



(12) 发明专利

(10) 授权公告号 CN 107943860 B

(45) 授权公告日 2020.10.27

(21) 申请号 201711092668.3

G06K 9/62 (2006.01)

(22) 申请日 2017.11.08

(56) 对比文件

(65) 同一申请的已公布的文献号

CN 107229684 A, 2017.10.03

申请公布号 CN 107943860 A

CN 107247702 A, 2017.10.13

CN 106462626 A, 2017.02.22

(43) 申请公布日 2018.04.20

US 8700552 B2, 2014.04.15

(73) 专利权人 北京奇艺世纪科技有限公司

审查员 吴瑶

地址 100080 北京市海淀区北一街2号爱奇艺

艺创新大厦10、11层

(72) 发明人 鲍新平 丁希晨

(74) 专利代理机构 北京润泽恒知识产权代理有

限公司 11319

代理人 莎日娜

(51) Int. Cl.

G06F 16/35 (2019.01)

G06F 16/36 (2019.01)

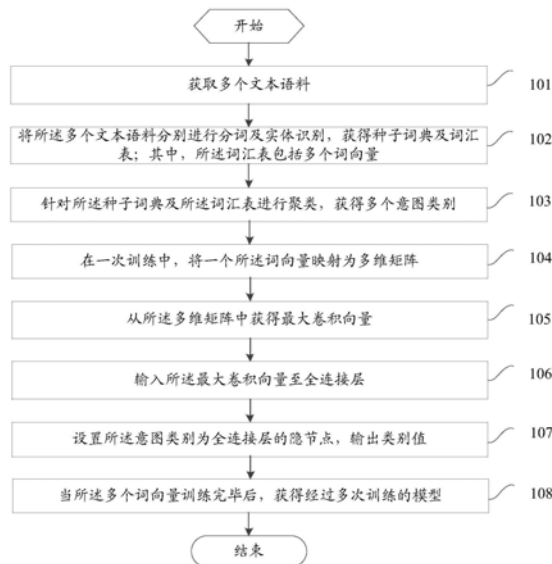
权利要求书2页 说明书14页 附图5页

(54) 发明名称

模型的训练方法、文本意图的识别方法及装置

(57) 摘要

本发明实施例提供了一种模型的训练方法、文本意图的识别方法及装置,获取多个文本语料;将所述多个文本语料分别进行分词及实体识别,获得种子词典及词汇表;针对所述种子词典及所述词汇表进行聚类,获得多个意图类别;在一次训练中,将一个所述词向量映射为多维矩阵;从所述多维矩阵中获得最大卷积向量;输入所述最大卷积向量至全连接层;设置所述意图类别为全连接层的隐节点,输出类别值;当所述多个词向量训练完毕后,获得经过多次训练的模型;本发明实施例中,提出了一种基于词向量的确定意图类别的方法,相比于传统人工设定和枚举的方法,善于发现新的意图类别;采用训练得到的模型使文本意图的识别率更加高。



1. 一种模型的训练方法,其特征在于,所述方法包括:

获取多个文本语料;

将所述多个文本语料分别进行分词及实体识别,获得种子词典及词汇表;其中,所述词汇表包括多个词向量;所述词汇表为对分词后的文本语料进行实体识别得到的实体组合;所述词向量为所述实体组合中的实体;所述种子词典由种子词组成,所述种子词为所述词汇表中包含意图的动词;

针对所述种子词典及所述词汇表进行聚类,获得多个意图类别;

在一次训练中,将一个所述词向量映射为多维矩阵;

从所述多维矩阵中获得最大卷积向量;

输入所述最大卷积向量至全连接层;

设置所述意图类别为全连接层的隐节点,输出类别值;

当所述多个词向量训练完毕后,获得经过多次训练的模型;

所述针对所述种子词典及词汇表进行聚类,获得多个意图类别的步骤包括:

将每个种子词及每个词向量分别进行组合,计算组合后的所述种子词及所述词向量之间的多个特征相似度;

分别判断各个所述特征相似度是否大于预设阈值;

当存在大于预设阈值的特征相似度时,则确定该大于预设阈值的所述特征相似度对应的种子词为意图类别,以及,确定该大于预设阈值的所述特征相似度对应的词向量为所述意图类别对应的意图词。

2. 根据权利要求1所述的方法,其特征在于,所述将所述多个文本语料进行分词及实体识别,获得种子词典及词汇表的步骤包括:

针对所述文本语料进行分词及实体识别,获得词汇表;

从所述词汇表中提取出多个种子词,组合成种子词典。

3. 根据权利要求1所述的方法,其特征在于,所述从所述多维矩阵中获得最大卷积向量的步骤包括:

从所述多维矩阵中选取不同数量的卷积核;

针对所述卷积核进行计算,得到卷积向量;

确定卷积向量的最大值为最大卷积向量。

4. 一种文本意图的识别方法,其特征在于,应用所述权利要求1-3任意一项的所述模型,所述方法包括:

获取输入文本;

将所述输入文本对应的词向量输入所述模型,得到所述模型输出的最大概率的类别值;获取所述模型输出的所述最大概率的类别值对应的意图类别;

依据所述意图类别获取所述输入文本的意图词。

5. 一种模型的训练装置,其特征在于,所述装置包括:

文本语料获取模块,用于获取多个文本语料;

种子词典及词汇表获得模块,用于将所述多个文本语料分别进行分词及实体识别,获得种子词典及词汇表;其中,所述词汇表包括多个词向量;所述词汇表为对分词后的文本语料进行实体识别得到的实体组合;所述词向量为所述实体组合中的实体;所述种子词典由

种子词组成,所述种子词为所述词汇表中包含意图的动词;

聚类模块,用于针对所述种子词典及所述词汇表进行聚类,获得多个意图类别;

第一映射模块,用于在一次训练中,将一个所述词向量映射为多维矩阵;

最大卷积向量获得模块,用于从所述多维矩阵中获得最大卷积向量;

第一输入模块,用于输入所述最大卷积向量至全连接层;

第一类别值输出模块,用于设置所述意图类别为全连接层的隐节点,输出类别值;

模型获得模块,用于当所述多个词向量训练完毕后,获得经过多次训练的模型;

所述聚类模块包括:

计算子模块,用于将每个种子词及每个词向量分别进行组合,计算组合后的所述种子词及所述词向量之间的多个特征相似度;

判断子模块,用于判断所述特征相似度是否大于预设阈值;

确定子模块,用于当存在大于预设阈值的特征相似度时,则确定该大于预设阈值的所述特征相似度对应的种子词为意图类别,以及,确定该大于预设阈值的所述特征相似度对应的词向量为所述意图类别对应的意图词。

6. 根据权利要求5所述的装置,其特征在於,所述种子词典及词汇表获得模块包括:

词汇表获得子模块,用于针对所述文本语料进行分词及实体识别,获得词汇表;

种子词典组合子模块,用于从所述词汇表提取出多个种子词,组合成种子词典。

7. 根据权利要求5所述的装置,其特征在於,所述最大卷积向量获得模块包括:

选取子模块,用于从所述多维矩阵中选取不同数量的卷积核;

卷积向量得到子模块,用于针对所述卷积核进行计算,得到卷积向量;

最大卷积向量确定子模块,用于确定卷积向量的最大值为最大卷积向量。

8. 一种文本意图的识别装置,其特征在於,应用所述权利要求1-3任意一项的所述模型,所述装置包括:

输入文本获取模块,用于获取输入文本;

输入模块,用于将所述输入文本对应的词向量输入所述模型,得到所述模型输出的最大概率的类别值;

意图类别输出模块,用于获取所述模型输出的所述最大概率的类别值对应的意图类别;

