



(21) 申请号 202011282073.6

(22) 申请日 2020.11.16

(65) 同一申请的已公布的文献号

申请公布号 CN 112347760 A

(43) 申请公布日 2021.02.09

(73) 专利权人 北京京东尚科信息技术有限公司

地址 100086 北京市海淀区知春路76号8层

专利权人 北京京东世纪贸易有限公司

(72) 发明人 乐雨泉 赵宇明 陈蒙

(74) 专利代理机构 中科专利商标代理有限责任

公司 11021

专利代理师 孙蕾

(51) Int. Cl.

G06F 40/211 (2020.01)

G06F 40/279 (2020.01)

(56) 对比文件

CN 110825866 A, 2020.02.21

CN 111159358 A, 2020.05.15

审查员 弓正

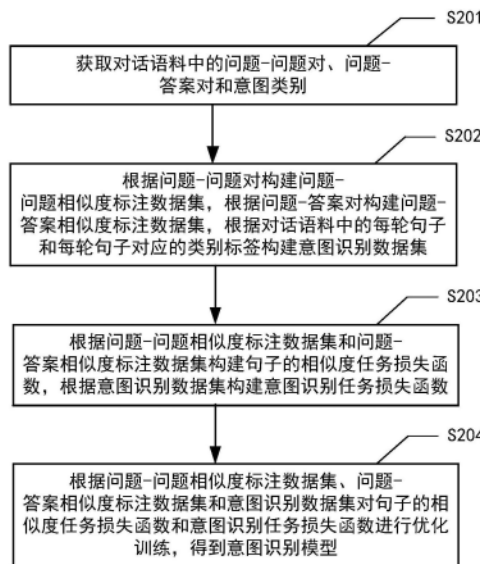
权利要求书3页 说明书14页 附图6页

(54) 发明名称

意图识别模型的训练方法及装置、意图识别方法及装置

(57) 摘要

本公开提供了一种意图识别模型的训练方法及装置、意图识别方法及装置。其中,训练方法包括:获取对话语料中的问题-问题对、问题-答案对和意图类别;根据问题-问题对构建问题-问题相似度标注数据集,根据问题-答案对构建问题-答案相似度标注数据集,根据对话语料中的每轮句子和每轮句子对应的类别标签构建意图识别数据集;根据问题-问题相似度标注数据集和问题-答案相似度标注数据集构建句子的相似度任务损失函数,根据意图识别数据集构建意图识别任务损失函数;以及根据问题-问题相似度标注数据集、问题-答案相似度标注数据集和意图识别数据集对句子的相似度任务损失函数和意图识别任务损失函数进行优化训练,得到意图识别模型。



1. 一种意图识别模型的训练方法,包括:

获取对话语料中的问题-问题对、问题-答案对和意图类别,其中,所述意图类别用于表征所述对话语料中的每轮句子的类别标签;

根据所述问题-问题对构建问题-问题相似度标注数据集;

根据所述问题-答案对构建问题-答案相似度标注数据集,包括:

获取两条问题-答案对: $(q1, a1)$ 和 $(q2, a2)$;

在 $f(q1, q2) + f(a1, a2) < k$ 的情况下,将所述两条问题-答案对中句子长度更长的问题-答案对的标签值配置为1,将所述两条问题-答案对中句子长度更短的问题-答案对的标签值配置为0,其中, $q1$ 表征以文本形式存储的第一问题, $a1$ 表征以文本形式存储的第一答案, $q2$ 表征以文本形式存储的第二问题, $a2$ 表征以文本形式存储的第二答案, f 为编辑距离函数, k 为过滤阈值;以及

根据所述问题-答案对及其标签构建所述问题-答案相似度标注数据集;

根据所述对话语料中的每轮句子和所述每轮句子对应的类别标签构建意图识别数据集;

根据所述问题-问题相似度标注数据集和所述问题-答案相似度标注数据集构建句子的相似度任务损失函数,根据所述意图识别数据集构建意图识别任务损失函数;以及

根据所述问题-问题相似度标注数据集、所述问题-答案相似度标注数据集和所述意图识别数据集对所述句子的相似度任务损失函数和所述意图识别任务损失函数进行优化训练,得到所述意图识别模型。

2. 根据权利要求1所述的方法,其中,所述获取对话语料中的问题-问题对、问题-答案对包括:

通过自动挖掘的方式从所述对话语料中获取所述问题-问题对和所述问题-答案对;

在所述问题-问题对中的两个问题对应的答案相同的情况下,将所述问题-问题对的标签值配置为1;以及

在所述问题-问题对中的两个问题对应的答案不同的情况下,将所述问题-问题对的标签值配置为0。

3. 根据权利要求2所述的方法,其中,根据所述问题-问题对构建问题-问题相似度标注数据集包括:

在所述问题-问题对中的两个问题的相似度小于预设阈值,且所述问题-问题对的标签值为1的情况下,变更所述问题-问题对的标签值至0或丢弃所述问题-问题对;

在所述问题-问题对中的两个问题的相似度大于预设阈值,且所述问题-问题对的标签值为0的情况下,变更所述问题-问题对的标签值至1或丢弃所述问题-问题对;以及

根据所述问题-问题对及其标签构建所述问题-问题相似度标注数据集。

4. 根据权利要求1所述的方法,其中,根据所述对话语料中的每轮句子和所述每轮句子对应的类别标签构建意图识别数据集包括:

获取包含有 n 轮句子的目标对话;

获取包含有 t 个窗口的滑窗,其中, $t < n$, t, n 为正整数;

通过所述滑窗对所述目标对话进行遍历,确定所述目标对话中的每轮句子对应的类别标签;以及

根据所述每轮句子和所述每轮句子对应的类别标签构建所述意图识别数据集。

5. 根据权利要求1所述的方法, 其中, 根据所述问题-问题相似度标注数据集和所述问题-答案相似度标注数据集构建句子的相似度任务损失函数为:

$$loss_{sim}(\theta) = -[p * \log(p_{pred}) + (1 - p) \log(1 - p_{pred})], \text{ 其中:}$$

$$p_{pred} = \text{sigmod}(\text{dense}_{sim}),$$

$$\text{dense}_{sim} = g(V_{cls}W_{sim} + b_{sim}),$$

其中, $loss_{sim}(\theta)$ 表示所述句子的相似度任务损失函数, θ 表示所述句子的相似度任务损失函数中的待优化参数, p 表示所述问题-问题对或所述问题-答案对的真实相似度标签, p_{pred} 表示所述问题-问题对或所述问题-答案对的预测相似度标签, V_{cls} 表示所述问题-问题对或所述问题-答案对, $W_{sim} \in R^{d_{cls} \times d_{sim}}$, d_{cls} 、 d_{sim} 表示矩阵R的维度, b_{sim} 表示偏置, g 代表 tanh 函数。

6. 根据权利要求1所述的方法, 其中, 根据所述意图识别数据集构建意图识别任务损失函数为:

$$loss_{intent}(\theta) = -[y * \log(y_{pred}) + (1 - y) \log(1 - y_{pred})], \text{ 其中:}$$

$$y_{pred} = \text{sigomd}(\text{dense}_{intent}),$$

$$\text{dense}_{intent} = g(V_{cls}W_{intent} + b_{intent}),$$

其中, $loss_{intent}(\theta)$ 表示所述意图识别任务损失函数, θ 表示所述意图识别任务损失函数中的待优化参数, y 表示所述对话中的句子对应的真实类别标签, y_{pred} 表示所述对话中的句子对应的预测类别标签, V_{cls} 表示所述对话中的句子, $W_{intent} \in R^{d_{cls} \times d_{intent}}$, d_{cls} 、 d_{intent} 表示矩阵R的维度, b_{intent} 表示偏置, g 代表 tanh 函数。

7. 根据权利要求1所述的方法, 其中, 根据所述问题-问题相似度标注数据集、所述问题-答案相似度标注数据集和所述意图识别数据集对所述句子的相似度任务损失函数和所述意图识别任务损失函数进行优化训练包括:

将所述问题-问题相似度标注数据集、所述问题-答案相似度标注数据集和所述意图识别数据集分别进行批量拆分, 得到多个子集合, 每个子集合存储有预设量的所述问题-问题相似度标注数据集、所述问题-答案相似度标注数据集或所述意图识别数据集;

将所述多个子集合混合合并至一个总集合; 以及

遍历所述总集合, 根据所述总集合中的每个所述子集合对所述句子的相似度任务损失函数或所述意图识别任务损失函数进行训练。

8. 一种意图识别方法, 包括:

获取意图识别模型; 以及

通过所述意图识别模型对待识别句子进行意图识别;

其中, 所述意图识别模型是通过如权利要求1至7中任一所述的训练方法训练得到。

9. 一种意图识别模型的训练装置, 包括:

第一获取模块, 用于获取对话语料中的问题-问题对、问题-答案对和意图类别, 其中, 所述意图类别用于表征所述对话语料中的每轮句子的类别标签;

数据集构建模块, 包括:

第一数据集构建单元, 用于根据所述问题-问题对构建问题-问题相似度标注数据集;

第二数据集构建单元, 用于根据所述问题-答案对构建问题-答案相似度标注数据集,

所述第二数据集构建单元包括：

第一获取子单元,用于获取两条问题-答案对: $(q1, a1)$ 和 $(q2, a2)$;

