关条件的至少一个

样本输入语义元素,将获得的至少一个样本输入语义元素确定为该训练输出语义位置对应的目标相关样本输入语义元素。

[0129] 作为一种实施例,确定与最相关样本输入语义元素满足预设相关条件的至少一个样本输入语义元素的方法有多种,具体可以根据实际使用场景设置预设相关条件,下面以其中的两种确定方法为例进行介绍。

[0130] 确定方法一:

[0131] 将样本输入语义位置与最相关样本输入语义元素的样本输入语义位置之间的距离在预设的距离范围内的至少一个样本输入语义元素,确定为一个训练输出语义位置对应的目标相关样本输入语义元素。

[0132] 针对一个训练输出语义位置,在样本输入语义元素序列中确定了对应的最相关样本输入语义元素之后,可以将最相关样本输入语义元素,以及排在最相关样本输入语义元素前后各一个样本输入语义元素确定为目标相关样本输入语义元素;或者,可以将排在最相关样本输入语义元素之后,且与最相关样本输入语义元素相距两个样本输入语义元素的样本输入语义元素确定为目标相关样本输入语义元素等,具体可以根据预设的距离范围确定,具体不做限制。

[0133] 例如,与“dating”对应的训练输出语义元素最相关的是“交往”对应的样本输入语义元素,可以将“交往”对应的样本输入语义元素,以及排在“交往”对应的样本输入语义元素之前的两个样本输入语义元素,即“一个”和“女孩”对应的样本输入语义元素确定为与“dating”对应的训练输出语义元素目标相关样本输入语义元素。

[0134] 确定方法二:

[0135] 将元素相关概率与最大元素相关概率之间的误差在预设的误差范围内的至少一个样本输入语义元素,确定为一个训练输出语义位置对应的目标相关样本输入语义元素。

[0136] 针对一个训练输出语义位置,在样本输入语义元素序列中确定了对应的最相关样本输入语义元素之后,可以将元素相关概率小于或等于最大元素相关概率的至少一个样本输入语义元素,确定为一个训练输出语义位置对应的目标相关样本输入语义元素,选择几个样本输入语义元素,可以根据实际情况对预设的误差范围进行设置,具体不做限制;或者,可以将元素相关概率与最大元素相关概率之间的误差为指定数值的至少一个样本输入语义元素,确定为一个训练输出语义位置对应的目标相关样本输入语义元素,具体不做限制。

[0137] 例如,与“dating”对应的训练输出语义元素最相关的是“交往”对应的样本输入语义元素,最大元素相关概率为0.68。可以将元素相关概率排在“交往”对应的最大元素相关概率之后的两个样本输入语义元素,即“A”和“跟”对应的样本输入语义元素确定为与“dating”对应的训练输出语义元素目标相关样本输入语义元素。或者,可以将元素相关概率与“交往”对应的最大元素相关概率之间的误差在0.63~0.64内的样本输入语义元素,即“一个”和“女孩”对应的样本输入语义元素确定为与“dating”对应的训练输出语义元素目标相关样本输入语义元素。

[0138] S403,基于获得的各个元素相关概率,确定各个训练输出语义位置各自对应的局部训练输出语义元素。

[0139] 在获得该训练输出语义位置对应的目标相关样本输入语义元素之后,服务端102

分别确定每个目标相关样本输入语义元素对应于该训练输出语义位置的第二元素权重。服务端102确定每个目标相关样本输入语义元素对应于该训练输出语义位置的第二元素权重的过程,与S401中服务端102确定每个样本输入语义元素对应于该训练输出语义位置的第一元素权重的过程相同,在此不再赘述。

[0140] 在获得各个目标相关样本输入语义元素对应于该训练输出语义位置的第二元素权重之后,服务端102基于各个第二元素权重,对各个目标相关样本输入语义元素进行加权求和处理,确定该训练输出语义位置对应的局部训练输出语义元素。

[0141] 应当说明的是,S401与S402~S403之间没有必然的先后顺序关系,可以同时执行,或者可以先执行S401再执行S402~S403,或者可以先执行S402~S403再执行S401等。

[0142] S303,基于获得的各个全局训练输出语义元素和各个局部训练输出语义元素,分别确定各个训练输出语义位置各自对应的目标训练输出语义元素,以获得训练输出语义元素序列。

[0143] 服务端102在获得各个全局训练输出语义元素和各个局部训练输出语义元素之后,服务端102可以针对训练输出语义元素序列中的各个训练输出语义位置,分别确定对应的目标训练输出语义元素,以获得训练输出语义元素序列。

[0144] 针对一个训练输出语义位置,服务端102可以获得该训练输出语义位置对应的全局训练输出语义元素的全局权重和局部训练输出语义元素的局部权重。在获得该训练输出语义位置对应的全局训练输出语义元素和局部训练输出语义元素之后,根据全局权重和局部权重,对全局训练输出语义元素和局部训练输出语义元素进行加权求和处理,获得该训练输出语义位置对应的目标训练输出语义元素。

[0145] 全局权重和局部权重之间满足预设的权重关系,预设的权重关系例如是全局权重和局部权重之和为预设值“1”等,在此不对所有可能的权重关系进行一一列举。

[0146] 全局权重和局部权重可以在数据转换模型的训练过程中,通过不断的调整数据转换模型的模型参数之后学习到。如果直接根据样本输入语义元素序列集合来学习全局权重和局部权重,那么对于其它输入语义元素序列集合来说,该数据转换模型需要基于其它输入语义元素序列集合重新训练全局权重和局部权重,使得数据转换模型具有一定的局限性。因此,可以将全局权重设置为与对应的训练输出语义位置相关,从而局部权重也与对应的训练输出语义位置相关。在数据转换模型学习到每个训练输出语义位置应该输出的训练输出语义元素之后,全局权重和局部权重就可以与训练输出语义位置相匹配,及时针对其他输入语义元素序列集合,也不需要重新训练全局权重和局部权重,提高了数据转换模型的灵活性。

[0147] 下面对将全局权重设置为与对应的训练输出语义位置相关的过程进行示例介绍。

[0148] 针对一训练输出语义位置,对该训练输出语义位置对应的位置标识进行线性变换处理,获得该训练输出语义位置对应的线性变换后的位置标识。位置标识用于表征该训练输出语义位置,例如是表征该训练输出语义位置在训练输出语义元素序列中顺序的数值,或者,是对应的训练输出语义元素等,具体不做限制。

[0149] 将经过线性变换后的位置标识,确定为该训练输出语义位置对应的全局训练输出语义元素的全局权重,请参考公式(1)。

[0150] $g = W * Q_i$ (1)

[0151] 其中, g 表示全局训练输出语义元素的全局权重, W 表示一种线性变换处理,例如一个常数或一个常数矩阵等, i 表示训练输出语义位置, Q_i 表示训练输出语义位置对应的位置标识。

[0152] 或者,对经过线性变换后的位置标识进行归一化处理,获得归一化后的位置标识,将归一化后的位置标识确定为该训练输出语义位置对应的全局训练输出语义元素的全局权重,请参考公式(2)。

$$[0153] \quad g = \delta(W * Q_i) \quad (2)$$

[0154] 其中, $\delta(\cdot)$ 表示一种归一化处理,例如sigmoid函数。

[0155] S304,基于样本输入语义元素序列和训练输出语义元素序列确定训练损失,并基于训练损失对数据转换模型的模型参数进行调整。

[0156] 在获得训练输出语义元素序列之后,可以基于样本输入语义元素序列和训练输出语义元素序列确定数据转换模型的训练损失。由于编码模型、数据转换模型和解码模型是共同训练的,因此,可以基于编码模型的输入和解码模型的输出确定编解码训练损失,将编解码训练损失作为数据转换模型的训练损失。在获得数据转换模型的训练损失之后,可以确定数据转换模型的训练损失是否满足预设的收敛条件,如果满足预设的收敛条件,那么执行S305;如果不满足预设的收敛条件,可以基于数据转换模型的训练损失对数据转换模型的模型参数进行调整,实现数据转换模型基于样本输入语义元素的一次学习过程。

[0157] 基于编码模型的输入和解码模型的输出确定编解码训练损失的过程可能有多种,例如基于初始数据和目标数据之间的相似度,确定编解码训练损失;或者基于目标数据相较于初始数据的最大似然估计,确定编解码训练损失等,具体不做限制。

[0158] S305,直到数据转换模型的训练损失满足预设的收敛条件为止,获得已训练的数据转换模型。

[0159] 在基于数据转换模型的训练损失对数据转换模型的模型参数进行调整,实现数据转换模型基于样本输入语义元素的一次学习过程之后,可以基于样本输入语义元素序列集合中的下一个样本输入语义元素序列继续对数据转换模型进行训练,直到数据转换模型的训练损失满足预设的收敛条件为止,获得已训练的数据转换模型。在训练过程中,并未引入新的模型参数,在不增加数据转换模型的训练负担的前提下,提高了数据转换模型的准确性。

[0160] 作为一种实施例,服务端102可以将训练过程中每次获得的数据发送至第二客户端1012进行数据展示或可视化显示,从而用户可以基于第二客户端1012直观的看到整个数据转换模型的训练过程,便于用户对训练过程进行分析,或对数据转换模型进行调整等。

[0161] 基于同一发明构思,本申请实施例还提供一种使用数据转换模型的方法,下面对使用数据转换模型的方法进行介绍,请参考图5a,为使用数据转换模型的方法的一种原理示意图。

[0162] 获得待处理输入语义元素序列,对待处理输入语义元素序列使用数据转换模型,待处理输入语义元素序列可以是采用编码模型获得的,或者可以采用其他模型或采用其他方式获得的,具体不做限制。获得转换输出语义元素序列中的各个转换输出语义位置,与待处理输入语义元素序列中各个待处理输入语义元素之间的元素相关概率。