意图词获取模块,用于依据所述意图类别获取所述输入文本的意图词。

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,其特征在於,所述处理器执行所述程序时实现权利要求1至4任一项所述方法的步骤。

10. 一种计算机可读存储介质,其上存储有计算机程序,其特征在於,该程序被处理器执行时实现权利要求1至4任一项所述方法的步骤。

## 模型的训练方法、文本意图的识别方法及装置

### 技术领域

[0001] 本发明涉及计算机的技术领域,特别是涉及一种模型的训练方法、一种文本意图的识别方法和一种模型的训练装置、一种文本意图的识别装置。

### 背景技术

[0002] 机器学习(Machine Learning,ML)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。

[0003] 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,主要使用归纳、综合而不是演绎。机器学习是一门让计算机在非精确编程下进行活动的科学。在过去十年,机器学习促成了无人驾驶车、高效语音识别、精确网络搜索及人类基因组认知的大力发展。

[0004] 文本的意图识别被广泛的应用于搜索引擎,Query(查询)识别,人机对话等领域。目前传统的基于机器学习的意图识别方法存在准确度不高,意图类别无法自动确定等等弊端。

### 发明内容

[0005] 鉴于上述问题,提出了本发明实施例以便提供一种克服上述问题或者至少部分地解决上述问题的一种模型的训练方法、一种文本意图的识别方法和一种模型训练的装置、一种文本意图的识别装置、一种电子设备、一种计算机可读存储介质。

[0006] 为了解决上述问题,本发明实施例公开了一种模型的训练方法,所述方法包括:

[0007] 获取多个文本语料;

[0008] 将所述多个文本语料分别进行分词及实体识别,获得种子词典及词汇表;其中,所述词汇表包括多个词向量;

[0009] 针对所述种子词典及所述词汇表进行聚类,获得多个意图类别;

[0010] 在一次训练中,将一个所述词向量映射为多维矩阵;

[0011] 从所述多维矩阵中获得最大卷积向量;

[0012] 输入所述最大卷积向量至全连接层;

[0013] 设置所述意图类别为全连接层的隐节点,输出类别值;

[0014] 当所述多个词向量训练完毕后,获得经过多次训练的模型。

[0015] 优选地,所述将所述多个文本语料进行分词及实体识别,获得种子词典及词汇表的步骤包括:

[0016] 针对所述文本语料进行分词及实体识别,获得词汇表;

[0017] 从所述词汇表提取出多个种子词,组合成种子词典。

[0018] 优选地,所述针对所述种子词典及词汇表进行聚类,获得多个意图类别的步骤包括:

- [0019] 将每个种子词及每个词向量分别进行组合,计算组合后的所述种子词及所述词向量之间的多个特征相似度;
- [0020] 分别判断各个所述特征相似度是否大于预设阈值;
- [0021] 当存在大于预设阈值的特征相似度时,则确定该大于预设阈值的所述特征相似度对应的种子词为意图类别,以及,确定该大于预设阈值的所述特征相似度对应的词向量为所述意图类别对应的意图词。
- [0022] 优选地,所述从所述多维矩阵中获得最大卷积向量的步骤包括:
- [0023] 从所述多维矩阵中选取不同数量的卷积核;
- [0024] 针对所述卷积核进行计算,得到卷积向量;
- [0025] 确定卷积向量的最大值为最大卷积向量。
- [0026] 本发明实施例还公开了一种文本意图的识别方法,应用上述模型,所述方法包括:
- [0027] 获取输入文本;
- [0028] 将所述输入文本对应的词向量输入所述模型;
- [0029] 获取所述模型输出的所述最大概率的类别值对应的意图类别;
- [0030] 依据所述意图类别获取所述输入文本的意图词。
- [0031] 本发明实施例还公开了一种模型的训练装置,所述装置包括:
- [0032] 文本语料获取模块,用于获取多个文本语料;
- [0033] 种子词典及词汇表获得模块,用于将所述多个文本语料分别进行分词及实体识别,获得种子词典及词汇表;其中,所述词汇表包括多个词向量;
- [0034] 聚类模块,用于针对所述种子词典及所述词汇表进行聚类,获得多个意图类别;
- [0035] 第一映射模块,用于在一次训练中,将一个所述词向量映射为多维矩阵;
- [0036] 最大卷积向量获得模块,用于从所述多维矩阵中获得最大卷积向量;
- [0037] 第一输入模块,用于输入所述最大卷积向量至全连接层;
- [0038] 第一类别值输出模块,用于设置所述意图类别为全连接层的隐节点,输出类别值;
- [0039] 模型获得模块,用于当所述多个词向量训练完毕后,获得经过多次训练的模型。
- [0040] 优选地,所述种子词典及词汇表获得模块包括:
- [0041] 词汇表获得子模块,用于针对所述文本语料进行分词及实体识别,获得词汇表;
- [0042] 种子词典组合子模块,用于从所述词汇表提取出多个种子词,组合成种子词典。
- [0043] 优选地,所述聚类模块包括:
- [0044] 计算子模块,用于将每个种子词及每个词向量分别进行组合,计算组合后的所述种子词及所述词向量之间的多个特征相似度;
- [0045] 判断子模块,用于判断所述特征相似度是否大于预设阈值;
- [0046] 确定子模块,用于当存在大于预设阈值的特征相似度时,则确定该大于预设阈值的所述特征相似度对应的种子词为意图类别,以及,确定该大于预设阈值的所述特征相似度对应的词向量为所述意图类别对应的意图词。
- [0047] 优选地,所述最大卷积向量获得模块包括:
- [0048] 选取子模块,用于从所述多维矩阵中选取不同数量的卷积核;
- [0049] 卷积向量得到子模块,用于针对所述卷积核进行计算,得到卷积向量;
- [0050] 最大卷积向量确定子模块,用于确定卷积向量的最大值为最大卷积向量。

- [0051] 本发明实施例还公开了一种文本意图的识别装置,所述装置包括:
- [0052] 输入文本获取模块,用于获取输入文本;
- [0053] 输入模块,用于将所述输入文本对应的词向量输入所述模型;
- [0054] 意图类别输出模块,用于获取所述模型输出的所述最大概率的类别值对应的意图类别;
- [0055] 意图词获取模块,用于依据所述意图类别获取所述输入文本的意图词。
- [0056] 本发明实施例还公开了一种电子设备,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上述方法的步骤。
- [0057] 本发明实施例还公开了一种计算机可读存储介质,其上存储有计算机程序,其特征在于,该程序被处理器执行时实现上述方法的步骤。
- [0058] 本发明实施例包括以下优点:
- [0059] 本发明实施例中,获取多个文本语料,将所述多个文本语料分别进行分词及实体识别,获得种子词典及词汇表,其中,所述词汇表包括多个词向量,针对所述种子词典及所述词汇表进行聚类,获得多个意图类别,在一次训练中,将一个所述词向量映射为多维矩阵,从所述多维矩阵中获得最大卷积向量,输入所述最大卷积向量至全连接层,设置所述意图类别为全连接层的隐节点,输出类别值,当所述多个词向量训练完毕后,获得经过多次训练的模型;本发明实施例中,提出了一种基于词向量的确定意图类别的方法,相比于传统人工设定和枚举的方法,泛化能力更强,善于发现新的意图类别;采用训练得到的模型使文本意图的识别率更加高。
- [0060] 进一步地,本发明实施例中,获取输入文本,将所述输入文本对应的词向量输入所述模型;获取所述模型输出的所述最大概率的类别值对应的意图类别;依据所述意图类别获取所述输入文本的意图词,相比于传统的机器学习进行文本意图识别准确率不高的情况,应用了CNN模型后,并考虑到词向量的信息,使得意图识别的准确率得到大幅提升。