配置子单元,用于在 $f(q1, q2) + f(a1, a2) < k$ 的情况下,将所述两条问题-答案对中句子长度更长的问题-答案对的标签值配置为1,将所述两条问题-答案对中句子长度更短的问题-答案对的标签值配置为0,其中, $q1$ 表征以文本形式存储的第一问题, $a1$ 表征以文本形式存储的第一答案, $q2$ 表征以文本形式存储的第二问题, $a2$ 表征以文本形式存储的第二答案, f 为编辑距离函数, k 为过滤阈值;以及

第二数据集构建子单元,用于根据所述问题-答案对及其标签构建所述问题-答案相似度标注数据集;

第三数据集构建单元,用于根据所述对话语料中的每轮句子和所述每轮句子对应的类别标签构建意图识别数据集;

损失函数构建模块,包括:

第一损失函数构建单元,用于根据所述问题-问题相似度标注数据集和所述问题-答案相似度标注数据集构建句子的相似度任务损失函数;

第二损失函数构建单元,用于根据所述意图识别数据集构建意图识别任务损失函数;以及

训练模块,用于根据所述问题-问题相似度标注数据集、所述问题-答案相似度标注数据集和所述意图识别数据集对所述句子的相似度任务损失函数和所述意图识别任务损失函数进行优化训练,得到所述意图识别模型。

10.一种意图识别装置,包括:

第二获取模块,用于获取意图识别模型;以及

识别模块,用于通过所述意图识别模型对待识别句子进行意图识别;

其中,所述意图识别模型是通过如权利要求1至7中任一所述的训练方法训练得到。

11.一种计算机系统,包括:

一个或多个处理器;

存储器,用于存储一个或多个程序,

其中,当所述一个或多个程序被所述一个或多个处理器执行时,使得所述一个或多个处理器实现权利要求1至7或权利要求8中任一项所述的方法。

12.一种计算机可读存储介质,其上存储有可执行指令,该指令被处理器执行时使处理器实现权利要求1至7或权利要求8中任一项所述的方法。

意图识别模型的训练方法及装置、意图识别方法及装置

技术领域

[0001] 本公开涉及计算机技术领域,更具体地,涉及一种意图识别模型的训练方法及训练装置、意图识别方法及意图识别装置、以及计算机系统和计算机可读存储介质。

背景技术

[0002] 随着计算机、互联网技术的快速发展,越来越多的智能化交互方式已逐渐深入日常生活。

[0003] 智能化交互的一种体现在于,智能化产品或应用可以理解用户需求,并依据需求为用户提供适当的响应。例如在客服场景的问答系统中,可以根据用户输入的句子识别用户的意图,并为用户提供合适的回应。

[0004] 但在实现本公开构思的过程中,发明人发现相关技术中至少存在如下问题,通过单一的任务数据训练得到模型,导致模型学习的表征缺乏一定泛化能力,以至于模型的意图识别性能较弱。

发明内容

[0005] 有鉴于此,本公开提供了一种意图识别模型的训练方法及训练装置、意图识别方法及意图识别装置、以及计算机系统和计算机可读存储介质。

[0006] 本公开的一个方面提供了一种意图识别模型的训练方法,包括:获取对话语料中的问题-问题对、问题-答案对和意图类别,其中,所述意图类别用于表征所述对话语料中的每轮句子的类别标签;根据所述问题-问题对构建问题-问题相似度标注数据集,根据所述问题-答案对构建问题-答案相似度标注数据集,根据所述对话语料中的每轮句子和所述每轮句子对应的类别标签构建意图识别数据集;根据所述问题-问题相似度标注数据集和所述问题-答案相似度标注数据集构建句子的相似度任务损失函数,根据所述意图识别数据集构建意图识别任务损失函数;以及根据所述问题-问题相似度标注数据集、所述问题-答案相似度标注数据集和所述意图识别数据集对所述句子的相似度任务损失函数和所述意图识别任务损失函数进行优化训练,得到所述意图识别模型。

[0007] 本公开的另一个方面提供了一种意图识别方法,包括:获取意图识别模型;以及通过所述意图识别模型对待识别句子进行意图识别;其中,所述意图识别模型是通过如上所述的训练方法训练得到。

[0008] 本公开的另一个方面提供了一种意图识别模型的训练装置,包括:第一获取模块,用于获取对话语料中的问题-问题对、问题-答案对和意图类别,其中,所述意图类别用于表征所述对话语料中的每轮句子的类别标签;数据集构建模块,包括:第一数据集构建单元,用于根据所述问题-问题对构建问题-问题相似度标注数据集;第二数据集构建单元,用于根据所述问题-答案对构建问题-答案相似度标注数据集;第三数据集构建单元,用于根据所述对话语料中的每轮句子和所述每轮句子对应的类别标签构建意图识别数据集;损失函数构建模块,包括:第一损失函数构建单元,用于根据所述问题-问题相似度标注数据集和

所述问题-答案相似度标注数据集构建句子的相似度任务损失函数;第二损失函数构建单元,用于根据所述意图识别数据集构建意图识别任务损失函数;以及训练模块,用于根据所述问题-问题相似度标注数据集、所述问题-答案相似度标注数据集和所述意图识别数据集对所述句子的相似度任务损失函数和所述意图识别任务损失函数进行优化训练,得到所述意图识别模型。

[0009] 本公开的另一个方面提供了一种意图识别装置,包括:第二获取模块,用于获取意图识别模型;以及识别模块,用于通过所述意图识别模型对待识别句子进行意图识别;其中,所述意图识别模型是通过如上所述的训练方法训练得到。

[0010] 本公开的另一方面提供了一种计算机系统,包括:一个或多个处理器;存储器,用于存储一个或多个程序,其中,当所述一个或多个程序被所述一个或多个处理器执行时,使得所述一个或多个处理器实现如上所述的意图识别模型的训练方法或意图识别方法。

[0011] 本公开的另一方面提供了一种计算机可读存储介质,其上存储有计算机可执行指令,所述指令在被执行时用于实现如上所述的意图识别模型的训练方法或意图识别方法。

[0012] 本公开的另一方面提供了一种计算机程序,所述计算机程序包括计算机可执行指令,所述指令在被执行时用于实现如上所述的意图识别模型的训练方法或意图识别方法。

[0013] 根据本公开的实施例,通过采用了获取对话语料中的问题-问题对、问题-答案对和意图类别;根据问题-问题对构建问题-问题相似度标注数据集,根据问题-答案对构建问题-答案相似度标注数据集,根据对话语料中的每轮句子和每轮句子对应的类别标签构建意图识别数据集;根据问题-问题相似度标注数据集和问题-答案相似度标注数据集构建句子的相似度任务损失函数,根据意图识别数据集构建意图识别任务损失函数;以及根据问题-问题相似度标注数据集、问题-答案相似度标注数据集和意图识别数据集对句子的相似度任务损失函数和意图识别任务损失函数进行优化训练,得到意图识别模型的技术手段,由于意图识别模型为根据问题-问题对、问题-答案对和意图类别三种任务数据联合训练得到,能够有效地利用客服场景中的多种任务语料,所以至少部分地克服了模型的训练数据单一的技术问题,进而达到了模型学习的表征更具泛化能力,并可有效增强意图识别性能的技术效果。

附图说明

[0014] 通过以下参照附图对本公开实施例的描述,本公开的上述以及其他目的、特征和优点将更为清楚,在附图中:

[0015] 图1示意性示出了根据本公开实施例的可以应用意图识别模型的训练方法的示例性系统架构;

[0016] 图2示意性示出了根据本公开实施例的意图识别模型的训练方法的流程图;

[0017] 图3示意性示出了根据本公开实施例的基于BERT多任务学习的意图识别模型的构建流程;

[0018] 图4示意性示出了根据本公开实施例的基于BERT多任务学习的意图识别模型的训练流程;

[0019] 图5示意性示出了根据本公开的实施例的意图识别模型的训练装置的框图;

[0020] 图6示意性示出了根据本公开的实施例的意图识别装置的框图;以及

[0021] 图7示意性示出了根据本公开实施例的适于实现意图识别模型的训练方法或意图识别方法的计算机系统的框图。

具体实施方式

[0022] 以下,将参照附图来描述本公开的实施例。但是应该理解,这些描述只是示例性的,而并非要限制本公开的范围。在下面的详细描述中,为便于解释,阐述了许多具体的细节以提供对本公开实施例的全面理解。然而,明显地,一个或多个实施例在没有这些具体细节的情况下也可以被实施。此外,在以下说明中,省略了对公知结构和技术的描述,以避免不必要地混淆本公开的概念。

[0023] 在此使用的术语仅仅是为了描述具体实施例,而并非意在限制本公开。在此使用的术语“包括”、“包含”等表明了所述特征、步骤、操作和/或部件的存在,但是并不排除存在或添加一个或多个其他特征、步骤、操作或部件。

[0024] 在此使用的所有术语(包括技术和科学术语)具有本领域技术人员通常所理解的含义,除非另外定义。应注意,这里使用的术语应解释为具有与本说明书的上下文相一致的含义,而不应以理想化或过于刻板的方式来解释。