[0163] 在获得各个元素相关概率之后,采用已训练的数据转换模型,基于获得的各个元

素相关概率,分别获得各个转换输出语义位置各自对应的全局转换输出语义元素和局部转换输出语义元素。在获得各个全局转换输出语义元素和各个局部转换输出语义元素之后,基于获得的各个全局转换输出语义元素和各个局部转换输出语义元素,分别确定各个转换输出语义位置各自对应的目标转换输出语义元素,以获得转换输出语义元素序列。

[0164] 作为一种实施例,每个转换输出语义位置对应的全局转换输出语义元素用于表征待处理输入语义元素序列中全部待处理输入语义元素针对相应的输出语义位置的语义特征,每个转换输出语义位置对应的局部转换输出语义元素用于表征待处理输入语义元素序列中部分待处理输入语义元素针对相应的转换输出语义位置的语义特征。

[0165] 本申请实施例中,采用数据转换模型,基于各个全局转换输出语义元素和各个局部转换输出语义元素,获得各个转换输出语义位置各自对应的目标转换输出语义元素,以获得转换输出语义元素序列。将对待处理输入语义元素序列多维度的描述一同作为获得转换输出语义元素序列的依据,使得获得的转换输出语义元素序列能够更加准确地表征待处理输入语义元素序列的语义,提高了数据转换模型的准确性。

[0166] 下面对本申请实施例提供的使用数据转换模型的方法进行介绍。

[0167] 本申请实施例中以服务端102基于编码模型获得待处理输入语义元素序列为例进行介绍。例如,服务端102接收第一客户端1011发送的初始数据,即中文语句“我爱喝水”,服务端102将初始数据输入编码模型,获得编码模型输出的待处理输入语义元素序列。例如,获得编码模型输出的每个词语“我”、“爱”、“喝”和“水”对应的待处理输入语义元素。

[0168] 服务端102将获得的待处理输入语义元素序列输入数据转换模型,获得数据转换模型输出的转换输出语义元素序列。服务端102将获得的转换输出语义元素序列输入解码模型,获得解码模型输出的各个转换输出语义元素对应的目标子数据。每个转换输出语义元素可以对应于一个目标子数据,即“I”、“like”、“drink”和“water”,从而获得目标数据“I like drink water”,实现中英文文本翻译。

[0169] 又例如,服务端102接收第一客户端1011发送的图像,图像中包括包含草坪的图像区域、包含人的图像区域和包含狗的图像区域。服务端102将图像输入编码模型,获得编码模型输出的每个图像区域对应的待处理输入语义元素,服务端102获得待处理输入语义元素序列。

[0170] 服务端102将获得的待处理输入语义元素序列输入数据转换模型,获得数据转换模型输出的转换输出语义元素序列。服务端102将获得的转换输出语义元素序列输入解码模型,获得解码模型输出的各个转换输出语义元素对应的目标子数据,包括“目标A”、“草坪”和“狗”,从而获得目标数据“目标A在草坪上遛狗”。

[0171] 服务端102基于编码模型,获得待处理输入语义元素序列的过程,与S201中介绍的服务端102基于编码模型,获得样本输入语义元素序列的过程类似,在此不再赘述。服务端102基于解码模型,获得目标数据的过程,与S203中介绍的服务端102基于解码模型,获得目标数据的过程类似,在此不再赘述。

[0172] 下面对使用数据转换模型的过程进行具体介绍,请参考图5b,为本申请实施例提供的使用数据转换模型的方法的一种流程示意图。

[0173] S501,获得待处理输入语义元素序列,以及转换输出语义元素序列中的各个转换输出语义位置,与待处理输入语义元素序列中各个待处理输入语义元素之间的元素相关概

率。

[0174] 服务端102获得转换输出语义元素序列中的各个转换输出语义位置,与待处理输入语义元素序列中各个待处理输入语义元素之间的元素相关概率的过程,与S301,服务端102获得训练输出语义元素序列中的各个训练输出语义位置,与样本输入语义元素序列中各个样本输入语义元素之间的元素相关概率的过程类似,在此不再赘述。

[0175] S502,采用数据转换模型,基于获得的各个元素相关概率,分别获得各个转换输出语义位置各自对应的全局转换输出语义元素和局部转换输出语义元素。

[0176] 服务端102采用数据转换模型,基于获得的各个元素相关概率,分别获得各个转换输出语义位置各自对应的全局转换输出语义元素和局部转换输出语义元素的过程,与S302,服务端102采用数据转换模型,基于获得的各个元素相关概率,分别获得各个训练输出语义位置各自对应的全局训练输出语义元素和局部训练输出语义元素的过程类似,在此不再赘述。

[0177] S503,基于获得的各个全局转换输出语义元素和各个局部转换输出语义元素,分别确定各个转换输出语义位置各自对应的目标转换输出语义元素,以获得转换输出语义元素序列。

[0178] 服务端102基于获得的各个全局转换输出语义元素和各个局部转换输出语义元素,分别确定各个转换输出语义位置各自对应的目标转换输出语义元素,以获得转换输出语义元素序列的过程,与S303,基于获得的各个全局训练输出语义元素和各个局部训练输出语义元素,分别确定各个训练输出语义位置各自对应的目标训练输出语义元素,以获得训练输出语义元素序列的过程类似,在此不再赘述。

[0179] 作为一种实施例,服务端102可以将使用数据转换模型的过程在第二客户端1012上进行可视化显示,从而用户可以基于第二客户端1012直观的看到整个数据转换模型的使用过程,便于用户对使用数据转换模型的过程进行分析,或对数据转换模型进行调整等。

[0180] 作为一种实施例,为了确定数据转换模型针对一个转换输出语义位置,是否将注意力分散在最相关待处理输入语义元素以外的其他相关待处理输入语义元素,服务端102在获得转换输出语义元素之后,可以确定数据转换模型的感知能力。确定数据转换模型的感知能力的方法有多种,下面以其中的一种为例进行示例介绍。请参考图6,为确定数据转换模型的感知能力的一种流程示意图。

[0181] S601,基于各个元素相关概率,获得转换输出语义元素序列对应的局部熵。

[0182] 服务端102基于每个转换输出语义位置与待处理输入语义元素序列中各个待处理输入语义元素之间的元素相关概率,可以获得每个转换输出语义位置对应的元素相关概率分布。

[0183] 例如,待处理输入语义元素序列为 $\{f_1, f_2, \dots, f_j\}$,转换输出语义元素序列为 $\{e_1, e_2, \dots, e_i\}$,一个转换输出语义位置pos与待处理输入语义元素序列中第一个待处理输入语义元素 f_1 之间的元素相关概率为 $P(f_1 | \text{pos})$,那么该转换输出语义位置pos对应的元素相关概率分布 P_{pos} ,请参考公式(3)。

[0184]
$$P_{\text{pos}} = \{P(f_1 | \text{pos}), P(f_2 | \text{pos}), \dots, P(f_j | \text{pos})\} \quad (3)$$

[0185] 作为一种实施例,如果解码模型包括多个解码层,那么可以分别针对每个解码层,确定每个转换输出语义位置对应的元素相关概率分布,请参考公式(4),为第n层解码层中,

转换输出语义位置pos对应的元素相关概率分布 P_{pos}^n 。

$$[0186] \quad P_{\text{pos}}^n = \{ P^n(f_1|\text{pos}), P^n(f_2|\text{pos}), \dots, P^n(f_1|\text{pos}) \} \quad (4)$$

[0187] 在获得每个转换输出语义位置对应的元素相关概率分布之后,请参考公式(5),确定转换输出语义元素序列对应的局部熵LE。

$$[0188] \quad LE = -\frac{1}{n \cdot i} \sum_{k \in [1, n]} \sum_{\text{pos} \in [1, i]} P_{\text{pos}}^n * \log_2 P_{\text{pos}}^n \quad (5)$$

[0189] 转换输出语义元素序列对应的局部熵LE越小,说明在获得转换输出语义元素序列时,数据转换模型的注意力越集中在一个待处理输入语义元素上;转换输出语义元素序列对应的局部熵LE越大,说明在获得转换输出语义元素序列时,数据转换模型的注意力越分散在不同的待处理输入语义元素上。

[0190] S602,基于获得的局部熵与预设的局部熵阈值的比对结果,获得数据转换模型的准确性分析结果。

[0191] 在获得转换输出语义元素序列对应的局部熵之后,服务端102可以比对获得的局部熵与预设的局部熵阈值,并基于比对结果,获得数据转换模型的准确性分析结果。例如,如果获得的局部熵大于预设的局部熵阈值,那么获得的数据转换模型的准确性分析结果表示数据转换模型的注意力分散在不同的待处理输入语义元素上,即所关注的待处理输入语义元素较为准确。或者,如果获得的局部熵大于预设的第一局部熵阈值,且小于预设的第二局部熵阈值,那么获得的数据转换模型的准确性分析结果表示数据转换模型的注意力较为集中,且分散在待处理输入语义元素序列中的部分待处理输入语义元素上,即所关注的待处理输入语义元素较为准确等。

[0192] 下面结合应用场景,对上述的使用数据转换模型的方法进行示例介绍。

[0193] 假设需要采用数据转换模型将中文语句“A在跟一个女孩交往”翻译为“A is dating a girl”。服务端102在接收第一客户端1011发送的中文语句“A在跟一个女孩交往”之后,将该中文语句作为初始数据输入编码模型中。服务端102获得编码模型输出的每个初始子数据“A”、“在”、“跟”、“一个”、“女孩”和“交往”分别对应的待处理输入语义元素。服务端102根据各个待处理输入语义元素获得待处理输入语义元素序列[“A”,“在”,“跟”,“一个”,“女孩”,“交往”]。

[0194] 以第三个关联位置,即目标子数据“dating”在目标数据中的位置为例,服务端102确定第三个关联位置与“A”之间的数据相关概率为0.18,第三个关联位置与“在”之间的数据相关概率为0.02,第三个关联位置与“跟”之间的数据相关概率为0.05,第三个关联位置与“一个”之间的数据相关概率为0.03,第三个关联位置与“女孩”之间的数据相关概率为0.04,第三个关联位置与“交往”之间的数据相关概率为0.68。