## 附图说明

- [0061] 图1是本发明实施例的一种模型的训练方法实施例一的步骤流程图;
- [0062] 图2是本发明实施例的一种模型的训练方法实施例二的步骤流程图;
- [0063] 图3是本发明实施例的一种文本意图的识别方法实施例的步骤流程图;
- [0064] 图4是本发明实施例的模型的训练方法及文本意图的识别方法的步骤流程图;
- [0065] 图5是本发明实施例的一种模型的训练装置实施例的结构框图;
- [0066] 图6是本发明实施例的一种文本意图的识别装置实施例的结构框图;
- [0067] 图7示出了本发明实施例提供的一种终端设备结构示意图。

## 具体实施方式

[0068] 为使本发明的上述目的、特征和优点能够更加明显易懂,下面结合附图和具体实施方式对本发明作进一步详细的说明。

[0069] 参照图1,示出了本发明的一种模型的训练方法实施例一的步骤流程图,具体可以包括如下步骤:

[0070] 步骤101,获取多个文本语料;

[0071] 本发明实施例中,所述模型可以包括卷积神经网络模型,该卷积神经网络模型相当于一个分类器,针对用户的输入依据训练好的卷积神经网络模型输出设置的类别值;具体地,该卷积神经网络模型可以包括词嵌入层、卷积层、池化层及全连接层,该词嵌入层主要用于将词向量映射为多维矩阵,该卷积层主要用于获得相应的卷积向量,该池化层主要用于从多维矩阵中获得最大卷积向量,该全连接层主要用于输入所述最大卷积向量至全连接层,设置所述类别值为全连接层的隐节点,训练后输出类别值。

[0072] 具体应用到本发明实施例中,所述训练方法首先采集足够多的文本语料,从该文本语料中提取出种子词典、词汇表及意图类别以供后续的模式训练使用。

[0073] 本发明实施例中,可以获取多个文本语料,需要说明的是,该文本语料的来源可以从网络上抓取的,本发明实施例中对此不作具体的限制。

[0074] 步骤102,将所述多个文本语料分别进行分词及实体识别,获得种子词典及词汇表;其中,所述词汇表包括多个词向量;

[0075] 实际应用中,获取多个文本语料后,采用如双向长短期记忆-条件随机场层模型(Bi-LSTM-CRF)将分词后的文本进行实体识别,获得种子词典及词汇表,即将识别出来的实体组合为词汇表,将该词汇表中的包含意图的动词确定为种子词;该词汇表中的实体以词向量的形式呈现,再将种子词组成为种子词典,该实体可以包括“人名”、“地名”、“组织”、“商业品牌”等,本发明实施例对此不作具体的限制。

[0076] 步骤103,针对所述种子词典及所述词汇表进行聚类,获得多个意图类别;

[0077] 作为应用本发明实施例的一种示例,获取到该种子词典及词汇表后,针对该种子词典及所述词汇表进行聚类,获得多个意图类别,具体地,将每个种子词及每个词向量分别进行组合,计算组合后的所述种子词及所述词向量之间的多个特征相似度;分别判断各个所述特征相似度是否大于预设阈值;当存在大于预设阈值的特征相似度时,则确定该大于预设阈值的所述特征相似度对应的种子词为意图类别,以及,确定该大于预设阈值的所述特征相似度对应的词向量为所述意图类别对应的意图词,这样,就可以将该意图词(词向量)归类至该意图类别(种子词)下,这样,一个意图类别可以对应一个或者多个意图词,建立起意图类别与意图词之间的对应关系,可以认为上述的步骤为采用模型模块进行训练前的准备步骤。

[0078] 步骤104,在一次训练中,将一个所述词向量映射为多维矩阵;

[0079] 本发明实施例中,在一次训练中,将一个所述词向量输入到CNN模型中进行训练,具体地,可以将词汇表中的词向量映射为多个维矩阵,如,当该词向量的个数为100个时,则将映射为100\*200维的矩阵。

[0080] 步骤105,从所述多维矩阵中获得最大卷积向量;

[0081] 具体到本发明实施例中,从所述多维矩阵中选取不同数量的卷积核,针对所述卷积核进行计算,得到卷积向量,确定卷积向量的最大值为最大卷积向量。

[0082] 步骤106,输入所述最大卷积向量至全连接层;

[0083] 步骤107,设置所述意图类别为全连接层的隐节点,输出类别值;

[0084] 实际应用到本发明实施例中,将该最大卷积向量作为全连接层的输入,设定2-3层的隐含层,及将意图类别的个数设置为全连接层的隐节点数,即每个意图类别对应一个隐节点,当该模型在训练完毕后,则会输出类别值,可以通过类别值获取到对应的意图类别,

需要说明的是,每个类别值对应一个意图类别。

[0085] 需要说明的是,从上述的将一个词向量映射为多维矩阵的步骤开始,至输出类别值的步骤为一次训练的过程;而模型训练通常是一个多次训练的过程,即在模型训练过程中,可以有多个训练样本,采用该多个训练样本进行训练,在本发明实施例中,可以以词向量为训练样本,针对模型进行多次的训练。

[0086] 步骤108,当所述多个词向量训练完毕后,获得经过多次训练的模型。

[0087] 本发明实施例中,以所述多个词向量为训练样本,在一次训练中,输入一个词向量,将该词向量映射为多维矩阵,直至输出对应的类别值;而当所有的训练样本都训练完毕后,即多个词向量都输入完毕后,可以获得该经过多次训练的模型。

[0088] 本发明实施例中,获取多个文本语料,将所述多个文本语料分别进行分词及实体识别,获得种子词典及词汇表,其中,所述词汇表包括多个词向量,针对所述种子词典及所述词汇表进行聚类,获得多个意图类别,在一次训练中,将一个所述词向量映射为多维矩阵,从所述多维矩阵中获得最大卷积向量,输入所述最大卷积向量至全连接层,设置所述意图类别为全连接层的隐节点,输出类别值,当所述多个词向量训练完毕后,获得经过多次训练的模型;本发明实施例中,提出了一种基于词向量的确定意图类别的方法,相比于传统人工设定和枚举的方法,泛化能力更强,善于发现新的意图类别;采用训练得到的模型使文本意图的识别率更加高。

[0089] 参照图2,示出了本发明的一种模型的训练方法实施例二的步骤流程图,具体可以包括如下步骤:

[0090] 步骤201,获取多个文本语料;

[0091] 本发明实施例中,可以获取多个文本语料,需要说明的是,该文本语料可以是直接从网络上直接抓取的,本发明实施例中,对文本语料的数量不作限制。

[0092] 步骤202,将所述多个文本语料分别进行分词及实体识别,获得种子词典及词汇表;其中,所述词汇表包括多个词向量;

[0093] 进一步地,获取多个文本语料后,可以对所述多个文本语料进行分词,及识别出所述分词后的文本语料中的实体,该实体可以包括“人名”、“地名”、“组织”、“商业品牌”等,本发明实施例对此不作具体的限制;具体而言,可以采用多个模型对所述文本语料进行分词及实体识别,如采用双向长短期记忆-条件随机场层模型(Bi-LSTM-CRF)进行实体识别,本发明实施例对此不作限制,提高了实体识别的准确率。

[0094] 本发明实施例的一种优选实施例中,所述将所述多个文本语料进行分词,获得种子词典及词汇表的步骤包括以下子步骤:

[0095] 子步骤S2021,针对所述文本语料进行分词及实体识别,获得词汇表;

[0096] 子步骤S2022,从所述词汇表提取出多个种子词,组合成种子词典。

[0097] 具体地,采用所述分词后的文本语料作为词汇表,其中,该词汇表中每个词的表现形式可以包括词向量,将词汇表中的包含意图的动词确定为种子词,将所述种子词组合成种子词典;即所述种子词典包括多个种子词。