[0025] 在使用类似于“A、B和C等中至少一个”这样的表述的情况下,一般来说应该按照本领域技术人员通常理解该表述的含义来予以解释(例如,“具有A、B和C中至少一个的系统”应包括但不限于单独具有A、单独具有B、单独具有C、具有A和B、具有A和C、具有B和C、和/或具有A、B、C的系统等)。在使用类似于“A、B或C等中至少一个”这样的表述的情况下,一般来说应该按照本领域技术人员通常理解该表述的含义来予以解释(例如,“具有A、B或C中至少一个的系统”应包括但不限于单独具有A、单独具有B、单独具有C、具有A和B、具有A和C、具有B和C、和/或具有A、B、C的系统等)。

[0026] 随着时代和技术的发展,客服场景中问答系统在日常生活中发挥着越来越重要的作用,而意图识别是问答系统中的核心技术。意图识别旨在识别对话系统中用户输入文本的意图,为系统下一步操作提供指导,精确的理解用户意图能够有效地提高用户对问答系统的体验。

[0027] 目前意图识别方法通常将其看作一个句子分类任务,主要分为基于机器学习的方法、基于神经网络的方法。基于机器学习的方法主要分为二步:首先提取句子的特征;然后送入机器学习分类模型。其中常见的特征比如TF-IDF(词频-逆文本频率指数)、One-hot(独热编码)等,常见的机器学习算法包含SVM(支持向量机)、贝叶斯等。由于深度学习的发展,开始利用神经网络搭建模型自动挖掘句子特征,其中比较有名的网络例如包含TextCNN(文本分类模型)、LSTM(Long Short-Term Memory,长短期记忆网络)、HAN、BERT(Bidirectional Encoder Representation from Transformers,一种语言表示模型)等。

[0028] 但在实现本公开构思的过程中,发明人发现,相关技术中的意图识别模型未能有效地利用客服场景中其它任务语料,利用孤立任务单独训练,导致学习的表征缺乏一定泛化能力。

[0029] 发明人还发现,基于机器学习的意图识别通过挖掘各种词法、句法特征,然后结合机器学习方法,该类方法需要大量的、繁琐的人工工作去提取语句特征,但是人工很难提取到全面的、有效地特征。虽然随着深度学习的发展,涌现出一些通过CNN(Convolutional

Neural Network,卷积神经网络)、RNN(Recurrent Neural Network,循环神经网络)、LSTM等神经网络的方法,但是基于深度学习的方法也需要海量意图标注语料来训练调优,而实际上电商场景中预先定义的意图类别十分庞大,动辄几十上百意图类别,使得标注人员在标注过程中十分吃力,因此难以获得大量高质量的意图识别标注语料。

[0030] 本公开的实施例提供了一种意图识别模型的训练方法及训练装置、意图识别方法及意图识别装置、以及计算机系统和计算机可读存储介质。其中,意图识别模型的训练方法包括:获取对话语料中的问题-问题对、问题-答案对和意图类别;根据问题-问题对构建问题-问题相似度标注数据集,根据问题-答案对构建问题-答案相似度标注数据集,根据对话语料中的每轮句子和每轮句子对应的类别标签构建意图识别数据集;根据问题-问题相似度标注数据集和问题-答案相似度标注数据集构建句子的相似度任务损失函数,根据意图识别数据集构建意图识别任务损失函数;以及根据问题-问题相似度标注数据集、问题-答案相似度标注数据集和意图识别数据集对句子的相似度任务损失函数和意图识别任务损失函数进行优化训练,得到意图识别模型。

[0031] 图1示意性示出了根据本公开实施例的可以应用意图识别模型的训练方法的示例性系统架构100。需要注意的是,图1所示仅为可以应用本公开实施例的系统架构的示例,以帮助本领域技术人员理解本公开的技术内容,但并不意味着本公开实施例不可以用于其他设备、系统、环境或场景。

[0032] 如图1所示,根据该实施例的系统架构100可以包括终端设备101、102、103,网络104和服务器105。网络104用以在终端设备101、102、103和服务器105之间提供通信链路的介质。网络104可以包括各种连接类型,例如有线和/或无线通信链路等等。

[0033] 用户可以使用终端设备101、102、103通过网络104与服务器105交互,以接收或发送消息等。终端设备101、102、103上可以安装有各种通讯客户端应用,例如购物类应用、网页浏览器应用、搜索类应用、即时通信工具、邮箱客户端和/或社交平台软件等。

[0034] 终端设备101、102、103可以是具有显示屏并且支持网页浏览的各种电子设备,包括但不限于智能手机、平板电脑、膝上型便携计算机和台式计算机等等。

[0035] 服务器105可以是提供各种服务的服务器,例如对用户利用终端设备101、102、103所开启的对话提供意图识别支持的后台管理服务器。后台管理服务器可以对接收到的用户输入语句等数据进行分析等处理,并将处理结果(例如根据用户输入语句获取或生成针对该用户输入语句的系统输出数据等)反馈给终端设备。

[0036] 需要说明的是,本公开实施例所提供的意图识别模型的训练方法一般可以由服务器105执行。相应地,本公开实施例所提供的意图识别模型的训练装置一般可以设置于服务器105中。本公开实施例所提供的意图识别模型的训练方法也可以由不同于服务器105且能够与终端设备101、102、103和/或服务器105通信的服务器或服务器集群执行。相应地,本公开实施例所提供的意图识别模型的训练装置也可以设置于不同于服务器105且能够与终端设备101、102、103和/或服务器105通信的服务器或服务器集群中。或者,本公开实施例所提供的意图识别模型的训练方法也可以由终端设备101、102、或103执行,或者也可以由不同于终端设备101、102、或103的其他终端设备执行。相应地,本公开实施例所提供的意图识别模型的训练装置也可以设置于终端设备101、102、或103中,或设置于不同于终端设备101、102、或103的其他终端设备中。

[0037] 例如,对话语料可以原本存储在终端设备101、102、或103中的任意一个(例如,终端设备101,但不限于此)之中,或者存储在外部存储设备上并可以导入到终端设备101中。然后,终端设备101可以在本地执行本公开实施例所提供的意图识别模型的训练方法,或者将对话语料发送到其他终端设备、服务器、或服务器集群,并由接收该对话语料的其他终端设备、服务器、或服务器集群来执行本公开实施例所提供的意图识别模型的训练方法。

[0038] 应该理解,图1中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需要,可以具有任意数目的终端设备、网络和服务器的。

[0039] 图2示意性示出了根据本公开实施例的意图识别模型的训练方法的流程图。

[0040] 如图2所示,该方法包括操作S201~S204。

[0041] 在操作S201,获取对话语料中的问题-问题对、问题-答案对和意图类别,其中,意图类别用于表征对话语料中的每轮句子的类别标签。

[0042] 根据本公开的实施例,首先收集对话语料,对话语料例如可以为来自客服场景的真实海量数据,该数据例如可以包括用户输入的句子、客服输出的句子以及系统输入或输出的句子等所有句子的集合。问题-问题对例如可以为对话语料中包含较多相似或相同词汇的两个句子或答案相同的两个句子。问题-答案对例如可以为分别来自用户和客服(或系统)的具有相似或相同词汇的两个句子,在该种情况下,该两个句子可以相邻或不相邻,两个句子可以来自同一段对话或不同段对话,用户和客服(或系统)可以是同一段对话或不同段对话中的用户和客服(或系统),问题-答案对例如还可以为在一定时间间隔内相邻的且分别来自用户和客服(或系统)的两个句子。意图类别例如可以包括产品信息类、物流运输类、地址类、时间类和和其他推荐类等多种类别,用于表征一句或一段句子的用户意图。

[0043] 在操作S202,根据问题-问题对构建问题-问题相似度标注数据集,根据问题-答案对构建问题-答案相似度标注数据集,根据对话语料中的每轮句子和每轮句子对应的类别标签构建意图识别数据集。

[0044] 根据本公开的实施例,在获取到上述问题-问题对、问题-答案对和意图类别等多个任务后,由于不同任务的数据结构并不完全相同,且不同任务的相似度计算方式完全不同,因此针对每一个任务构建与该任务相对应的数据集,以便可以根据该些不同类型的数据集训练得到相应的损失函数。

[0045] 在操作S203,根据问题-问题相似度标注数据集和问题-答案相似度标注数据集构建句子的相似度任务损失函数,根据意图识别数据集构建意图识别任务损失函数。

[0046] 根据本公开的实施例,由于问题-问题对、问题-答案对均是由两个句子构成的成对的数据结构,可以具有相似的值,而意图类别通常表征一个句子和类别标签,与前两者(问题-问题对、问题-答案对)的参数值存在差异,因此可以针对前两者(问题-问题对、问题-答案对)构建相同的句子的相似度任务损失函数,针对后一者(意图类别)构建单独的意图识别任务损失函数,从而可以根据两个函数和不同的训练数据集训练得到用于表征上述意图识别模型的目标损失函数,该目标损失函数例如可以包括上述句子的相似度任务损失函数和上述意图识别任务损失函数。

[0047] 在操作S204,根据问题-问题相似度标注数据集、问题-答案相似度标注数据集和意图识别数据集对句子的相似度任务损失函数和意图识别任务损失函数进行优化训练,得到意图识别模型。