[0195] 那么,服务端102获得第三个转换输出语义位置与“A”对应的待处理输入语义元素之间的元素相关概率为0.18,第三个转换输出语义位置与“在”对应的待处理输入语义元素之间的元素相关概率为0.02,“第三个转换输出语义位置与“跟”对应的待处理输入语义元素之间的元素相关概率为0.05,第三个转换输出语义位置与“一个”对应的待处理输入语义元素之间的元素相关概率为0.03,第三个转换输出语义位置与“女孩”对应的待处理输入语义元素之间的元素相关概率为0.04,第三个转换输出语义位置与“交往”对应的待处理输入语义元素之间的元素相关概率为0.68。

[0196] 服务端102将元素相关概率作为对应的待处理输入语义元素的第一元素权重,对待处理输入语义元素进行加权求和处理,获得第三个转换输出语义位置对应的全局转换输出语义元素。从而,服务端102可以获得每个转换输出语义位置对应的全局转换输出语义元素。

[0197] 根据预设相关条件,确定目标相关输入语义元素为“一个”、“女孩”和“交往”分别对应的待处理输入语义元素。服务端102确定第三个关联位置与“一个”之间的数据相关概率为0.14,第三个关联位置与“女孩”之间的数据相关概率为0.15,第三个关联位置与“交往”之间的数据相关概率为0.71。那么,服务端102获得第三个转换输出语义位置与“一个”对应的待处理输入语义元素之间的元素相关概率为0.14,第三个转换输出语义位置与“女孩”对应的待处理输入语义元素之间的元素相关概率为0.15,第三个转换输出语义位置与“交往”对应的待处理输入语义元素之间的元素相关概率为0.71。

[0198] 服务端102将元素相关概率作为对应的目标相关待处理输入语义元素的第二元素权重,对待处理输入语义元素进行加权求和处理,获得第三个转换输出语义位置对应的局部转换输出语义元素。从而,服务端102可以获得每个转换输出语义位置对应的局部转换输出语义元素。

[0199] 服务端102基于全局权重和局部权重,分别对每个转换输出语义位置对应的全局转换输出语义元素和局部转换输出语义元素进行加权求和处理,获得每个转换输出语义位置对应的目标转换输出语义元素。从而,服务端102可以获得转换输出语义元素序列。

[0200] 服务端102基于数据转换模型获得转换输出语义元素序列之后,可以将转换输出语义元素序列输入解码模型,获得解码模型针对每个转换输出语义元素输出的目标子数据“A”、“is”、“dating”、“a”和“girl”,从而,获得目标数据“A is dating a girl”。服务端102在获得目标数据“A is dating a girl”之后,可以发送至第二客户端1012,以使第二客户端1012显示目标数据“A is dating a girl”。可见,数据转换模型将更多的注意力分散在了“一个”和“女孩”上,使得在对“交往”进行翻译时,不会翻译为“socializing with”,而是会翻译为“dating”,显然翻译为“dating”更为准确。

[0201] 数据转换模型可以应用于任何模型中,例如递归神经网络RNN,请参考图7a(1),或自关注神经网络SAN,请参考图7a(2)。RNN每个目标子数据只与对应的初始子数据相关,目标数据的准确性较低。而SAN基于每个初始子数据获得每个目标子数据,目标数据的准确性较高。

[0202] 自关注神经网络SAN包括自回归机器翻译模型(Autoregressive Machine Translation,AT),请参考图7b(1),或非自回归神经机器翻译模型(Non-Autoregressive Machine Translation,NAT),请参考图7b(2)等。以Transformer模型为例,请参考图7c,为Transformer模型的结构示意图。其中的遮掩多头自注意力层用于遮掩目标数据中右侧的目标子数据,保证学习过程是自左向右的,如遮掩未输出的目标子数据,保证学习过程是基于已输出的目标子数据的。

[0203] 对于AT,在初始数据 $X = X_1, X_2, \dots, X_j$,目标数据 $Y = Y_1, Y_2, \dots, Y_i$ 的情况下,目标数据的条件概率 $P(Y|X; \theta)$,请参考公式(7)。

$$[0204] \quad P(Y|X; \theta) = \prod_{k=1}^K P(Y_k | Y < k, X; \theta) \quad (7)$$

[0205] 其中,表示AT的模型参数, $Y < k$ 表示部分初始子数据。

[0206] 在训练AT的过程中,确定条件概率的最大似然估计 $L(\theta)$,请参考公式(8)。

$$[0207] \quad L(\theta) = \arg \max_{\theta} \sum_{m=1}^M \log P(Y^m|X^m; \theta) \quad (8)$$

[0208] 对于NAT,条件概率 $P(Y|X)$ 则解除了 $Y \ll k$ 的条件依赖,请参考公式(9)。

$$[0209] \quad P(Y|X) = \prod_{t=1}^T P(Y_t|X) \quad (9)$$

[0210] 由于AT是逐个初始子数据产生对应的目标子数据,每个初始子数据依赖之前产生的目标子数据,请参考表1。AT模型复杂度较高,从而使得获得目标数据的效率较低,而NAT是每个初始子数据独立产生对应的目标子数据,可以同时生成所有目标子数据,获得目标数据,获得目标数据的效率较高,请参考表2。

[0211] 表1

	左侧目标子数据的 准确率(%)	右侧目标子数据的 准确率(%)
[0212] 自左向右解码方式	40.21	35.10
自右向左解码方式	35.67	39.47

[0213] 表2

	时间(ms)	BLEU评价指标
[0214] NAT	210	4.76
AT	353	2.83

[0215] 本申请实施例中以数据转换模型应用于条件掩码语言模型(Conditional Masked Language Models, CMLMs)中为例,对本申请实施例提供的数据转换模型进行示例介绍。训练条件掩码语言模型的方法在此不再赘述,下面对训练数据转换模型的方法进行示例介绍,在一次训练过程中,针对样本输入语义元素序列集合中的样本输入语义元素序列,可以执行以下操作。

[0216] 获得训练输出语义元素序列中的各个训练输出语义位置,与样本输入语义元素序列中各个样本输入语义元素之间的元素相关概率 φ_{pos} ,请参考公式(10)。

$$[0217] \quad \varphi_{\text{pos}} = Q_{\text{pos}} K^T \quad (10)$$

[0218] 其中, Q_{pos} 表示训练输出语义位置, K 表示样本输入语义元素在样本输入语义元素序列中的关联位置。

[0219] 基于各个元素相关概率 φ_{pos} ,针对一个训练输出语义位置 Q_{pos} ,确定该训练输出语义位置的全局训练输出语义元素 $\text{ATT}(\varphi_{\text{pos}}, V)$,请参考公式(11)。

$$[0220] \quad \text{ATT}(\varphi_{\text{pos}}, V) = \text{softmax}(\varphi_{\text{pos}}) V \quad (11)$$

[0221] 其中, V 表示样本输入语义元素。

[0222] 基于部分元素相关概率 $L(\varphi_{\text{pos}})$,确定该训练输出语义位置的局部训练输出语义元素 $\text{ATT}(L(\varphi_{\text{pos}}), V)$ 。其中,部分元素相关概率 $L(\varphi_{\text{pos}})$,请参考公式(12)。

$$[0223] \quad L(\varphi_{\text{pos}}) = \begin{cases} \varphi_{\text{pos}, r} & r_0 - \text{win} \leq r \leq r_0 + \text{win} \\ -\infty & \text{otherwise} \end{cases} \quad (12)$$

[0224] 其中, r 表示目标相关样本输入语义元素在样本输入语义元素序列中的关联位置, r_0 表示与训练输出语义位置 Q_{pos} 最相关的样本输入语义元素在样本输入语义元素序列中的关联位置, win 表示预设相关条件, $\varphi_{\text{pos}, r}$ 表示部分元素相关概率。

[0225] 基于全局训练输出语义元素 $\text{ATT}(\varphi_{\text{pos}}, V)$ 和局部训练输出语义元素 $\text{ATT}(L(\varphi_{\text{pos}}), V)$, 确定目标训练输出语义元素 $\text{CCAN}(Q_{\text{pos}}, K, V)$, 请参考公式(13)。

$$[0226] \quad \text{CCAN}(Q_{\text{pos}}, K, V) = g \cdot \text{ATT}(\varphi_{\text{pos}}, V) + (1 - g) \cdot \text{ATT}(L(\varphi_{\text{pos}}), V) \quad (13)$$

[0227] 其中, 全局训练输出语义元素的权重可以与 Q_{pos} 的线性变换相关, 请参考公式(14)。

$$[0228] \quad g = \delta(WQ_{\text{pos}}) \quad (14)$$

[0229] 其中, W 表示线性变换的参数, $\delta(\cdot)$ 表示sigmoid函数。

[0230] 基于样本输入语义元素序列和训练输出语义元素序列确定数据转换模型的训练损失, 并基于数据转换模型的训练损失对数据转换模型的模型参数进行调整。

[0231] 在条件掩码语言模型结合数据转换模型训练完成之后, 可以在数据集上对已训练的条件掩码语言模型结合数据转换模型进行测试实验。数据集可以采用小规模的王MT16罗马尼亚语-英语Romanian-English (Ro-En), 中等规模的王MT14英语-德语English-German (En-De), 大规模的王MT17中文-英语(Zh-En)和语序不同的WAT17日语-英语(Ja-En)等。在将初始数据输入Transformer模型之前, 可以先通过BPE字词切分使用32K的合并操作对数据进行预处理, 最后使用BLEU作为评价指标, 并进行统计学显著性检验。

[0232] 作为一种实施例, 可以使用知识蒸馏方法来简化训练数据。在教师模型方面, 利用初始数据集对BASE和BIG Transformer模型进行训练。在大模型中, 采用big batch策略(458K个词每batch)来优化性能。条件掩码语言模型可以包括6层编码层和6层解码层, 其中解码层采用条件掩码语言模型方式训练。条件掩码语言模型8个头为512, 前馈网络为2048。

[0233] 请参考表3, 在王MT16 RO-EN、王MT14 En-De、王MT17 ZH-EN和王AT17 JA-EN数据集上, 条件掩码语言模型结合数据转换模型与传统的NAT模型比较。根据表3, 条件掩码语言模型结合数据转换模型的翻译性能更高。