[0098] 步骤203,将每个种子词及每个词向量分别进行组合,计算组合后的所述种子词及所述词向量之间的多个特征相似度;

[0099] 具体应用到本发明实施例中,遍历所述种子词及所述词向量,可以对该种子词典

中的种子词及词汇表中的词向量计算两者的特征相似度,该特征相似度可以为cosine相似度,本发明实施例对此不作具体地限制。

[0100] 具体地,将每个种子词及每个词向量分别进行组合,计算组合后的所述种子词及所述词向量之间的多个特征相似度。

[0101] 步骤204,分别判断各个所述特征相似度是否大于预设阈值;

[0102] 进一步地,可以判断每个特征相似度是否大于预设阈值,需要说明的是,该预设阈值可以是本领域技术人员根据实际情况而设定的任何数值,本发明实施例对此不作限制。

[0103] 步骤205,当存在大于预设阈值的特征相似度时,则确定该大于预设阈值的所述特征相似度对应的种子词为意图类别,以及,确定该大于预设阈值的所述特征相似度对应的词向量为所述意图类别对应的意图词;

[0104] 实际应用到本发明实施例中,当存在大于预设阈值的特征相似度时,则确定该大于预设阈值的所述特征相似度对应的种子词为意图类别,以及,确定该大于预设阈值的所述特征相似度对应的词向量为所述意图类别对应的意图词,即遍历所有的种子词与所有的词向量,将种子词与词向量进行组合,计算组合后的特征相似度,获得意图类别及对应的意图词,举例而言,意图类别可以为:账户付费类、购买类、观看类等。

[0105] 步骤206,在一次训练中,将一个所述词向量映射为多维矩阵;

[0106] 进一步地,本发明实施例中,在一次训练中,将一个所述词向量输入到CNN模型中进行训练,具体地,可以将词汇表中的词向量映射为多个维矩阵,如,当该词向量的个数为100个时,则将映射为100\*200维的矩阵。

[0107] 步骤207,从所述多维矩阵中获得最大卷积向量;

[0108] 实际而言,将该多维矩阵进行卷积,得到不同的卷积向量,从不同的卷积向量得到最大卷积向量,具体地,将所述卷积向量的最大值进行提取,构成最大卷积向量。

[0109] 本发明实施例的另一种优选实施例中,所述从所述多维矩阵中获得最大卷积向量的步骤包括以下子步骤:

[0110] 子步骤S2071,从所述多维矩阵中选取不同数量的卷积核;

[0111] 子步骤S2072,针对所述卷积核进行计算,得到卷积向量;

[0112] 子步骤S2073,确定卷积向量的最大值为最大卷积向量。

[0113] 步骤208,输入所述最大卷积向量至全连接层;

[0114] 步骤209,设置所述意图类别为全连接层的隐节点,输出类别值;

[0115] 本发明实施例中,将该最大卷积向量作为全连接层的输入,设定2-3层的隐含层,并将意图类别的个数设置为全连接层的隐节点数,隐节点相当于一个感知器,产生一个输出值,即每个意图类别对应一个隐节点,当该模型在训练完毕后,则会输出类别值,可以通过类别值获取到对应的意图类别,需要说明的是,每个类别值对应一个意图类别。

[0116] 步骤210,当所述多个词向量训练完毕后,获得经过多次训练的模型。

[0117] 实际应用到本发明实施例中,以所述多个词向量为训练样本,在一次训练中,输入一个词向量,将该词向量映射为多维矩阵,直至输出对应的类别值;而当所有的训练样本都训练完毕后,即多个词向量都输入后,可以获得该经过多次训练的模型。

[0118] 本发明实施例中,获取多个文本语料,将所述多个文本语料分别进行分词及实体识别,获得种子词典及词汇表,其中,所述词汇表包括多个词向量,将每个种子词及每个词

向量分别进行组合,计算组合后的所述种子词及所述词向量之间的多个特征相似度;分别判断各个所述特征相似度是否大于预设阈值;当存在大于预设阈值的特征相似度时,则确定该大于预设阈值的所述特征相似度对应的种子词为意图类别,以及,确定该大于预设阈值的所述特征相似度对应的词向量为所述意图类别对应的意图词,在一次训练中,将一个所述词向量映射为多维矩阵,从所述多维矩阵中获得最大卷积向量,输入所述最大卷积向量至全连接层,设置所述意图类别为全连接层的隐节点,输出类别值,当所述多个词向量训练完毕后,获得经过多次训练的模型;本发明实施例中,提出了一种基于词向量的确定意图类别的方法,相比于传统人工设定和枚举的方法,泛化能力更强,善于发现新的意图类别;使用特定模型进行语料的实体识别,考虑到了实体和分词之间的转移关系,提高了识别的准确度,基于种子词与词向量的相似度进行意图词的聚类,减少了偏差;采用训练得到的模型使文本意图的识别率更加高。

[0119] 参照图3,示出了本发明的一种文本意图的识别方法实施例的步骤流程图,具体可以包括如下步骤:

[0120] 步骤301,获取输入文本;

[0121] 本发明实施例中,可以获取用户的输入文本,如用户在搜索框中输入“为了在爱奇艺看电视剧《老九门》,去给会员充值”,程序可以自动从相关网页的搜索框进行抓取。

[0122] 步骤302,将所述输入文本对应的词向量输入所述模型;

[0123] 本发明实施例中,将该输入文本对应的词向量输入所述模型;因为模型是经过训练的,则此模型会进行相应的运算,输出类别值。

[0124] 具体而言,将所述输入文本对应的词向量映射为多维矩阵;从所述多维矩阵中获得最大卷积向量;输入所述最大卷积向量至全连接层;获得所述全连接层中的最大概率的类别值。

[0125] 本发明实施例的一种优选实施例中,所述将所述输入文本对应的词向量映射为多维矩阵的步骤包括以下子步骤:将所述输入文本进行分词及实体识别,获得所述输入文本对应的词向量;将所述词向量映射为多维矩阵。

[0126] 首先,将该输入文本进行分词及实体识别,对识别出来的分词识别不同的标签,举例而言,上述的输入文本“为了在爱奇艺看电视剧《老九门》,去给会员充值”如“为了[nt]在[nt]爱奇艺[nbz]看[nt]电视剧[nt]《[nt]老九门[nz]》[nt],[nt]去[nt]给[nt]会员[nt]充值[nt]”。一共自动发现了两个实体,一个品牌实体“爱奇艺[nbz]”,一个影视作品实体“老九门[nz]”,其中,所述标签包括:n-人名、p-地点、o-组织、nbz-商业品牌实体、nt-非实体、nz-影视实体等;该实体的形式可以是词向量的形式,获得所述输入文本对应的词向量,该将所述词向量输入到训练好的CNN模型中进行识别,具体地,可以将词汇表中的词向量映射为多个维矩阵,如,当该词向量的个数为50个时,则将映射为50\*200维的矩阵。

[0127] 实际而言,将该多维矩阵进行卷积,得到不同的卷积向量,从不同的卷积向量得到最大卷积向量,具体地,将所述卷积向量的最大值进行提取,构成最大卷积向量。

[0128] 进一步地,从所述多维矩阵中获得最大卷积向量之后,可以将该最大卷积向量作为对全连接层的输入,获得所述全连接层中的最大概率的类别值。

[0129] 步骤303,获取所述模型输出的所述最大概率的类别值对应的意图类别;

[0130] 本发明实施例中,将该最大卷积向量作为全连接层的输入,计算该全连接层中的

每个类别值的概率大小,选取出最大概率对应的类别值,可以获取所述全连接层中的最大概率的类别值。

[0131] 实际应用到本发明实施例中,因为在训练模型之前已经设置意图类别为全连接层的隐节点,即已经建立意图类别与全连接层的隐节点的对应关系,获取所述模型输出的所述最大概率的类别值,即可以获得该类别值对应的意图类别。