[0048] 根据本公开的实施例,通过联合问题-问题相似度匹配任务(问题-问题对)、问题-答案相似度匹配任务(问题-答案对)和意图识别任务(意图类别),进一步构建数据集并形成上述目标损失函数,最终训练得到上述意图识别模型。

[0049] 通过本公开的上述实施例,由于采用多任务学习方式联合问题-问题相似度匹配任务、问题-答案相似度匹配任务和意图识别任务,可以有效地将这些孤立任务联系在一起。同时,结合问题-问题相似度匹配任务侧重学习用户句子之间的匹配相似性这一优势,可以有效提高意图分类的泛化性;并结合问题-答案相似度匹配任务侧重学习用户句子和回答答案之间的匹配度这一优势,可以有效提高意图识别的准确性。本公开的上述实施例,通过联合三个任务同时训练,共享参数和特征,可以有效地增强意图识别性能。

[0050] 根据本公开的实施例,上述意图识别模型例如可以是基于BERT多任务学习模型训练并完成构建。

[0051] 图3示意性示出了根据本公开实施例的基于BERT多任务学习的意图识别模型的构建流程。

[0052] 图4示意性示出了根据本公开实施例的基于BERT多任务学习的意图识别模型的训练流程。

[0053] 下面参考图3和图4,结合具体实施例对图2所示的方法做进一步说明。

[0054] 根据本公开的实施例,上述操作S201中的获取对话语料中的问题-问题对、问题-答案对包括:通过自动挖掘的方式从对话语料中获取问题-问题对和问题-答案对;在问题-问题对中的两个问题对应的答案相同的情况下,将问题-问题对的标签值配置为1;以及在问题-问题对中的两个问题对应的答案不同的情况下,将问题-问题对的标签值配置为0。

[0055] 根据本公开的一具体实施例,如图3所示,在需要获取问题-问题对的情况下,通过自动挖掘的方式从海量对话语料中挖掘出问题-问题对,挖掘准则为:若问题a与问题b的答案相同,则问题a与问题b相似,标签为1;若问题a与问题的答案不同,则问题a与问题b不相似,标签为0。在需要获取问题-答案对的情况下,通过自动挖掘的方式从海量对话语料中挖掘出问题-答案对,挖掘准则为:若该问题后面有回复,则抽取该问题-答案添加到问题-答案对候选集。

[0056] 根据本公开的实施例,上述操作S202中的根据问题-问题对构建问题-问题相似度标注数据集包括:在问题-问题对中的两个问题的相似度小于预设阈值,且问题-问题对的标签值为1的情况下,变更问题-问题对的标签值至0或丢弃问题-问题对;在问题-问题对中的两个问题的相似度大于预设阈值,且问题-问题对的标签值为0的情况下,变更问题-问题对的标签值至1或丢弃问题-问题对;以及根据问题-问题对及其标签构建问题-问题相似度标注数据集。

[0057] 根据本公开的一具体实施例,再请参见图3所示,在构建问题-问题相似度标注数据集的过程中,需要结合TF-IDF等方法筛选噪声对并人工标注,该过程例如可以包括:

[0058] a.通过TF-IDF、word embedding(一种将文本数据转换为数值型数据的方式)等方法计算问题-问题对的相似度,筛选出可能的噪声对,进行人工标注,筛选规则为:若问题a与问题b的相似度小于阈值p并且标签为1,则问题a和问题b可能相似度低,需要人工校正;若问题a与问题b的相似度大于阈值p并且标签为0,该种情况下问题a和问题b可能相似度高,也需要人工校正。

[0059] b.对于筛选出来的噪声对,人工进行标注,并对筛选出的语料进行预处理,包括移除语义无关的标点符号,得到问题-问题相似度标注数据集。

[0060] 需要说明的是,其中的人工校正例如可以为修改噪声对的标志值或者丢弃噪声对。

[0061] 根据本公开的实施例,上述操作S202中的根据问题-答案对构建问题-答案相似度标注数据集包括:获取两条问题-答案对: (q_1, a_1) 和 (q_2, a_2) ;在 $f(q_1, q_2) + f(a_1, a_2) < k$ 的情况下,将两条问题-答案对中句子长度更长的问题-答案对的标签值配置为1,将两条问题-答案对中句子长度更短的问题-答案对的标签值配置为0,其中, q_1 表征以文本形式存储的第一问题, a_1 表征以文本形式存储的第一答案, q_2 表征以文本形式存储的第二问题, a_2 表征以文本形式存储的第二答案, f 为编辑距离函数, k 为过滤阈值;以及根据问题-答案对及其标签构建问题-答案相似度标注数据集。

[0062] 根据本公开的一具体实施例,再请参见图3所示,在构建问题-答案相似度标注数据集的过程中,需要编辑距离过滤冗余样本,该过程例如可以包括:

[0063] 根据编辑距离过滤冗余的问题-答案对,过滤方法为:若两条问题-答案对为 (q_1, a_1) 和 (q_2, a_2) ,若 $f(q_1, q_2) + f(a_1, a_2) < k$,则表明该两条问题-答案对的样本出现冗余,在该情况下,保留两条问题-答案对中句子较长(标签为1)的样本,并通过负采样获取两条问题-答案对中句子较短(标签为0)的样本,得到问题-答案相似度标注数据集。

[0064] 根据本公开的实施例,上述操作S202中的根据对话语料中的每轮句子和每轮句子对应的类别标签构建意图识别数据集包括:获取包含有 n 轮句子的目标对话;获取包含有 t 个窗口的滑窗,其中, $t < n$, t, n 为正整数;通过滑窗对目标对话进行遍历,确定目标对话中的每轮句子对应的类别标签;以及根据每轮句子和每轮句子对应的类别标签构建意图识别数据集。

[0065] 根据本公开的一具体实施例,再请参见图3所示,在构建意图识别数据集的过程中,需要人工标注,该过程例如可以包括:

[0066] a.预先定义 M 个电商场景意图类别。

[0067] b.真实场景下的对话一般包含多轮句子,假设某一轮对话句子列表为: $\{sen_1, sen_2, \dots, sen_n\}$,其中 n 表示一共包含 n 句对话。

[0068] c.从左至右,按照滑窗为 t 的方式构建待标注意图识别语料: $(sen_1, sen_2, \dots, sen_t)$, $(sen_2, sen_3, \dots, sen_{t+1})$, $\dots$, $(sen_{n-t}, sen_{n-t+1}, \dots, sen_n)$

[0069] d.对于c中获得的待标注意图识别语料进行人工标注,标注准则为:结合上文语句,判断最后一句话属于 M 个意图其中一个。

[0070] e.根据句子和其所属意图构建意图识别数据集。

[0071] 需要说明的是,上述人工标注的过程均可直接替换为具有相应功能或准则的程序指令。

[0072] 通过本公开的上述实施例,由于采用自动挖掘的方式从海量对话语料中挖掘出问题-问题、问题-答案对,并结合相似度算法过滤得到高质量的样本对,可以有效减少人工标注数据的时间、快速获取大量相似度标注语料。

[0073] 根据本公开的实施例,上述操作S203中的根据问题-问题相似度标注数据集和问题-答案相似度标注数据集构建句子的相似度任务损失函数为:

[0074] $\text{loss}_{\text{sim}}(\theta) = -[p \cdot \log(p_{\text{pred}}) + (1-p) \log(1-p_{\text{pred}})]$, 其中:

[0075] $p_{\text{pred}} = \text{sigmod}(\text{dense}_{\text{sim}})$,

[0076] $\text{dense}_{\text{sim}} = g(V_{\text{cls}}W_{\text{sim}} + b_{\text{sim}})$,

[0077] 其中, $\text{loss}_{\text{sim}}(\theta)$ 表示句子的相似度任务损失函数, θ 表示句子的相似度任务损失函数中的待优化参数, p 表示问题-问题对或问题-答案对的真实相似度标签, p_{pred} 表示问题-问题对或问题-答案对的预测相似度标签, V_{cls} 表示问题-问题对或问题-答案对, $W_{\text{sim}} \in R^{d_{\text{cls}} \times d_{\text{sim}}}$, d_{cls} 、 d_{sim} 表示矩阵 R 的维度, b_{sim} 表示偏置, g 代表 $\tanh$ 函数。

[0078] 根据本公开的一具体实施例, 参见图4所示, 上述构建问题-问题相似度任务损失函数、问题-答案相似度任务损失函数的过程例如可以包括:

[0079] a. 假设 $(\text{sen}_1, \text{sen}_2)$ 和 p 为句子对 (问题-问题或者问题-答案对) 和相似度标签, 将句子对拼接得到句子对字序列: $X = \{x_1, x_2, \dots, x_t\}$, 作为图4所示模型的输入, 其中, $x_i \in W$, t 表示输入字序列长度, x_i 表示输入句子对中的某个字, W 表示整个字表大小。

[0080] b. 通过图4所示基于BERT多任务学习的意图识别模型得到句子对的表示 V_{cls} , 其维度为 d_{cls} 。

[0081] c. 计算相似度任务预测结果 p_{pred} 。

[0082] d. 采用交叉熵构建并输出问题-问题相似度任务损失函数和问题-答案相似度任务损失函数, 即上述句子的相似度任务损失函数 $\text{loss}_{\text{sim}}(\theta)$ 。