[0234] 表3

	模型	Ro-En	En-De	Zh-En	
	1	Transformer-BASE	34.1	27.3	24.4
	2	Transformer-BIG	/	29.2	25.3
	3	NAT	31.4	19.2	/
	4	Iterative	30.2	21.6	/
[0235]	5	Iterative NAT	33.3	26.8	/
	6	DisCo	33.3	27.3	/
	7	CMLMs	33.3	27.0	23.2
	8	CMLMs+ 数据转换模型	33.7	27.5	24.6

[0236] 通过设置全局权重 g ,将局部权重转换为元素相关概率,即重要程度。计算每个解码器层的局部注意力,请参考图8。在对数据转换的过程中,localness持续下降,表示数据转换模型的注意力持续分散,直到倒数第二层增加,数据转换模型的注意力开始集中。也就是说,中间层主要负责数据转换所以注意力分散,顶层需要生成目标数据所以注意力开始集中。

[0237] 数据转换模型应该将更多的注意力集中在最相关待处理输入语义元素,以及其他相关待处理输入语义元素。为了数据转换模型注意力的准确性,请参考图9,通过设置不同的win,获得不同的数据转换模型注意力的精度。条件掩码语言模型结合数据转换模型始终优于基线(Δ Accuracy),说明增强了条件掩码语言模型的注意力的精度。

[0238] 为了验证条件掩码语言模型结合数据转换模型学习到的语言特性,可以采用探究性任务从语言学角度进行定量。探究性任务可以分为三种类型,“表层任务”用于从句子嵌入中学到的简单表层属性;“句法任务”用于量化句法保留能力;“语义”用于评估更深层次的语义表征能力,请参考表4。可见,条件掩码语言模型结合数据转换模型保留了丰富的语法和语义信息。

[0239] 表4

	Surface		Syntactic			Semantic				
	SeLen	WC	TrDep	ToCo	BShif	Tense	SubNm	ObjNm	SoMo	
[0240]	CMLMs	93.2	79.4	45.8	79.4	73.5	88.9	87.5	85.4	52.7
	CMLMs+ 数据转换 模型	92.8	78.1	46.5	79.7	74.1	88.3	88.1	85.7	52.9

[0241] 基于同一发明构思,本申请实施例提供一种训练数据转换模型的装置,该装置相当于前文论述的服务端102,能够实现前述训练数据转换模型的方法对应的功能。请参考图

10,该装置包括训练模块1001和获取模块1002,其中:

[0242] 训练模块1001:用于采用样本输入语义元素序列集合对数据转换模型进行训练,获得已训练的数据转换模型;其中,在一次训练过程中,针对样本输入语义元素序列集合中的样本输入语义元素序列,至少执行以下操作:

[0243] 获取模块1002:用于获得训练输出语义元素序列中的各个训练输出语义位置,与样本输入语义元素序列中各个样本输入语义元素之间的元素相关概率;

[0244] 训练模块1001还用于:采用数据转换模型,基于获得的各个元素相关概率,分别获得各个训练输出语义位置各自对应的全局训练输出语义元素和局部训练输出语义元素,并基于获得的各个全局训练输出语义元素和各个局部训练输出语义元素,分别确定各个训练输出语义位置各自对应的目标训练输出语义元素,以获得训练输出语义元素序列,其中,全局训练输出语义元素与每个样本输入语义元素相关,局部训练输出语义元素与部分样本输入语义元素相关。

[0245] 在一种可能的实施例中,获取模块1002具体用于:

[0246] 基于编码模型和解码模型学习的目标子数据的排列规则,分别确定各个训练输出语义位置各自对应的目标子数据在目标数据中的关联位置,与初始数据中各个样本输入语义位置各自对应的各个初始子数据之间的数据相关概率,其中,样本输入语义元素序列是采用编码模型对初始数据进行编码处理获得的,目标数据是采用与编码模型对应的解码模型对训练输出语义元素序列进行解码处理获得的;

[0247] 基于获得的各个数据相关概率,确定训练输出语义元素序列中的各个训练输出语义位置,与样本输入语义元素序列中各个样本输入语义元素之间的元素相关概率。

[0248] 在一种可能的实施例中,训练模块1001具体用于:

[0249] 针对各个训练输出语义位置,分别执行以下操作:

[0250] 基于获得的各个元素相关概率,分别获得样本输入语义元素序列中的各个样本输入语义元素,对应于各个训练输出语义位置中的一个训练输出语义位置的第一元素权重;

[0251] 基于各个样本输入语义元素对应的第一元素权重,对各个样本输入语义元素进行加权求和处理,获得一个训练输出语义位置对应的全局训练输出语义元素。

[0252] 在一种可能的实施例中,训练模块1001具体用于:

[0253] 针对各个训练输出语义位置,分别执行以下操作:

[0254] 基于获得的各个元素相关概率,在样本输入语义元素序列中,确定各个训练输出语义位置中的一个训练输出语义位置对应的目标相关样本输入语义元素;

[0255] 基于获得的各个元素相关概率,分别获得样本输入语义元素序列中的各个目标相关样本输入语义元素,对应于一个训练输出语义位置的第二元素权重;

[0256] 基于各个目标相关样本输入语义元素对应的第二元素权重,对各个目标相关样本输入语义元素进行加权求和处理,获得一个训练输出语义位置对应的局部训练输出语义元素。

[0257] 在一种可能的实施例中,训练模块1001具体用于:

[0258] 基于获得的各个元素相关概率,确定一个训练输出语义位置对应的最大元素相关概率,获得样本输入语义元素序列中,一个训练输出语义位置对应的最相关样本输入语义元素;

[0259] 在样本输入语义元素序列中,将与最相关样本输入语义元素满足预设相关条件的至少一个样本输入语义元素,确定为一个训练输出语义位置对应的目标相关样本输入语义元素。

[0260] 在一种可能的实施例中,训练模块1001具体用于:

[0261] 在样本输入语义元素序列中,将样本输入语义位置与最相关样本输入语义元素的样本输入语义位置之间的距离在预设的距离范围内的至少一个样本输入语义元素,确定为一个训练输出语义位置对应的目标相关样本输入语义元素;或者,

[0262] 在样本输入语义元素序列中,将元素相关概率与最大元素相关概率之间的误差在预设的误差范围内的至少一个样本输入语义元素,确定为一个训练输出语义位置对应的目标相关样本输入语义元素。

[0263] 在一种可能的实施例中,训练模块1001具体用于:

[0264] 针对各个训练输出语义位置,分别执行以下操作:

[0265] 获得各个训练输出语义位置中的一个训练输出语义位置对应的全局训练输出语义元素的全局权重和局部训练输出语义元素的局部权重;

[0266] 基于获得的全局权重和局部权重,对一个训练输出语义位置对应的全局训练输出语义元素和局部训练输出语义元素进行加权求和处理,获得一个训练输出语义位置对应的目标训练输出语义元素。

[0267] 在一种可能的实施例中,训练模块1001具体用于:

[0268] 对一个训练输出语义位置对应的位置标识进行线性变换处理,获得一个训练输出语义位置对应的线性变换后的位置标识;

[0269] 将经过线性变换后的位置标识,确定为一个训练输出语义位置对应的全局训练输出语义元素的全局权重;

[0270] 将与获得的全局权重满足预设权重关系的权重系数,确定为一个训练输出语义位置对应的局部训练输出语义元素的局部权重。

[0271] 在一种可能的实施例中,训练模块1001还用于:

[0272] 在获得训练输出语义元素序列之后,基于训练输出语义元素序列对应的目标数据与样本输入语义元素序列对应的初始数据之间的相似度,确定编码模型和解码模型对应的编解码训练损失;

[0273] 将编解码训练损失确定为数据转换模型的训练损失,并基于训练损失对数据转换模型的模型参数进行调整。

[0274] 基于同一发明构思,本申请实施例提供一种使用数据转换模型的装置,该装置相当于前文论述的服务端102,能够实现前述使用数据转换模型的方法对应的功能。请参考图11,该装置包括获取模块1101和转换模块1102,其中:

[0275] 获取模块1101:用于获得待处理输入语义元素序列,以及转换输出语义元素序列中的各个转换输出语义位置,与待处理输入语义元素序列中各个待处理输入语义元素之间的元素相关概率;

[0276] 转换模块1102:用于采用数据转换模型,基于获得的各个元素相关概率,分别获得各个转换输出语义位置各自对应的全局转换输出语义元素和局部转换输出语义元素,并基于获得的各个全局转换输出语义元素和各个局部转换输出语义元素,分别确定各个转换输

出语义位置各自对应的目标转换输出语义元素,以获得转换输出语义元素序列,其中,全局转换输出语义元素与每个待处理输入语义元素相关,局部转换输出语义元素与部分待处理输入语义元素相关。

[0277] 在一种可能的实施例中,获取模块1101具体用于:

[0278] 基于编码模型和解码模型学习的目标子数据的排列规则,分别确定各个转换输出语义位置各自对应的目标子数据在目标数据中的关联位置,与初始数据中各个待处理输入语义位置各自对应的各个初始子数据之间的数据相关概率,其中,待处理输入语义元素序列是采用编码模型对初始数据进行编码处理获得的,目标数据是采用与编码模型对应的解码模型对转换输出语义元素序列进行解码处理获得的;

[0279] 基于获得的各个数据相关概率,确定转换输出语义元素序列中的各个转换输出语义位置,与待处理输入语义元素序列中各个待处理输入语义元素之间的元素相关概率。

[0280] 在一种可能的实施例中,转换模块1102具体用于:

[0281] 针对各个转换输出语义位置,采用数据转换模型,分别执行以下操作:

[0282] 基于获得的各个元素相关概率,分别获得待处理输入语义元素序列中的各个待处理输入语义元素,对应于各个转换输出语义位置中的一个转换输出语义位置的第一元素权重;