[0132] 步骤304,依据所述意图类别获取所述输入文本的意图词。

[0133] 本发明实施例中,所述依据所述意图类别获取所述输入文本的意图词的步骤包括:将所述意图类别与所述种子词典进行匹配,获得所述意图类别对应的意图词。

[0134] 本发明实施例中,获取输入文本,将所述输入文本对应的词向量输入所述模型;获取所述模型输出的所述最大概率的类别值对应的意图类别;依据所述意图类别获取所述输入文本的意图词,相比于传统的机器学习进行文本意图识别准确率不高的情况,应用了CNN模型后,并考虑到词向量的信息,使得意图识别的准确率得到大幅提升。

[0135] 为使本领域人员更好地理解本发明实施例,以下对模型训练及文本意图识别的具体示例进行说明。

[0136] 参照图4,示出了本发明实施例的模型的训练方法及文本意图的识别方法,该模型可以包括CNN模型。

[0137] 步骤一、分词及实体识别;

[0138] 针对获取的多个文本语料,首先进行分词和实体识别。对分好词的文本,进行实体识别的序列化标注,利用双向长短期记忆-条件随机场层模型(Bi-LSTM-CRF),对每个词word标记一个对应的tag标签。一个具体的应用实例是,对分本标注的标签包括:n-人名,p-地点,o-组织,nbz-商业品牌实体,nt-非实体,等等,得到种子词典(可以由包含意图的动词组成)及词汇表(所有词);Bi-LSTM-CRF模型是一个多层的深度神经网络模型,其基本结构包括:前向LSTM层,后向LSTM层,条件随机场层。本发明实施例中基于Tensorflow开源框架,拓展了Bi-LSTM-CRF模型,在测试语料的实体识别上取得了93.6%的准确率。

[0139] 具体地,Bi-LSTM-CRF模型的参数和输入输出如下:

[0140] LSTM单一结构的公式:

$$[0141] \quad i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i)$$

$$[0142] \quad f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f)$$

$$[0143] \quad c_t = f_t \otimes c_{t-1} + i_t \otimes \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c)$$

$$[0144] \quad o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o)$$

$$[0145] \quad h_t = o_t \otimes \tanh(c_t)$$

[0146] 双向LSTM公式:

[0147] 针对一个词序列,在任意一个时间点t,前向LSTM和后向LSTM的隐藏层进行拼接,得到一个长度为2\*n的词向量:

$$[0148] \quad [f_\theta]_{1,y_1,t} = W_h[\vec{h}_t, \overleftarrow{h}_t]$$

[0149] CRF条件随机场层:

[0150] 输入为任意一个标注序列,定义一个得分score为两部分,即转移Transition Score加上网络的network score。

$$[0151] \quad s([x]_1^T, [y]_1^T, \theta) = \sum_{i=1}^T ([A]_{[y]_{i-1}, [y]_i} + [f_\theta]_{[y]_i, i})$$

[0152] 条件随机场定义损失函数为,目标是求解最小化

$$[0153] \quad C(f_\theta, A) = \min_{\forall [j]_1^T} ([x]_1^T, [j]_1^T, \theta) - s([x]_1^T, [y]_1^T, \theta)$$

[0154] 步骤二、聚类以确定意图类别;

[0155] 针对意图类别,在一个包含种子词的种子词典V,由包含有意图的动词(种子词)构成,如V={下载,充值,买会员,……等等}。

[0156] 针对每个种子词,在词汇表中,找相似度高于一阈值alpha的词,对种子词进行拓展。

[0157] 假设每个词,都可以表示成一个d维的词向量X<sub>i</sub>,计算词表中的任意一个动词和种子词V<sub>i</sub>的cosine相似度,

$$[0158] \quad \text{similarity} = \text{cosine}(X_i, V_i) = \frac{X_i \cdot V_i}{|X_i| \times |V_i|}$$

[0159] 当相似度similarity>alpha阈值时,将X<sub>i</sub>添加入种子词典V,对所有在拓展后的意图动词进行层次聚类,确定一个K个类别,对层次聚类树进行切割,得到K个不同的意图类别,每个类别有该类下语义相关的意图词构成。

[0160] 步骤三、训练CNN模型;

[0161] 1. 嵌入层中将文本分词替换为语向量。如:“为了在爱奇艺看电视剧《老九门》,去给会员充值”有10个分词,将其替换为相应的词向量,最终将映射为10X200维矩阵。

[0162] 2. 卷积层

[0163] 2.1. 设定卷积核大小:分为2-gram卷积核,3-gram卷积,4-gram卷积等等,具体情况选取。2-gram卷积核就是长度方向一次卷积2个分词,宽度是200维。

[0164] 2.2设定不同卷积核的数量:我们可以设定2-gram卷积核有20组,3-gram的卷积核有20组,其他类似。具体值可以根据效果确定。针对2中每组的卷积核的进行扫描计算,我们都可以得到相应的卷积向量。

[0165] 3. 池化层,针对2.2中,获得卷积向量,我们选取每个卷积层的最大值作为最能体现特征的隐节点。为此我们可以得到由各层的最大值构成的一个最大卷积向量。

[0166] 4. 全连接层,利用3步骤中的最大卷积向量作为全连接层的输入,设定2-3层隐含层,最后一层为softmax层(意图类别个数即为softmax隐节点数)输出隐节点对应的类别值,采用多个文本语料训练后,可以得到训练好的CNN模型。

[0167] 步骤四、文本意图的识别。

[0168] 对于训练好的CNN模型,可以按照以下方式应用于文本意图的识别,针对一个输入文本,如“为了在爱奇艺看电视剧《老九门》,去给会员充值”中。文本意图的识别可以遵循下列流程。

[0169] 4.1分词和预处理,将文本分为分词的一个列表。如“为了在爱奇艺看电视剧《老九门》,去给会员充值”。

[0170] 4.2文本意图中的实体识别,利用预先训练好的Bi-LSTM-CRF模型,对文本进行序列化标注,得到[word/tag]对的列表结果,如“为了[nt]在[nt]爱奇艺[nbz]看[nt]电视剧

[nt]《[nt]老九门[nz]》[nt], [nt]去[nt]给[nt]会员[nt]充值[nt]”。一共自动发现了两个实体,一个品牌实体“爱奇艺[nbz]”,一个影视作品实体“老九门[nz]”。

[0171] 4.3利用训练好的CNN模型进行文本意图识别,判定一个输入文本会在多个意图类别中,属于最高概率的类别值对应的意图类别,如属于[账户付费]类。

[0172] 4.4整合结构化信息,抽取出了该输入文本提及了“爱奇艺[nbz]”这个实体,用户的意图类别属于[账户付费]类,与意图词的种子词典匹配后,得到的意图词是“充值[Intention]”。

[0173] 需要说明的是,对于方法实施例,为了简单描述,故将其都表述为一系列的动作组合,但是本领域技术人员应该知悉,本发明实施例并不受所描述的动作顺序的限制,因为依据本发明实施例,某些步骤可以采用其他顺序或者同时进行。其次,本领域技术人员也应该知悉,说明书中所描述的实施例均属于优选实施例,所涉及的动作并不一定是本发明实施例所必须的。

[0174] 参照图5,示出了本发明实施例的一种模型的训练装置实施例的结构框图,具体可以包括如下模块:

[0175] 文本语料获取模块501,用于获取多个文本语料;

[0176] 种子词典及词汇表获得模块502,用于将所述多个文本语料分别进行分词及实体识别,获得种子词典及词汇表;其中,所述词汇表包括多个词向量;

[0177] 聚类模块503,用于针对所述种子词典及所述词汇表进行聚类,获得多个意图类别;