[0083] 根据本公开的实施例, 上述操作S203中的根据意图识别数据集构建意图识别任务损失函数为:

[0084] $\text{loss}_{\text{intent}}(\theta) = -[y \cdot \log(y_{\text{pred}}) + (1-y) \log(1-y_{\text{pred}})]$, 其中:

[0085] $y_{\text{pred}} = \text{sigomd}(\text{dense}_{\text{intent}})$,

[0086] $\text{dense}_{\text{intent}} = g(V_{\text{cls}}W_{\text{intent}} + b_{\text{intent}})$,

[0087] 其中, $\text{loss}_{\text{intent}}(\theta)$ 表示意图识别任务损失函数, θ 表示意图识别任务损失函数中的待优化参数, y 表示对话中的句子对应的真实类别标签, y_{pred} 表示对话中的句子对应的预测类别标签, V_{cls} 表示对话中的句子, $W_{\text{intent}} \in R^{d_{\text{cls}} \times d_{\text{intent}}}$, d_{cls} 、 d_{intent} 表示矩阵 R 的维度, b_{intent} 表示偏置, g 代表 $\tanh$ 函数。

[0088] 根据本公开的一具体实施例, 再次参见图4所示, 上述构建意图识别任务损失函数的过程例如可以包括:

[0089] a. 假设 sentence 和 y 为句子和该句子对应的类别标签, 将句子分字, 得到对应的字序列: $U = \{u_1, u_2, \dots, u_k\}$, 作为图4所示模型的输入, 其中, $u_i \in W$, k 表示输入字序列长度, u_i 表示输入句子对中的某个字, W 表示整个字表大小。

[0090] b. 通过图4所示基于BERT多任务学习的意图识别模型得到句子的表示 V_{cls} , 其维度为 d_{cls} 。

[0091] c. 计算意图类别预测结果 y_{pred} 。

[0092] d. 采用交叉熵构建意图识别任务损失函数 $\text{loss}_{\text{intent}}(\theta)$ 。

[0093] 根据本公开的实施例, 上述操作S204包括: 将问题-问题相似度标注数据集、问题-答案相似度标注数据集和意图识别数据集分别进行批量拆分, 得到多个子集合, 每个子集合存储有预设量的问题-问题相似度标注数据集、问题-答案相似度标注数据集或意图识别

数据集;将多个子集合混合合并至一个总集合;以及遍历总集合,根据总集合中的每个子集合对句子的相似度任务损失函数或意图识别任务损失函数进行训练。

[0094] 根据本公开的一具体实施例,使用上述操作S202获得的问题-问题相似度标注数据集、问题-答案相似度标注数据集和意图识别数据集,在开源BERT-Base预训练模型上设计多任务学习进行微调优化,该过程例如可以包括:

[0095] a.假设 $data_{qa}$ 、 $data_{qq}$ 和 $data_{intent}$ 分别为问题-答案相似度标注数据集、问题-问题相似度标注数据集和意图识别数据集,将三个任务数据集按batch拆分(批量拆分)后合并,然后进行shuffle(混合)操作:

[0096] $data_list_{qa} = (batch_1^{qa}, batch_2^{qa}, \dots, batch_r^{qa}),$

[0097] $data_list_{qq} = (batch_1^{qq}, batch_2^{qq}, \dots, batch_s^{qq}),$

[0098] $data_list_{intent} = (batch_1^{intent}, batch_2^{intent}, \dots, batch_t^{intent}),$

[0099] $data_merge_list = (batch_1^{qa}, batch_i^{qq}, \dots, batch_j^{intent}, \dots),$

[0100] 其中, $data_list_{qa}$ 、 $data_list_{qq}$ 和 $data_list_{intent}$ 分别对应问题-答案相似度标注数据集、问题-问题相似度标注数据集、意图识别数据集按batch拆分的集合,r、s、t分别对应集合中的batch数量, $data_merge_list$ 为三个集合合并到一起进行shuffle之后的集合。

[0101] b.对于 $data_merge_list$ 里面任意 $batch_i^{tag}$,执行以下优化过程:

[0102] $loss = \begin{cases} loss_{sim}(\theta), & \text{if } tag == qa \text{ or } qq \\ loss_{intent}(\theta), & \text{if } tag = intent \end{cases},$

[0103] $\theta = \theta - \alpha \nabla(\theta),$

[0104] 其中, θ 为模型待优化参数, $\nabla(\theta)$ 为梯度, α 为学习率。

[0105] c.采用Adam(一种可以替代传统随机梯度下降过程的一阶优化算法)优化器优化模型参数,所有数据训练K轮之后输出训练完成的意图识别模型。

[0106] 通过本公开的上述实施例,由于在训练过程中,联合采用batch和shuffle操作,可以防止网络过拟合。在此基础上采用多任务学习方式联合句子相似度任务和意图识别任务,在相似度任务上微调训练BERT模型,以及将训练好的模型作为意图识别模型的初始化模型的方式,共享相似度任务和意图识别任务参数和特征,可以进一步增强意图识别性能。

[0107] 根据本公开的上述实施例,上述意图识别方法包括:获取意图识别模型;以及通过意图识别模型对待识别句子进行意图识别;其中,意图识别模型是通过如上所述的训练方法训练得到。

[0108] 根据本公开的实施例,该意图识别模型例如可以为图4所示的模型,通过该模型,可以输出针对用户输入的一句话或一段话的类别标签,然后可以根据该类别标签定位到系统的回答,从而回答用户问题。

[0109] 通过本公开的上述实施例,联合问题-答案相似度匹配任务、问题-问题相似度匹配任务和意图识别任务三个任务同时训练,共享参数和特征,从而可以有效地增强意图识别性能。

[0110] 图5示意性示出了根据本公开的实施例的意图识别模型的训练装置的框图。

[0111] 如图5所示,训练装置500包括第一获取模块510、数据集构建模块520、损失函数构建模块530和训练模块540,其中,数据集构建模块520包括第一数据集构建单元、第二数据集构建单元和第三数据集构建单元,损失函数构建模块530包括第一损失函数构建单元和第二损失函数构建单元。

[0112] 第一获取模块510,用于获取对话语料中的问题-问题对、问题-答案对和意图类别,其中,意图类别用于表征对话语料中的每轮句子的类别标签。

[0113] 第一数据集构建单元,用于根据问题-问题对构建问题-问题相似度标注数据集。

[0114] 第二数据集构建单元,用于根据问题-答案对构建问题-答案相似度标注数据集。

[0115] 第三数据集构建单元,用于根据对话语料中的每轮句子和每轮句子对应的类别标签构建意图识别数据集。

[0116] 第一损失函数构建单元,用于根据问题-问题相似度标注数据集和问题-答案相似度标注数据集构建句子的相似度任务损失函数。

[0117] 第二损失函数构建单元,用于根据意图识别数据集构建意图识别任务损失函数。

[0118] 训练模块540,用于根据问题-问题相似度标注数据集、问题-答案相似度标注数据集和意图识别数据集对句子的相似度任务损失函数和意图识别任务损失函数进行优化训练,得到意图识别模型。

[0119] 根据本公开的实施例,上述第一获取模块包括获取单元、第一配置单元和第二配置单元。

[0120] 获取单元,用于通过自动挖掘的方式从对话语料中获取问题-问题对和问题-答案对。

[0121] 第一配置单元,用于在问题-问题对中的两个问题对应的答案相同的情况下,将问题-问题对的标签值配置为1。

[0122] 第二配置单元,用于在问题-问题对中的两个问题对应的答案不同的情况下,将问题-问题对的标签值配置为0。

[0123] 根据本公开的实施例,上述第一数据集构建单元包括第一校正子单元、第二校正子单元和第一数据集构建子单元。

[0124] 第一校正子单元,用于在问题-问题对中的两个问题的相似度小于预设阈值,且问题-问题对的标签值为1的情况下,变更问题-问题对的标签值至0或丢弃问题-问题对。

[0125] 第二校正子单元,用于在问题-问题对中的两个问题的相似度大于预设阈值,且问题-问题对的标签值为0的情况下,变更问题-问题对的标签值至1或丢弃问题-问题对。

[0126] 第一数据集构建子单元,用于根据问题-问题对及其标签构建问题-问题相似度标注数据集。

[0127] 根据本公开的实施例,上述第二数据集构建单元包括第一获取子单元、配置子单元和第二数据集构建子单元。

[0128] 第一获取子单元,用于获取两条问题-答案对: $(q1, a1)$ 和 $(q2, a2)$ 。

[0129] 配置子单元,用于在 $f(q1, q2) + f(a1, a2) < k$ 的情况下,将两条问题-答案对中句子长度更长的问题-答案对的标签值配置为1,将两条问题-答案对中句子长度更短的问题-答案对的标签值配置为0,其中, $q1$ 表征以文本形式存储的第一问题, $a1$ 表征以文本形式存储的第一答案, $q2$ 表征以文本形式存储的第二问题, $a2$ 表征以文本形式存储的第二答案, f 为