[0283] 基于各个待处理输入语义元素对应的第一元素权重,对各个待处理输入语义元素进行加权求和处理,获得一个转换输出语义位置对应的全局转换输出语义元素。

[0284] 在一种可能的实施例中,转换模块1102具体用于:

[0285] 针对各个转换输出语义位置,分别执行以下操作:

[0286] 基于获得的各个元素相关概率,在待处理输入语义元素序列中,确定各个转换输出语义位置中的一个转换输出语义位置对应的目标相关待处理输入语义元素;

[0287] 基于获得的各个元素相关概率,分别获得待处理输入语义元素序列中的各个目标相关待处理输入语义元素,对应于各个转换输出语义位置中的一个转换输出语义位置的第二元素权重;

[0288] 基于各个目标相关待处理输入语义元素对应的第二元素权重,对各个目标相关待处理输入语义元素进行加权求和处理,获得一个转换输出语义位置对应的局部转换输出语义元素。

[0289] 在一种可能的实施例中,转换模块1102具体用于:

[0290] 基于获得的各个元素相关概率,确定一个转换输出语义位置对应的最大元素相关概率,获得待处理输入语义元素序列中,一个转换输出语义位置对应的最相关待处理输入语义元素;

[0291] 在待处理输入语义元素序列中,将与最相关待处理输入语义元素满足预设相关条件的至少一个待处理输入语义元素,确定为一个转换输出语义位置对应的目标相关输入语义元素。

[0292] 在一种可能的实施例中,转换模块1102具体用于:

[0293] 在待处理输入语义元素序列中,将待处理输入语义位置与最相关待处理输入语义元素的待处理输入语义位置之间的距离在预设的距离范围内的至少一个待处理输入语义元素,确定为一个转换输出语义位置对应的目标相关待处理输入语义元素;或者,

[0294] 在待处理输入语义元素序列中,将元素相关概率与最大元素相关概率之间的误差在预设的误差范围内的至少一个待处理输入语义元素,确定为一个转换输出语义位置对应的目标相关待处理输入语义元素。

[0295] 在一种可能的实施例中,转换模块1102具体用于:

[0296] 针对各个转换输出语义位置,分别执行以下操作:

[0297] 获得各个转换输出语义位置中的一个转换输出语义位置对应的全局转换输出语义元素的全局权重和局部转换输出语义元素的局部权重;

[0298] 基于获得的全局权重和局部权重,对一个转换输出语义位置对应的全局转换输出语义元素和局部转换输出语义元素进行加权求和处理,获得一个转换输出语义位置对应的目标转换输出语义元素。

[0299] 在一种可能的实施例中,转换模块1102还用于:

[0300] 在获得转换输出语义元素序列之后,基于获得的各个元素相关概率,获得转换输出语义元素序列对应的局部熵;

[0301] 基于获得的局部熵与预设的局部熵阈值的比对结果,获得数据转换模型的准确性分析结果,其中,准确性分析结果用于表征在确定各个转换输出语义位置各自对应的目标转换输出语义元素时,所关注的待处理输入语义元素是否准确。

[0302] 基于同一发明构思,本申请实施例提供一种计算机设备,下面对该计算机设备1200进行介绍。

[0303] 请参照图12,上述训练、使用数据转换模型的装置可以运行在计算机设备1200上,训练、使用数据转换模型的程序的当前版本和历史版本以及训练、使用数据转换模型的程序对应的应用软件可以安装在计算机设备1200上,该计算机设备1200包括显示单元1240、处理器1280以及存储器1220,其中,显示单元1240包括显示面板1241,用于显示由用户交互操作界面等。

[0304] 在一种可能的实施例中,可以采用液晶显示器(Liquid Crystal Display,LCD)或有机发光二极管OLED(Organic Light-Emitting Diode)等形式来配置显示面板1241。

[0305] 处理器1280用于读取计算机程序,然后执行计算机程序定义的方法,例如处理器1280读取训练、使用数据转换模型的程序或文件等,从而在该计算机设备1200上运行训练、使用数据转换模型的程序,在显示单元1240上显示对应的界面。处理器1280可以包括一个或多个通用处理器,还可包括一个或多个DSP(Digital Signal Processor,数字信号处理器),用于执行相关操作,以实现本申请实施例所提供的技术方案。

[0306] 存储器1220一般包括内存和外存,内存可以为随机存储器(RAM),只读存储器(ROM),以及高速缓存(CACHE)等。外存可以为硬盘、光盘、USB盘、软盘或磁带机等。存储器1220用于存储计算机程序和其他数据,该计算机程序包括各客户端对应的应用程序等,其他数据可包括操作系统或应用程序被运行后产生的数据,该数据包括系统数据(例如操作系统的配置参数)和用户数据。本申请实施例中程序指令存储在存储器1220中,处理器1280执行存储其中1220中的程序指令,实现前文图论述的任意的一种训练、使用数据转换模型的方法。

[0307] 上述显示单元1240用于接收输入的数字信息、字符信息或接触式触摸操作/非接触式手势,以及产生与计算机设备1200的用户设置以及功能控制有关的信号输入等。具体

地,本申请实施例中,该显示单元1240可以包括显示面板1241。显示面板1241例如触摸屏,可收集用户在其上或附近的触摸操作(比如用户使用手指、触笔等任何适合的物体或附件在显示面板1241上或在显示面板1241的操作),并根据预先设定的程式驱动相应的连接装置。

[0308] 在一种可能的实施例中,显示面板1241可包括触摸检测装置和触摸控制器两个部分。其中,触摸检测装置检测玩家的触摸方位,并检测触摸操作带来的信号,将信号传送给触摸控制器;触摸控制器从触摸检测装置上接收触摸信息,并将它转换成触点坐标,再送给处理器1280,并能接收处理器1280发来的命令并加以执行。

[0309] 其中,显示面板1241可以采用电阻式、电容式、红外线以及表面声波等多种类型实现。除了显示单元1240,计算机设备1200还可以包括输入单元1230,输入单元1230可以包括图形输入设备1231和其他输入设备1232,其中其他输入设备可以但不限于包括物理键盘、功能键(比如音量控制按键、开关按键等)、轨迹球、鼠标、操作杆等中的一种或多种。

[0310] 除以上之外,计算机设备1200还可以包括用于给其他模块供电的电源1290、音频电路1260、近场通信模块1270和RF电路1210。计算机设备1200还可以包括一个或多个传感器1250,例如加速度传感器、光传感器、压力传感器等。音频电路1260具体包括扬声器1261和麦克风1262等,例如计算机设备1200可以通过麦克风1262采集用户的声音,进行相应的操作等。

[0311] 作为一种实施例,处理器1280的数量可以是一个或多个,处理器1280和存储器1220可以是耦合设置,也可以是相对独立设置。

[0312] 作为一种实施例,图12中的处理器1280可以用于实现如图10中的训练模块1001和获取模块1002的功能,也可以用于实现如图11中的获取模块1101和转换模块1102的功能。

[0313] 作为一种实施例,图12中的处理器1280可以用于实现前文论述的测试设备103对应的功能。

[0314] 本领域普通技术人员可以理解:实现上述方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成,前述的程序可以存储于一计算机可读取存储介质中,该程序在执行时,执行包括上述方法实施例的步骤;而前述的存储介质包括:移动存储设备、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0315] 或者,本发明上述集成的单元如果以软件功能模块的形式实现并作为独立的产品销售或使用,也可以存储在一个计算机可读取存储介质中。基于这样的理解,本发明实施例的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机、服务器、或者网络设备)执行本发明各个实施例所述方法的全部或部分。而前述的存储介质包括:移动存储设备、ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