[0178] 第一映射模块504,用于在一次训练中,将一个所述词向量映射为多维矩阵;

[0179] 最大卷积向量获得模块505,用于从所述多维矩阵中获得最大卷积向量;

[0180] 第一输入模块506,用于输入所述最大卷积向量至全连接层;

[0181] 第一类别值输出模块507,用于设置所述意图类别为全连接层的隐节点,输出类别值;

[0182] 模型获得模块508,用于当所述多个词向量训练完毕后,获得经过多次训练的模型。

[0183] 优选地,所述种子词典及词汇表获得模块包括:

[0184] 词汇表获得子模块,用于针对所述文本语料进行分词及实体识别,获得词汇表;

[0185] 种子词典组合子模块,用于从所述词汇表提取出多个种子词,组合成种子词典。

[0186] 优选地,所述聚类模块包括:

[0187] 计算子模块,用于将每个种子词及每个词向量分别进行组合,计算组合后的所述种子词及所述词向量之间的多个特征相似度;

[0188] 判断子模块,用于分别判断各个所述特征相似度是否大于预设阈值;

[0189] 确定子模块,用于当存在大于预设阈值的特征相似度时,则确定该大于预设阈值的所述特征相似度对应的种子词为意图类别,以及,确定该大于预设阈值的所述特征相似度对应的词向量为所述意图类别对应的意图词。

[0190] 优选地,所述最大卷积向量获得模块包括:

[0191] 选取子模块,用于从所述多维矩阵中选取不同数量的卷积核;

[0192] 卷积向量得到子模块,用于针对所述卷积核进行计算,得到卷积向量;

[0193] 最大卷积向量确定子模块,用于确定卷积向量的最大值为最大卷积向量。

[0194] 参照图6,示出了本发明实施例的一种文本意图的识别装置实施例的结构框图,具体可以包括如下模块:

[0195] 输入文本获取模块601,用于获取输入文本;

[0196] 输入模块602,用于将所述输入文本对应的词向量输入所述模型;

[0197] 意图类别输出模块603,用于获取所述模型输出的所述最大概率的类别值对应的意图类别;

[0198] 意图词获取模块604,用于依据所述意图类别获取所述输入文本的意图词。

[0199] 参照图7,其示出了本发明实施例提供的终端设备的结构示意图。该电子设备用于实施上述实施例中提供的模型的训练方法及文本意图的识别方法,具体来讲:

[0200] 电子设备700可以包括RF (Radio Frequency, 射频) 电路710、包括有一个或一个以上计算机可读存储介质的存储器720、输入单元730、显示单元740、传感器750、音频电路760、短距离无线传输模块770、包括有一个或者一个以上处理核心的处理器780、以及电源790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。其中:

[0201] RF电路710可用于收发信息或通话过程中,信号的接收和发送,特别地,将基站的下行信息接收后,交由一个或者一个以上处理器780处理;另外,将涉及上行的数据发送给基站。通常,RF电路710包括但不限于天线、至少一个放大器、调谐器、一个或多个振荡器、收发信机、耦合器、LNA (Low Noise Amplifier, 低噪声放大器)、双工器等。此外,RF电路710还可以通过无线通信与网络和其他设备通信。所述无线通信可以使用任一通信标准或协议,包括但不限于GSM (Global System of Mobile communication, 全球移动通讯系统)、GPRS (General Packet Radio Service, 通用分组无线服务)、CDMA (Code Division Multiple Access, 码分多址)、WCDMA (Wideband Code Division Multiple Access, 宽带码分多址)、LTE (Long Term Evolution, 长期演进)、电子邮件、SMS (Short Messaging Service, 短消息服务) 等。存储器720可用于存储软件程序以及模块,比如,存储器720可以用于存储采集语音信号的软件程序、实现关键词识别的软件程序、实现连续语音识别的软件程序以及实现设置提醒事项的软件程序等等。处理器780通过运行存储在存储器720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,比如本发明实施例中“获取多个文本语料;将所述多个文本语料分别进行分词及实体识别,获得种子词典及词汇表;其中,所述词汇表包括多个词向量;针对所述种子词典及所述词汇表进行聚类,获得多个意图类别;在一次训练中,将一个所述词向量映射为多维矩阵;从所述多维矩阵中获得最大卷积向量;输入所述最大卷积向量至全连接层;设置所述意图类别为全连接层的隐节点,输出类别值;当所述多个词向量训练完毕后,获得经过多次训练的模型”的功能等等。存储器720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700的使用所创建的数据(比如音频数据、电话本等)等。此外,存储器720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720还可以包括存储器控制器,以提供处理器780和输入单元730对存储器720的访问。

[0202] 输入单元730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。具体地,输入单元730可包括触敏表面731以及其他输入设备732。触敏表面731,也称为触摸显示屏或者触控板,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触敏表面731上或在触敏表面731附近的操作),并根据预先设定的程式驱动相应的连接装置。可选的,触敏表面731可包括触摸检测装置和触摸控制器两个部分。其中,触摸检测装置检测用户的触摸方位,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780,并能接收处理器780发来的命令并加以执行。此外,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触敏表面731。除了触敏表面731,输入单元730还可以包括其他输入设备732。具体地,其他输入设备732可以包括但不限于物理键盘、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。

[0203] 显示单元740可用于显示由用户输入的信息或提供给用户的信息以及电子设备700的各种图形用户接口,这些图形用户接口可以由图形、文本、图标、视频和其任意组合来构成。显示单元740可包括显示面板741,可选的,可以采用LCD(Liquid Crystal Display,液晶显示器)、OLED(Organic Light-Emitting Diode,有机发光二极管)等形式来配置显示面板741。进一步的,触敏表面731可覆盖在显示面板741之上,当触敏表面731检测到在其上或附近的触摸操作后,传送给处理器780以确定触摸事件的类型,随后处理器780根据触摸事件的类型在显示面板741上提供相应的视觉输出。虽然在图7中,触敏表面731与显示面板741是作为两个独立的部件来实现输入和输入功能,但是在某些实施例中,可以将触敏表面731与显示面板741集成而实现输入和输出功能。

[0204] 电子设备700还可包括至少一种传感器750,比如光传感器、运动传感器以及其他传感器。具体地,光传感器可包括环境光传感器及接近传感器,其中,环境光传感器可根据环境光线的明暗来调节显示面板741的亮度,接近传感器可在电子设备700移动到耳边时,关闭显示面板741和/或背光。作为运动传感器的一种,重力加速度传感器可检测各个方向上(一般为三轴)加速度的大小,静止时可检测出重力的大小及方向,可用于识别手机姿态的应用(比如横竖屏切换、相关游戏、磁力计姿态校准)、振动识别相关功能(比如计步器、敲击)等;至于电子设备700还可配置的陀螺仪、气压计、湿度计、温度计、红外线传感器等其他传感器,在此不再赘述。

[0205] 音频电路760、扬声器761,传声器762可提供用户与电子设备700之间的音频接口。音频电路760可将接收到的音频数据转换后的电信号,传输到扬声器761,由扬声器761转换为声音信号输出;另一方面,传声器762将收集的声音信号转换为电信号,由音频电路760接收后转换为音频数据,再将音频数据输出处理器780处理后,经RF电路710以发送给另一终端,或者将音频数据输出至存储器720以便进一步处理。音频电路760还可能包括耳塞插孔,以提供外设耳机与电子设备700的通信。

[0206] 短距离无线传输模块770可以是WIFI(wireless fidelity,无线保真)模块或者蓝牙模块等。电子设备700通过短距离无线传输模块770可以帮助用户收发电子邮件、浏览网页和访问流式媒体等,它为用户提供了无线的宽带互联网访问。虽然图7示出了短距离无线传输模块770,但是可以理解的是,其并不属于电子设备700的必须构成,完全可以根据需要