编辑距离函数,k为过滤阈值。

[0130] 第二数据集构建子单元,用于根据问题-答案对及其标签构建问题-答案相似度标注数据集。

[0131] 根据本公开的实施例,上述第三数据集构建单元包括第二获取子单元、第三获取子单元、确定子单元和第三数据集构建子单元。

[0132] 第二获取子单元,用于获取包含有n轮句子的目标对话。

[0133] 第三获取子单元,用于获取包含有t个窗口的滑窗,其中, $t < n$,t,n为正整数。

[0134] 确定子单元,用于通过滑窗对目标对话进行遍历,确定目标对话中的每轮句子对应的类别标签。

[0135] 第三数据集构建子单元,用于根据每轮句子和每轮句子对应的类别标签构建意图识别数据集。

[0136] 根据本公开的实施例,通过上述第一损失函数构建单元构建的上述句子的相似度任务损失函数为:

[0137] $\text{loss}_{\text{sim}}(\theta) = -[p \cdot \log(p_{\text{pred}}) + (1-p) \log(1-p_{\text{pred}})]$,其中:

[0138] $p_{\text{pred}} = \text{sigmod}(\text{dense}_{\text{sim}})$,

[0139] $\text{dense}_{\text{sim}} = g(V_{\text{cls}} W_{\text{sim}} + b_{\text{sim}})$,

[0140] 其中, $\text{loss}_{\text{sim}}(\theta)$ 表示句子的相似度任务损失函数, θ 表示句子的相似度任务损失函数中的待优化参数,p表示问题-问题对或问题-答案对的真实相似度标签, p_{pred} 表示问题-问题对或问题-答案对的预测相似度标签, V_{cls} 表示问题-问题对或问题-答案对, $W_{\text{sim}} \in R^{d_{\text{cls}} \times d_{\text{sim}}}$, d_{cls} 、 d_{sim} 表示矩阵R的维度, b_{sim} 表示偏置,g代表tanh函数。

[0141] 根据本公开的实施例,通过上述第二损失函数构建单元构建的上述意图识别任务损失函数为:

[0142] $\text{loss}_{\text{intent}}(\theta) = -[y \cdot \log(y_{\text{pred}}) + (1-y) \log(1-y_{\text{pred}})]$,其中:

[0143] $y_{\text{pred}} = \text{sigomd}(\text{dense}_{\text{intent}})$,

[0144] $\text{dense}_{\text{intent}} = g(V_{\text{cls}} W_{\text{intent}} + b_{\text{intent}})$,

[0145] 其中, $\text{loss}_{\text{intent}}(\theta)$ 表示意图识别任务损失函数, θ 表示意图识别任务损失函数中的待优化参数,y表示对话中的句子对应的真实类别标签, y_{pred} 表示对话中的句子对应的预测类别标签, V_{cls} 表示对话中的句子, $W_{\text{intent}} \in R^{d_{\text{cls}} \times d_{\text{intent}}}$, d_{cls} 、 d_{intent} 表示矩阵R的维度, b_{intent} 表示偏置,g代表tanh函数。

[0146] 根据本公开的实施例,上述训练模块包括拆分单元、合并单元和训练单元。

[0147] 拆分单元,用于将问题-问题相似度标注数据集、问题-答案相似度标注数据集和意图识别数据集分别进行批量拆分,得到多个子集合,每个子集合存储有预设量的问题-问题相似度标注数据集、问题-答案相似度标注数据集或意图识别数据集。

[0148] 合并单元,用于将多个子集合混合合并至一个总集合。

[0149] 训练单元,用于遍历总集合,根据总集合中的每个子集合对句子的相似度任务损失函数或意图识别任务损失函数进行训练。

[0150] 图6示意性示出了根据本公开的实施例的意图识别装置的框图。

[0151] 如图6所示,意图识别装置600包括第二获取模块610和识别模块620。

[0152] 第二获取模块610,用于获取意图识别模型,其中,所述意图识别模型是通过如上所述的意图识别模型的训练方法训练得到。

[0153] 识别模块620,用于通过所述意图识别模型对待识别句子进行意图识别。

[0154] 根据本公开的实施例的模块、单元、子单元中的任意多个、或其中任意多个的至少部分功能可以在一个模块中实现。根据本公开实施例的模块、单元、子单元中的任意一个或多个可以被拆分成多个模块来实现。根据本公开实施例的模块、单元、子单元中的任意一个或多个可以至少被部分地实现为硬件电路,例如现场可编程门阵列(FPGA)、可编程逻辑阵列(PLA)、片上系统、基板上的系统、封装上的系统、专用集成电路(ASIC),或可以通过对电路进行集成或封装的任何其他的合理方式的硬件或固件来实现,或以软件、硬件以及固件三种实现方式中任意一种或以其中任意几种的适当组合来实现。或者,根据本公开实施例的模块、单元、子单元中的一个或多个可以至少被部分地实现为计算机程序模块,当该计算机程序模块被运行时,可以执行相应的功能。

[0155] 例如,第一获取模块510、数据集构建模块520、损失函数构建模块530和训练模块540中的任意多个可以合并在一个模块/单元/子单元中实现,或者其中的任意一个模块/单元/子单元可以被拆分成多个模块/单元/子单元。或者,这些模块/单元/子单元中的一个或多个模块/单元/子单元的至少部分功能可以与其他模块/单元/子单元的至少部分功能相结合,并在一个模块/单元/子单元中实现。根据本公开的实施例,第一获取模块510、数据集构建模块520、损失函数构建模块530和训练模块540中的至少一个可以至少被部分地实现为硬件电路,例如现场可编程门阵列(FPGA)、可编程逻辑阵列(PLA)、片上系统、基板上的系统、封装上的系统、专用集成电路(ASIC),或可以通过对电路进行集成或封装的任何其他的合理方式等硬件或固件来实现,或以软件、硬件以及固件三种实现方式中任意一种或以其中任意几种的适当组合来实现。或者,第一获取模块510、数据集构建模块520、损失函数构建模块530和训练模块540中的至少一个可以至少被部分地实现为计算机程序模块,当该计算机程序模块被运行时,可以执行相应的功能。

[0156] 需要说明的是,本公开的实施例中意图识别模型的训练装置部分与本公开的实施例中意图识别模型的训练方法部分是相对应的,意图识别模型的训练装置部分的描述具体参考意图识别模型的训练方法部分,在此不再赘述。

[0157] 相应地,本公开的实施例中意图识别装置部分与本公开的实施例中意图识别方法部分是相对应的,意图识别装置部分的描述具体参考意图识别方法部分,在此不再赘述。

[0158] 图7示意性示出了根据本公开实施例的适于实现意图识别模型的训练方法或意图识别方法的计算机系统的框图。图7示出的计算机系统仅仅是一个示例,不应对本公开实施例的功能和使用范围带来任何限制。

[0159] 如图7所示,根据本公开实施例的计算机系统700包括处理器701,其可以根据存储在只读存储器(ROM)702中的程序或者从存储部分708加载到随机访问存储器(RAM)703中的程序而执行各种适当的动作和处理。处理器701例如可以包括通用微处理器(例如CPU)、指令集处理器和/或相关芯片组和/或专用微处理器(例如,专用集成电路(ASIC)),等等。处理器701还可以包括用于缓存用途的板载存储器。处理器701可以包括用于执行根据本公开实施例的方法流程的不同动作的单一处理单元或者是多个处理单元。

[0160] 在RAM 703中,存储有系统700操作所需的各种程序和数据。处理器701、ROM 702以

及RAM 703通过总线704彼此相连。处理器701通过执行ROM 702和/或RAM 703中的程序来执行根据本公开实施例的方法流程的各种操作。需要注意,所述程序也可以存储在除ROM 702和RAM 703以外的一个或多个存储器中。处理器701也可以通过执行存储在所述一个或多个存储器中的程序来执行根据本公开实施例的方法流程的各种操作。

[0161] 根据本公开的实施例,系统700还可以包括输入/输出(I/O)接口705,输入/输出(I/O)接口705也连接至总线704。系统700还可以包括连接至I/O接口705的以下部件中的一项或多项:包括键盘、鼠标等的输入部分706;包括诸如阴极射线管(CRT)、液晶显示器(LCD)等以及扬声器等的输出部分707;包括硬盘等的存储部分708;以及包括诸如LAN卡、调制解调器等的网络接口卡的通信部分709。通信部分709经由诸如因特网的网络执行通信处理。驱动器710也根据需要连接至I/O接口705。可拆卸介质711,诸如磁盘、光盘、磁光盘、半导体存储器等等,根据需要安装在驱动器710上,以便于从其上读出的计算机程序根据需要被安装入存储部分708。

[0162] 根据本公开的实施例,根据本公开实施例的方法流程可以被实现为计算机软件程序。例如,本公开的实施例包括一种计算机程序产品,其包括承载在计算机可读存储介质上的计算机程序,该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中,该计算机程序可以通过通信部分709从网络上被下载和安装,和/或从可拆卸介质711被安装。在该计算机程序被处理器701执行时,执行本公开实施例的系统中限定的上述功能。根据本公开的实施例,上文描述的系统、设备、装置、模块、单元等可以通过计算机程序模块来实现。