[0316] 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。

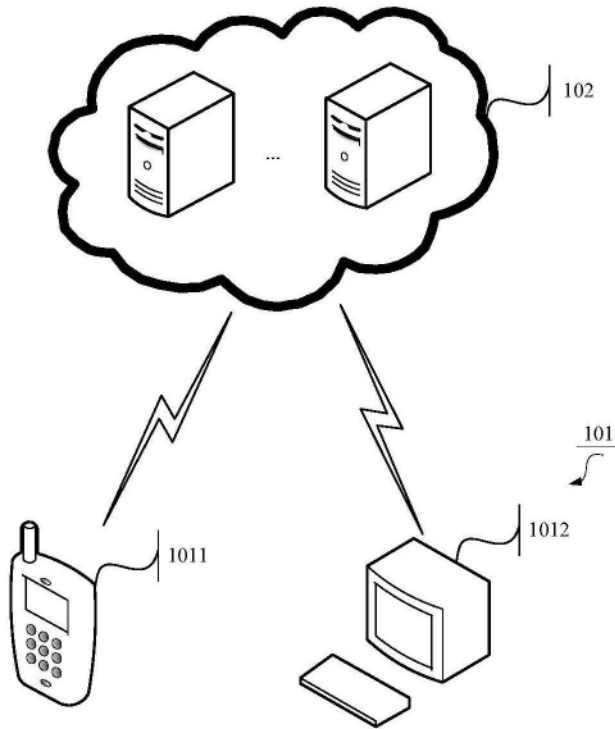


图1

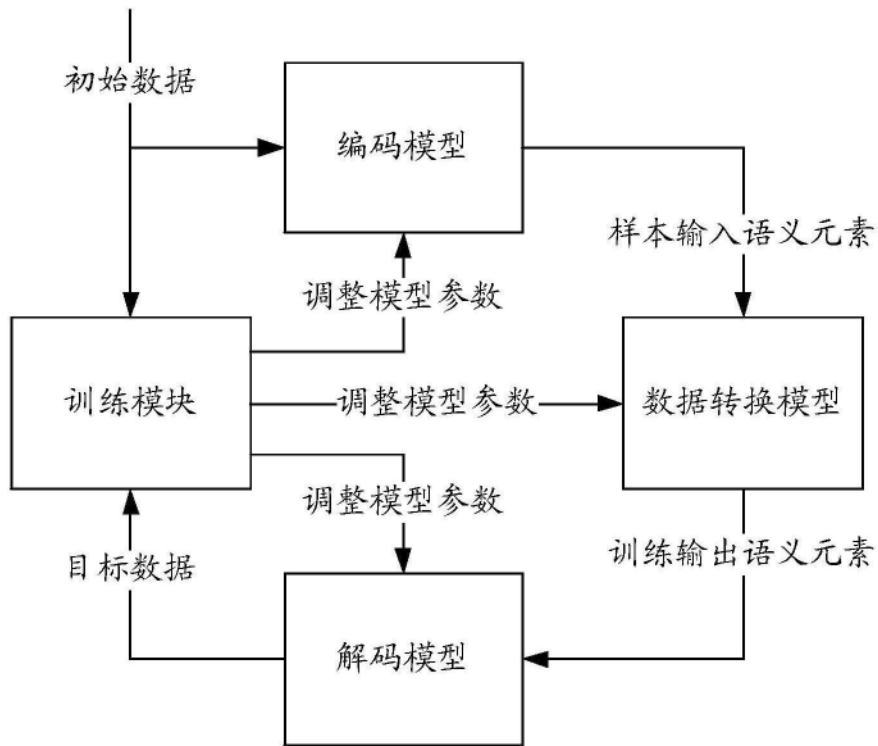


图2a

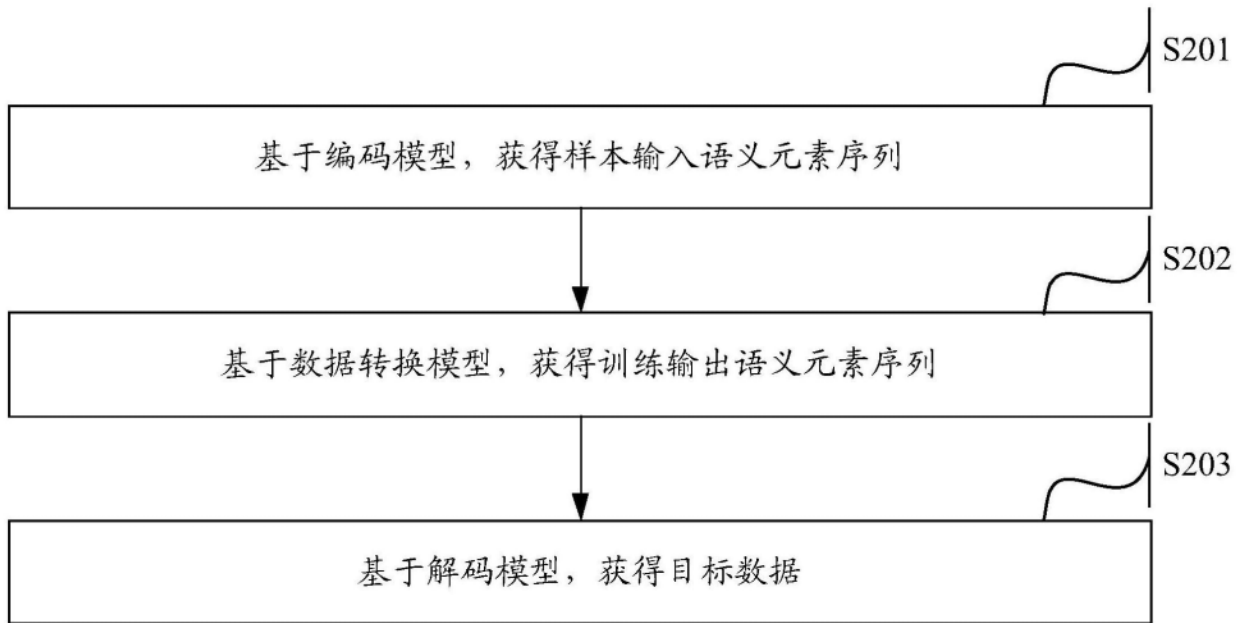


图2b

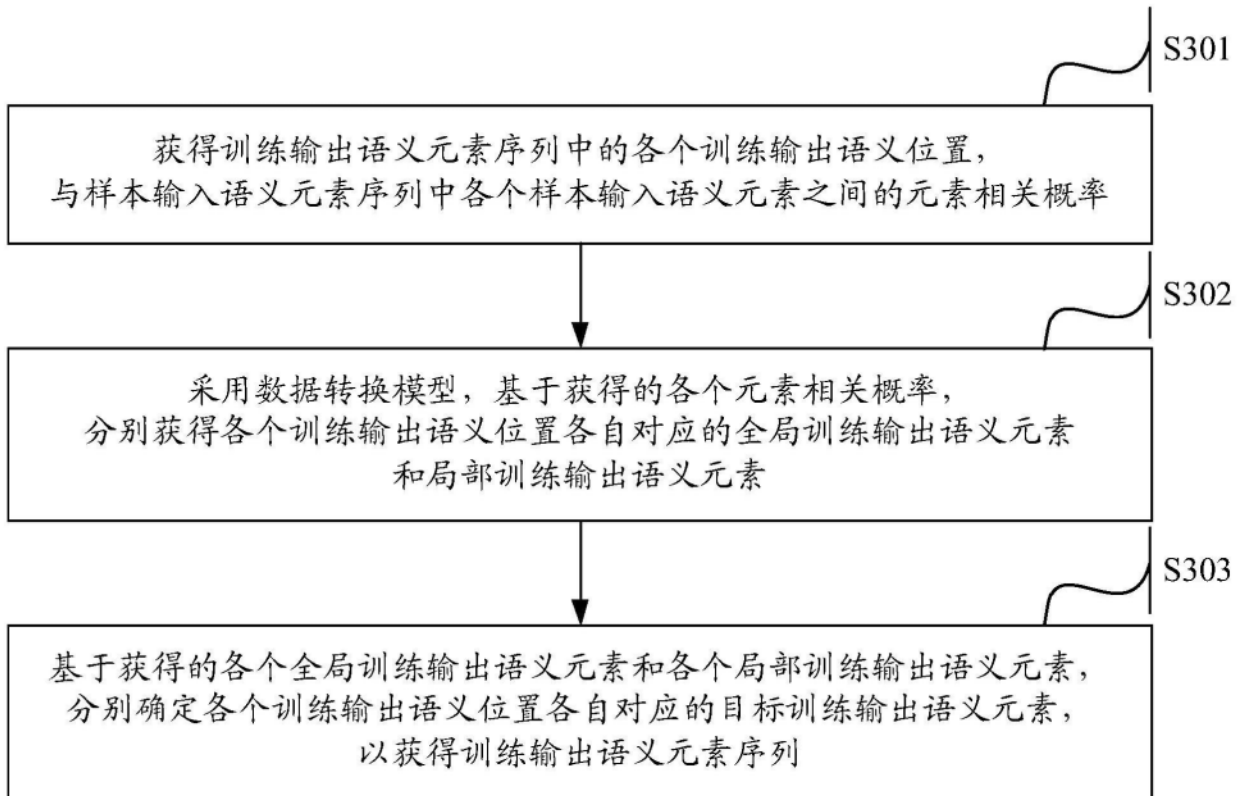


图3