在不改变发明的本质的范围内而省略。

[0207] 处理器780是电子设备700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720内的软件程序和/或模块,以及调用存储在存储器720内的数据,执行电子设备700的各种功能和处理数据,从而对电子设备进行整体监控。可选的,处理器780可包括一个或多个处理核心;优选的,处理器780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780中。

[0208] 电子设备700还包括给各个部件供电的电源790(比如电池),优选的,电源可以通过电源管理系统与处理器780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790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。

[0209] 尽管未示出,电子设备700还可以包括摄像头、蓝牙模块等,在此不再赘述。具体在本实施例中,电子设备700的显示单元是触摸屏显示器。

[0210] 本说明书中的各个实施例均采用递进的方式描述,每个实施例重点说明的都是与其他实施例的不同之处,各个实施例之间相同相似的部分互相参见即可。

[0211] 本领域内的技术人员应明白,本发明实施例的实施例可提供为方法、装置、或计算机程序产品。因此,本发明实施例可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本发明实施例可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0212] 本发明实施例还公开了一种计算机可读存储介质,其上存储有计算机程序,该程序被处理器执行时可以实现图1和/或图2和/或图3所述方法。

[0213] 对于装置实施例而言,由于其与方法实施例基本相似,所以描述的比较简单,相关之处参见方法实施例的部分说明即可。

[0214] 本说明书中的各个实施例均采用递进的方式描述,每个实施例重点说明的都是与其他实施例的不同之处,各个实施例之间相同相似的部分互相参见即可。

[0215] 本领域内的技术人员应明白,本发明实施例的实施例可提供为方法、装置、或计算机程序产品。因此,本发明实施例可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本发明实施例可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0216] 本发明实施例是参照根据本发明实施例的方法、终端设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理终端设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理终端设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0217] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理终端设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品,该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

[0218] 这些计算机程序指令也可装载到计算机或其他可编程数据处理终端设备上,使得在计算机或其他可编程终端设备上执行一系列操作步骤以产生计算机实现的处理,从而在计算机或其他可编程终端设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0219] 尽管已描述了本发明实施例的优选实施例,但本领域内的技术人员一旦得知了基本创造性概念,则可对这些实施例做出另外的变更和修改。所以,所附权利要求意欲解释为包括优选实施例以及落入本发明实施例范围的所有变更和修改。

[0220] 最后,还需要说明的是,在本文中,诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来,而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者终端设备不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者终端设备所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括所述要素的过程、方法、物品或者终端设备中还存在另外的相同要素。

[0221] 以上对本发明所提供的一种模型的训练方法、一种文本意图的识别方法和一种模型的训练装置、一种文本意图的识别装置,进行了详细介绍,本文中应用了具体个例对本发明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发明的方法及其核心思想;同时,对于本领域的一般技术人员,依据本发明的思想,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发明的限制。