[0163] 本公开还提供了一种计算机可读存储介质,该计算机可读存储介质可以是上述实施例中描述的设备/装置/系统中所包含的;也可以是单独存在,而未装配入该设备/装置/系统中。上述计算机可读存储介质承载有一个或者多个程序,当上述一个或者多个程序被执行时,实现根据本公开实施例的方法。

[0164] 根据本公开的实施例,计算机可读存储介质可以是非易失性的计算机可读存储介质。例如可以包括但不限于:便携式计算机磁盘、硬盘、随机访问存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、便携式紧凑磁盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本公开中,计算机可读存储介质可以是任何包含或存储程序的有形介质,该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。

[0165] 例如,根据本公开的实施例,计算机可读存储介质可以包括上文描述的ROM 702和/或RAM 703和/或ROM 702和RAM 703以外的一个或多个存储器。

[0166] 附图中的流程图和框图,图示了按照本公开各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分,上述模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意,在有些作为替换的实现中,方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,两个接连地表示的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意,框图或流程图中的每个方框、以及框图或流程图中的方框的组合,可以用执行规定的功能或操作的专用的基于硬件的系统来实现,或者可以用专用硬件与计算机指令的组

合来实现。本领域技术人员可以理解,本公开的各个实施例和/或权利要求中记载的特征可以进行多种组合和/或结合,即使这样的组合或结合没有明确记载于本公开中。特别地,在不脱离本公开精神和教导的情况下,本公开的各个实施例和/或权利要求中记载的特征可以进行多种组合和/或结合。所有这些组合和/或结合均落入本公开的范围。

[0167] 以上对本公开的实施例进行了描述。但是,这些实施例仅仅是为了说明的目的,而并非为了限制本公开的范围。尽管在以上分别描述了各实施例,但是这并不意味着各个实施例中的措施不能有利地结合使用。本公开的范围由所附权利要求及其等同物限定。不脱离本公开的范围,本领域技术人员可以做出多种替代和修改,这些替代和修改都应落在本公开的范围之内。