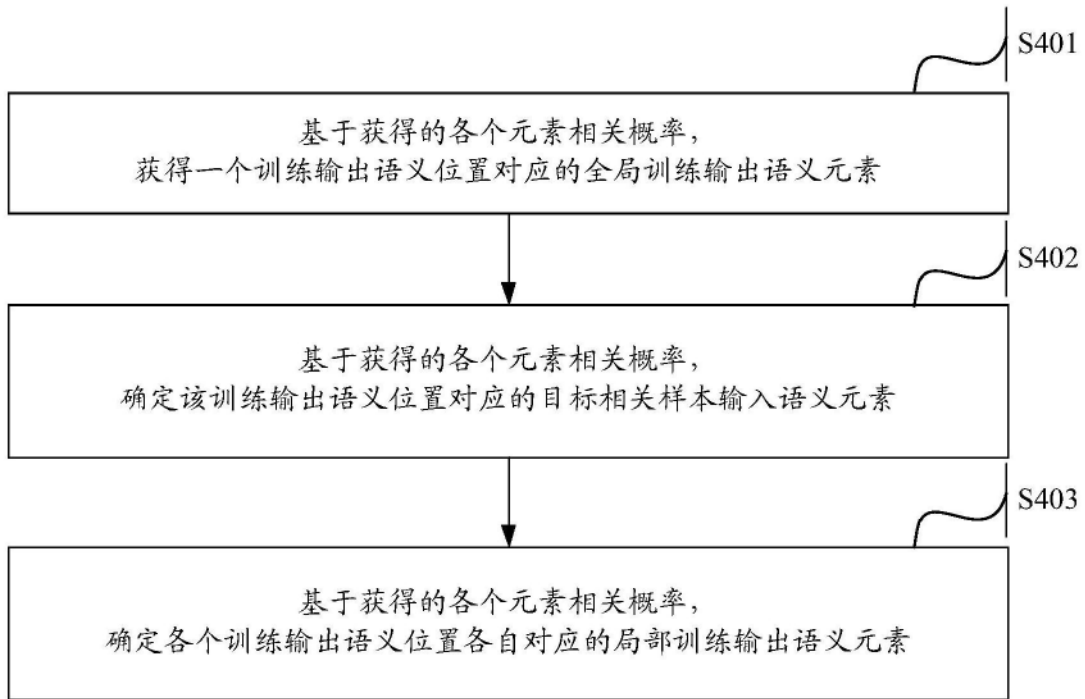


图4

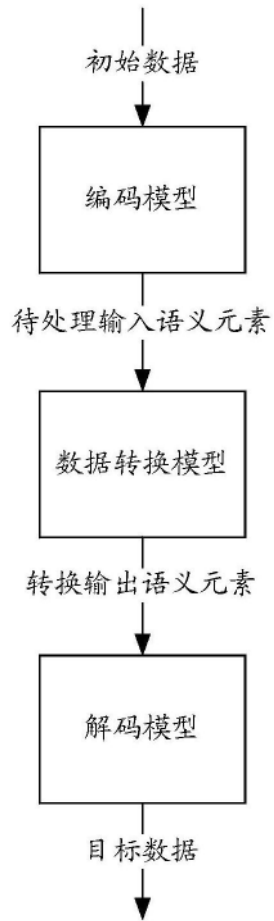


图5a

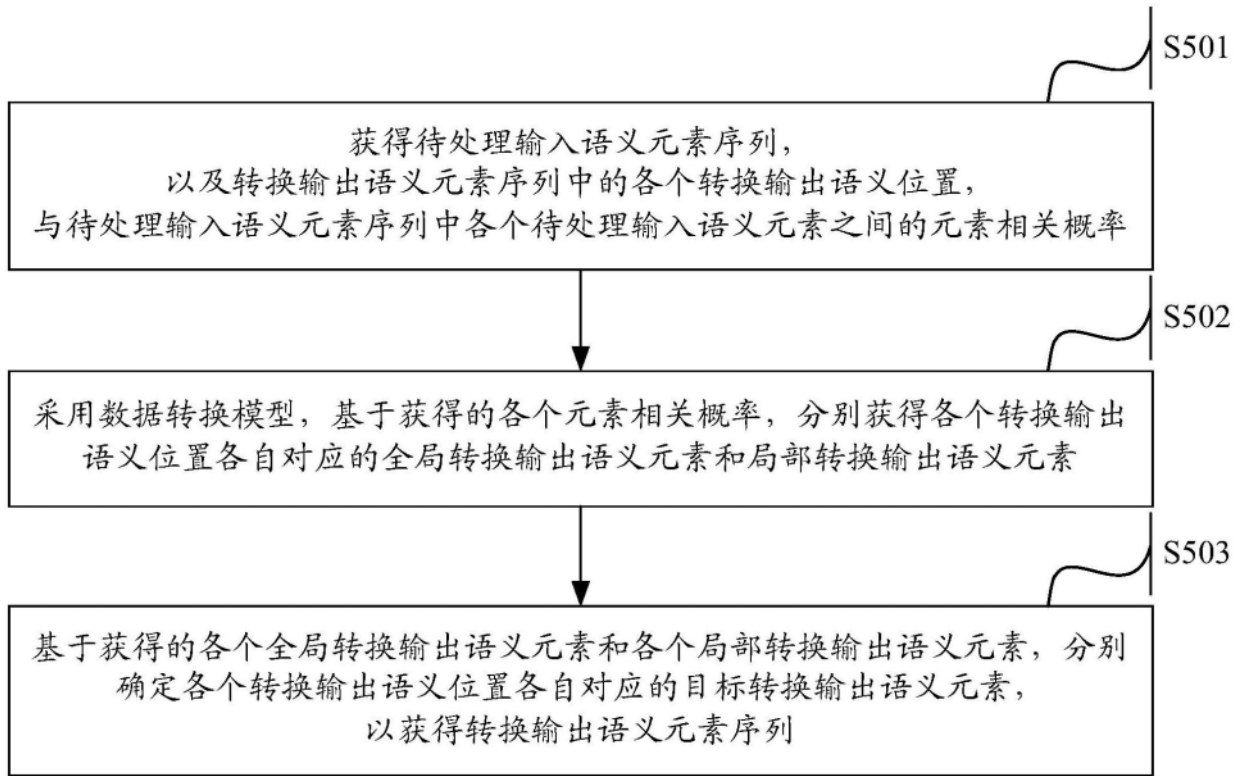


图5b

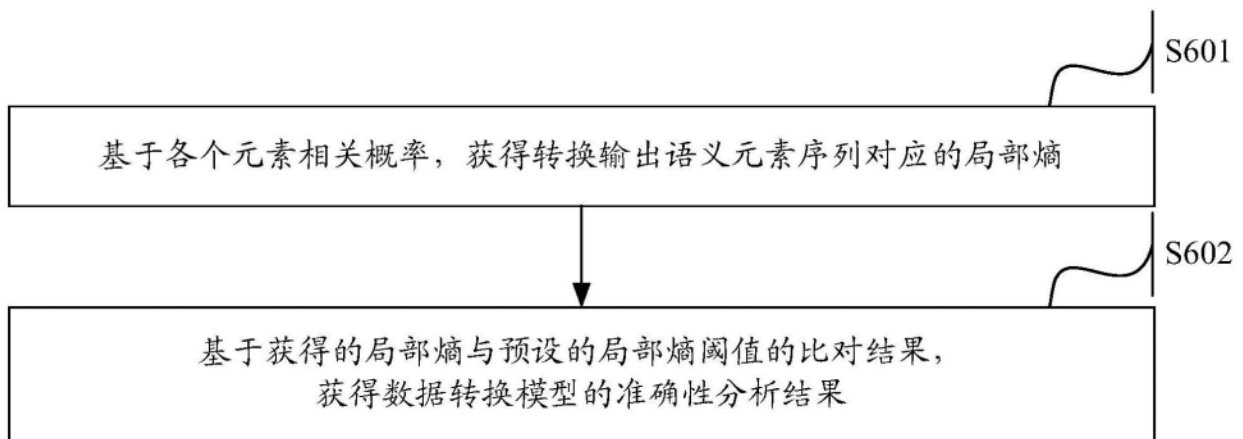


图6

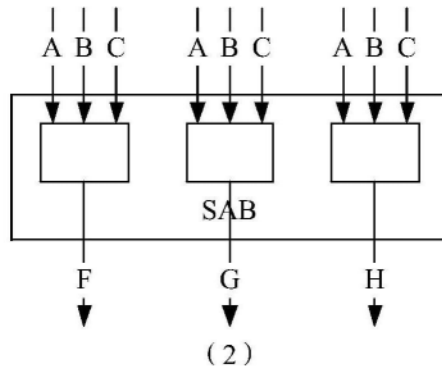
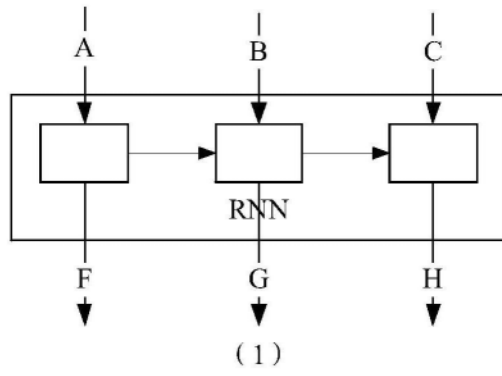