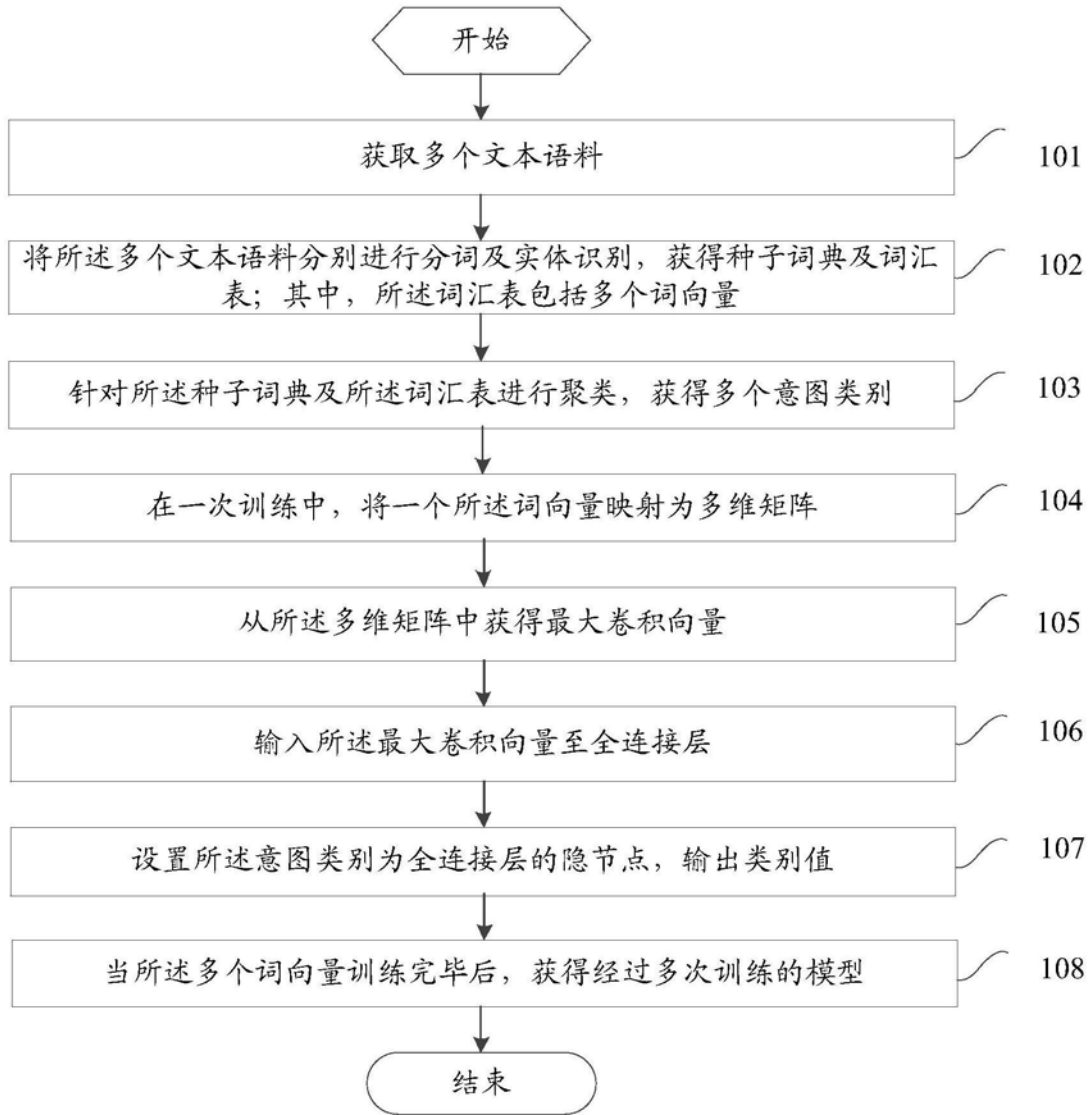


图1

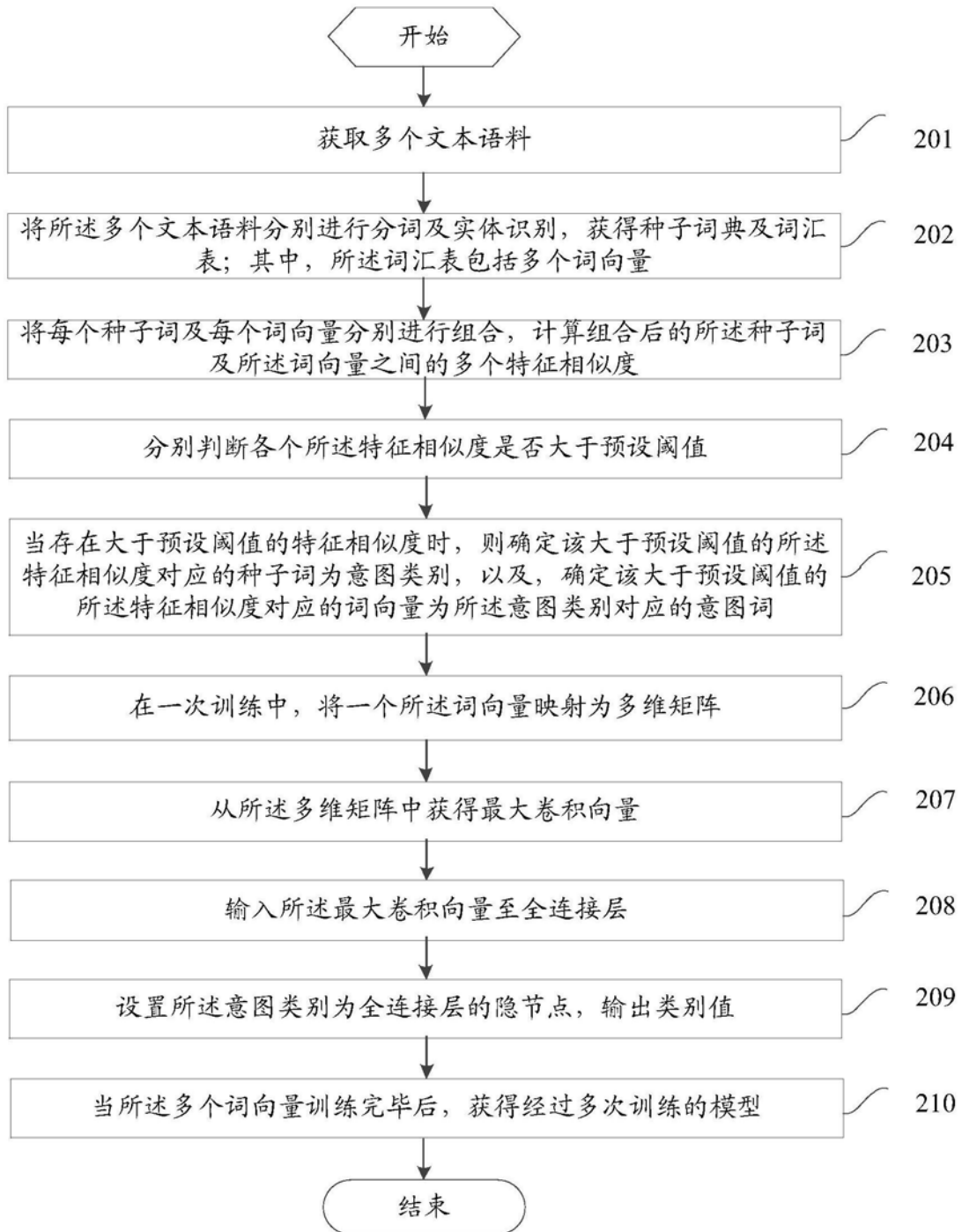


图2

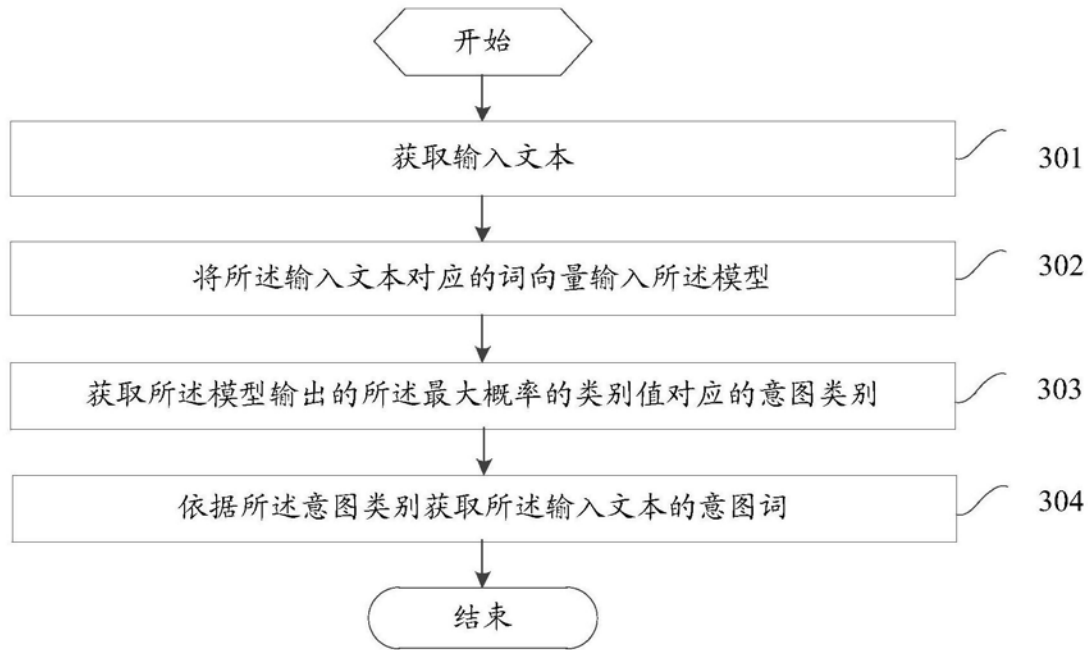


图3

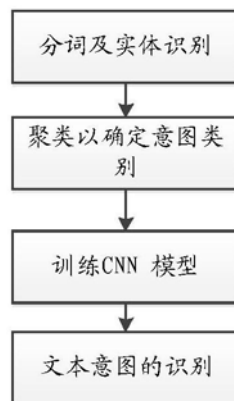


图4



图5

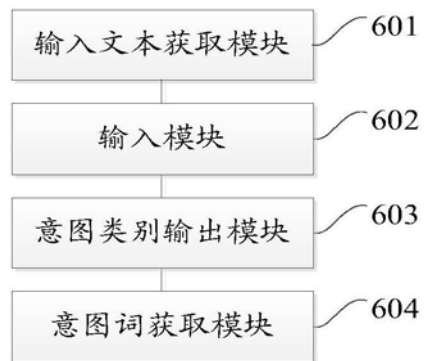


图6

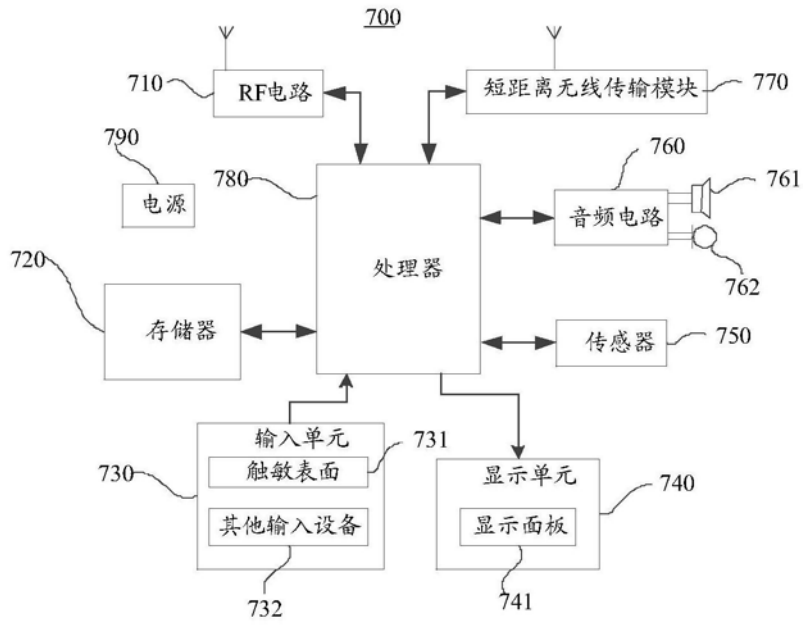


图7