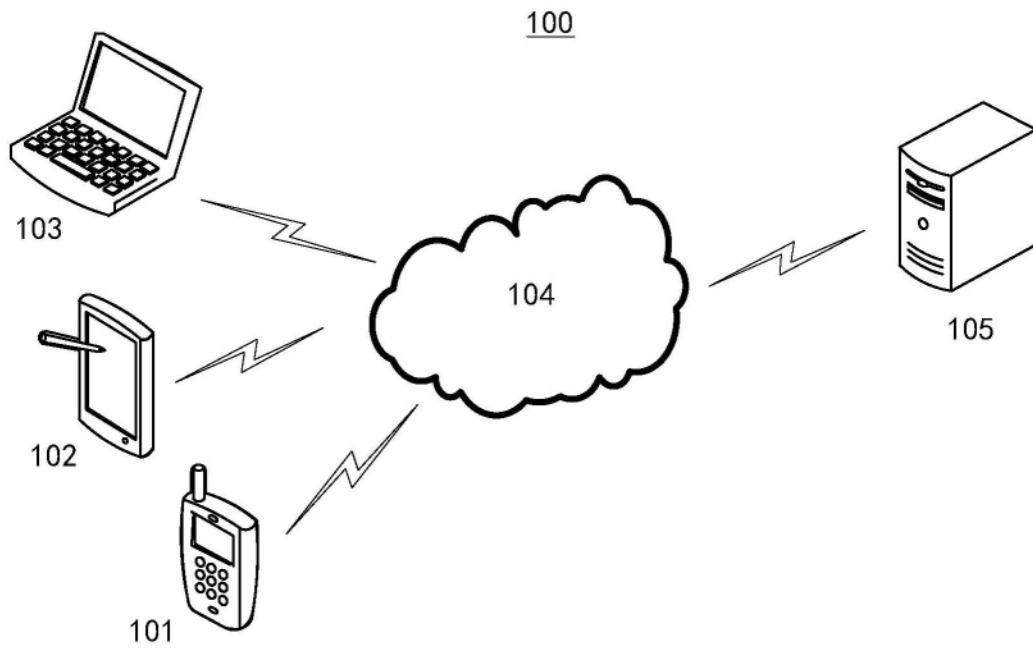


图1

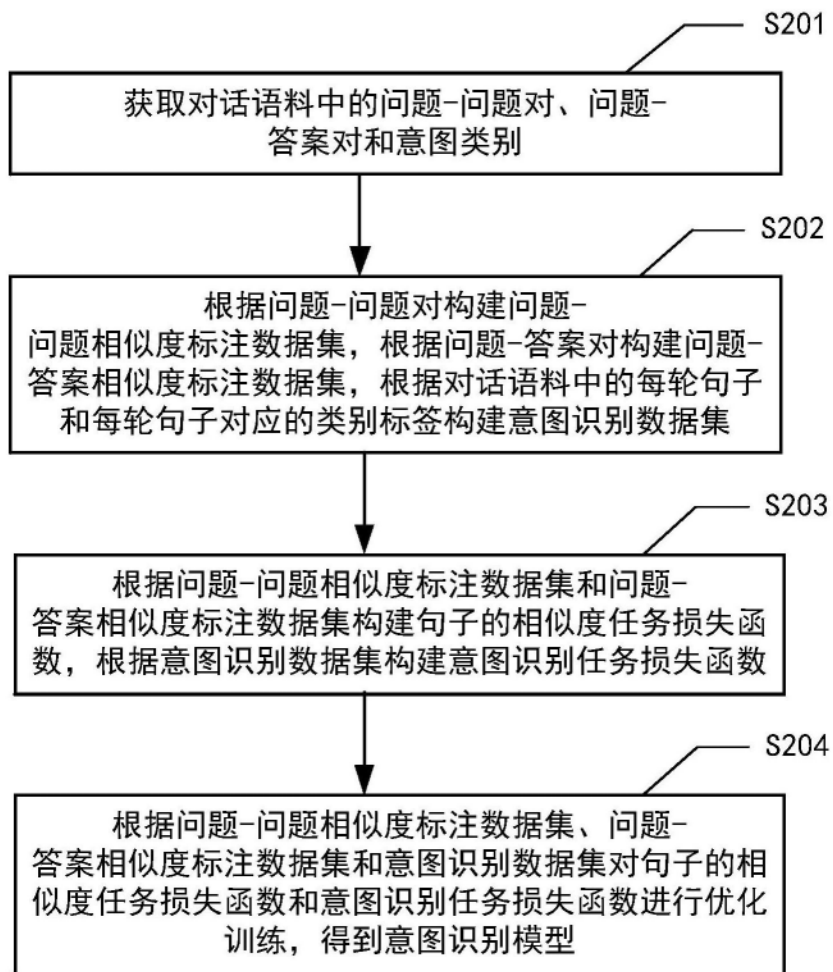


图2

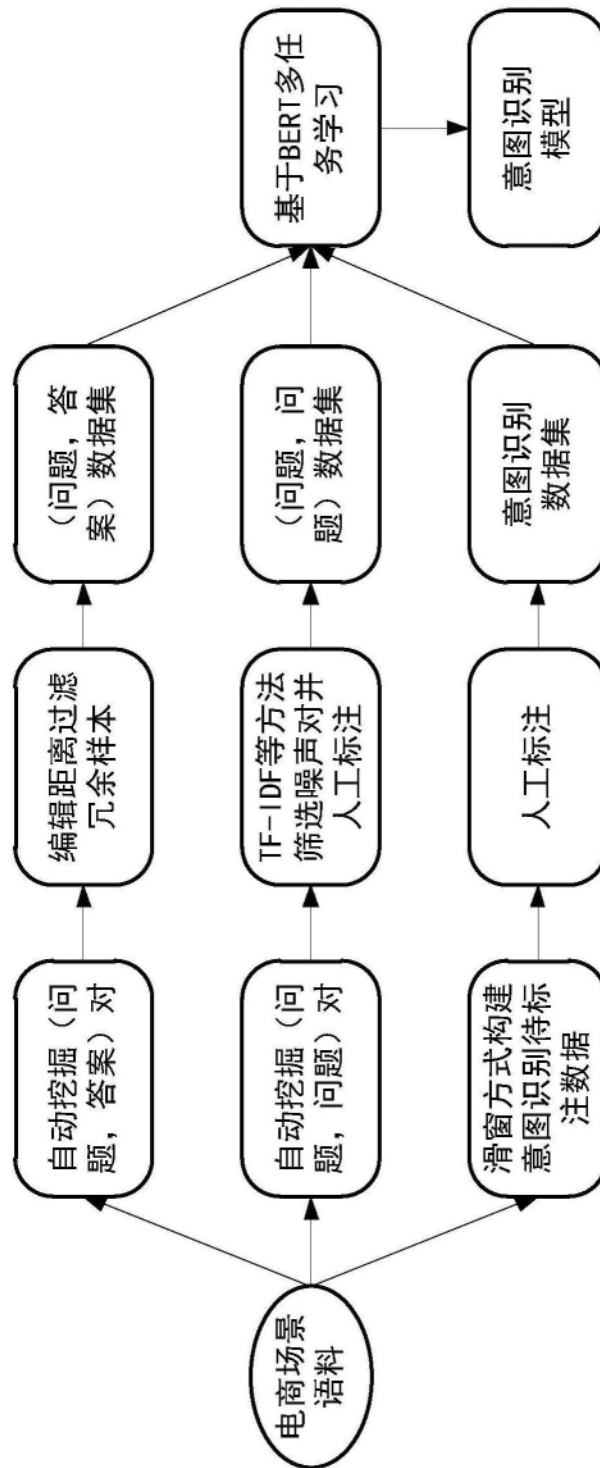


图3

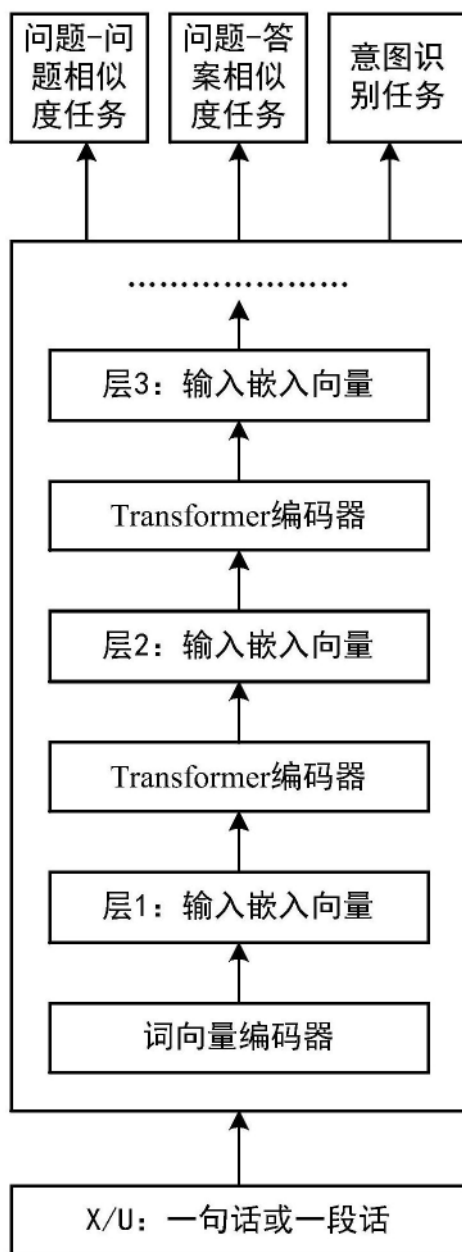


图4

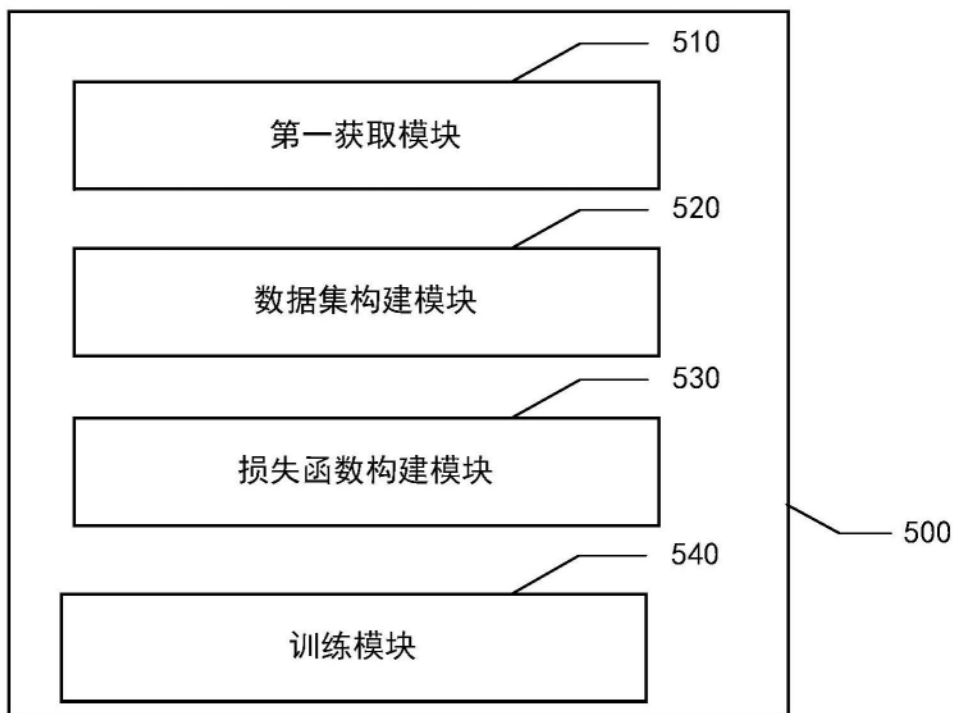


图5

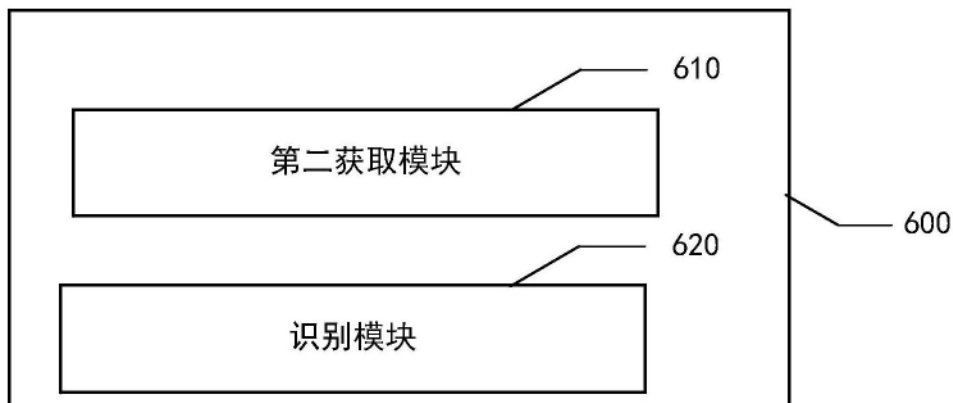


图6

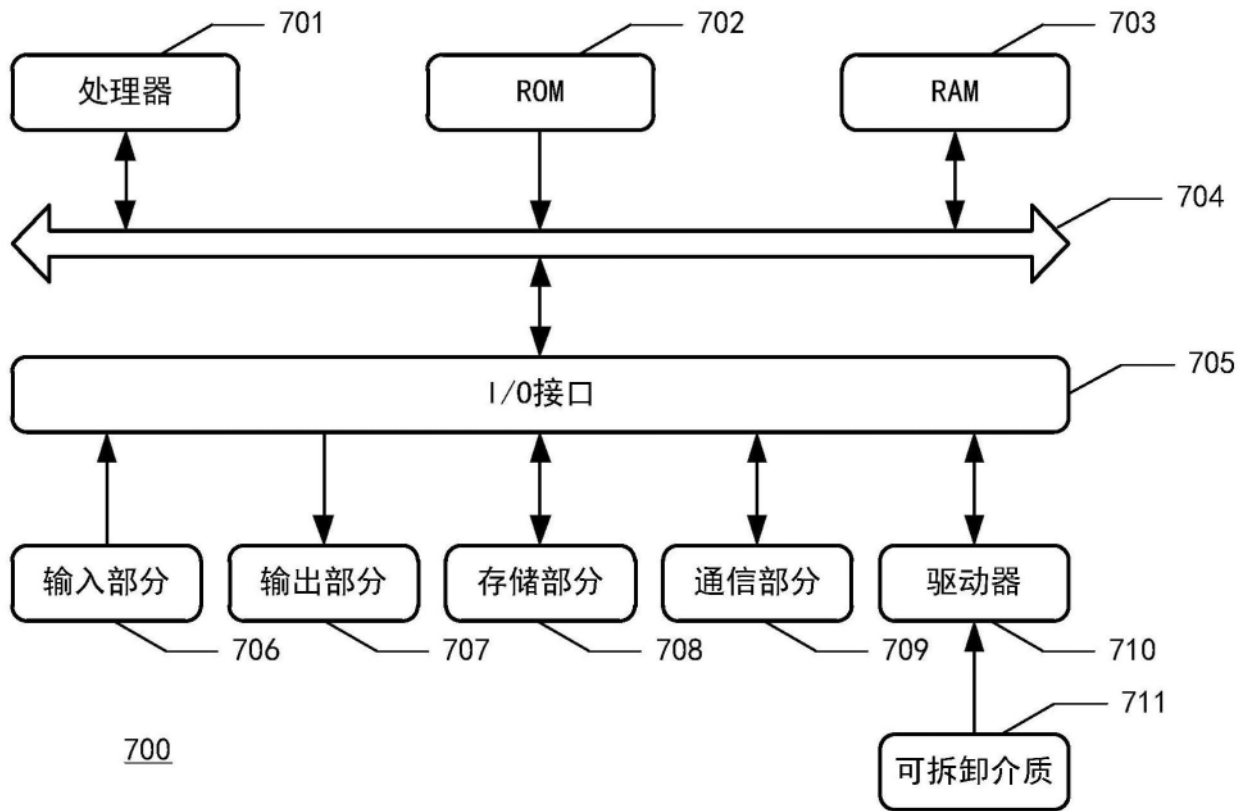


图7