图7a

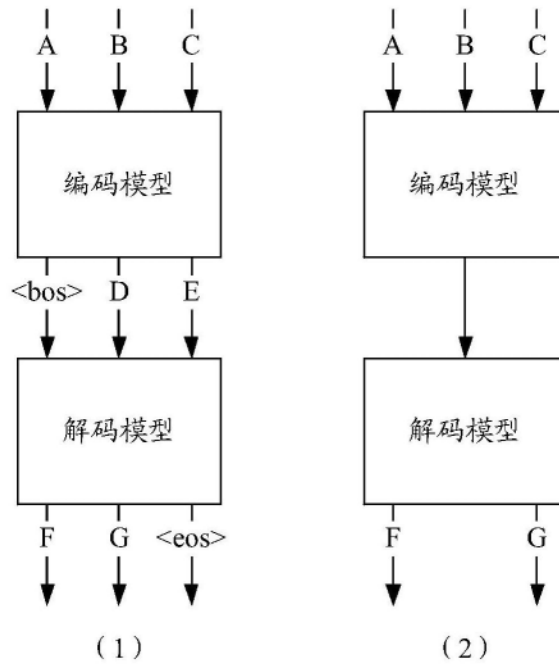


图7b

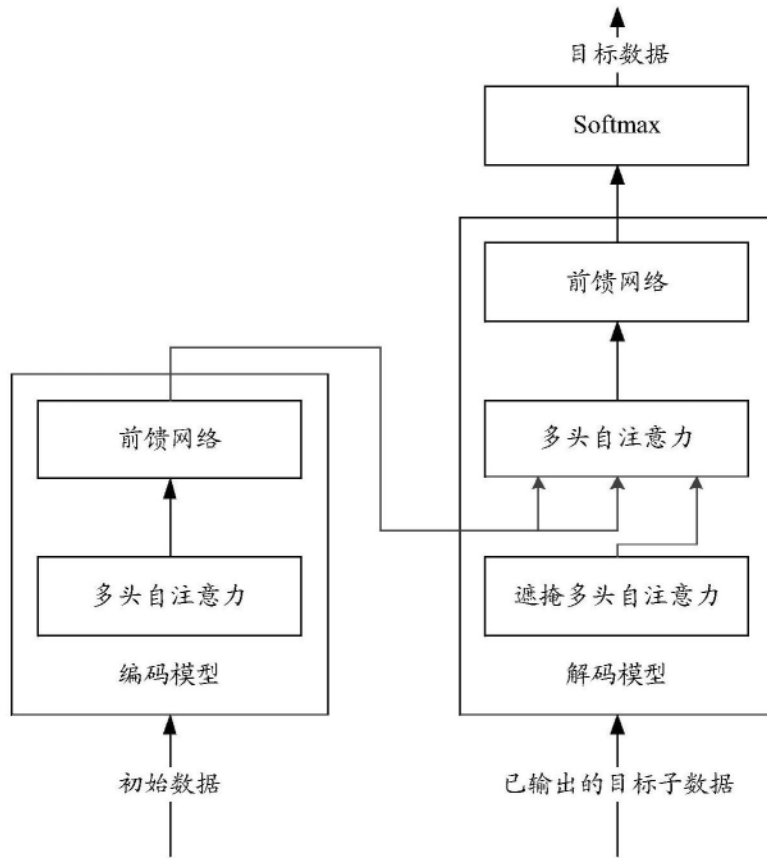


图7c

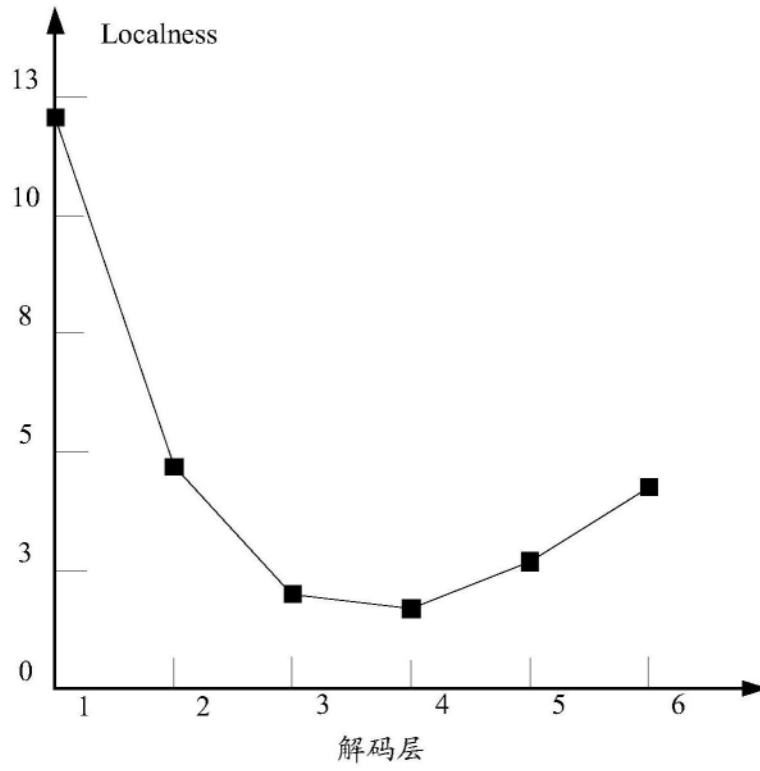


图8

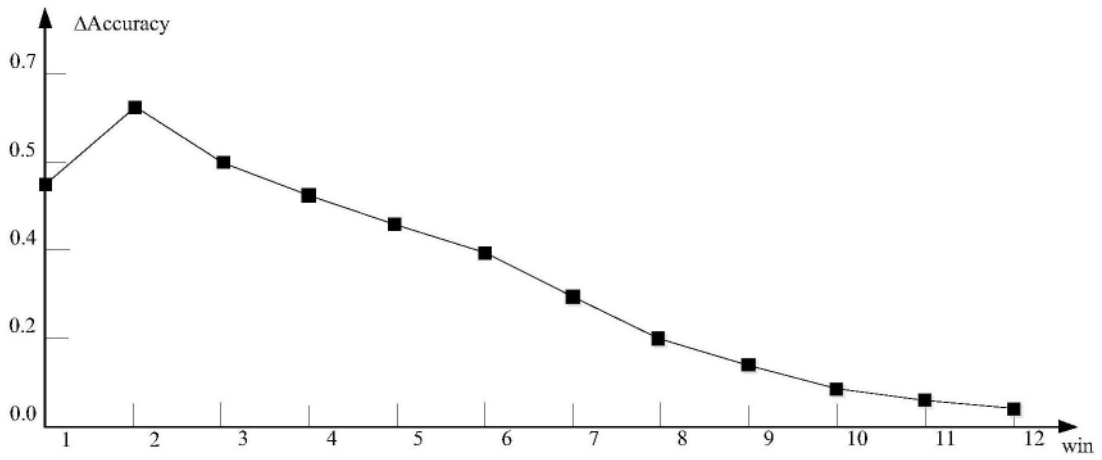


图9

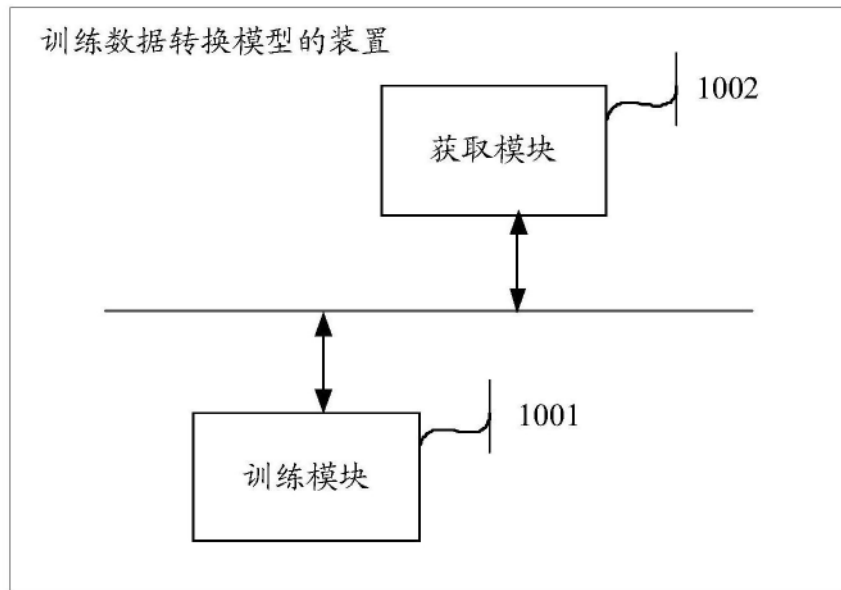


图10

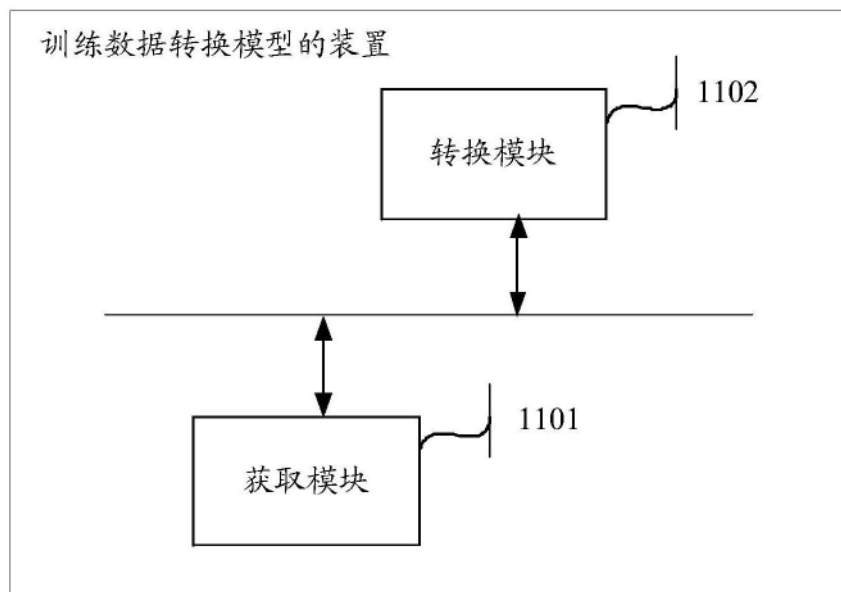


图11

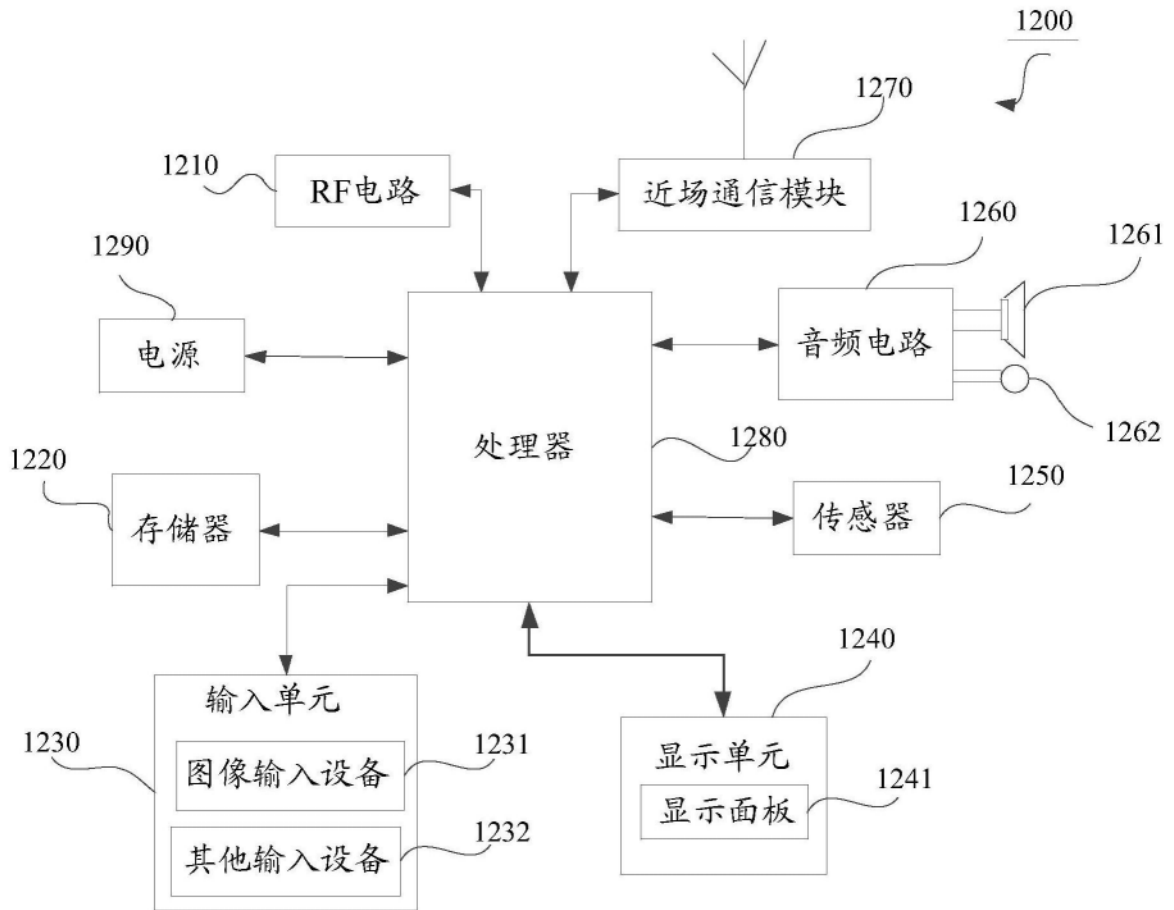


图12