



(12) 发明专利

(10) 授权公告号 CN 114818665 B

(45) 授权公告日 2023.05.12

(21) 申请号 202210432349.7

G06F 40/284 (2020.01)

(22) 申请日 2022.04.22

G06F 40/30 (2020.01)

G06F 18/241 (2023.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 114818665 A

(56) 对比文件

CN 110532558 A, 2019.12.03

(43) 申请公布日 2022.07.29

审查员 何俊伟

(73) 专利权人 电子科技大学

地址 611731 四川省成都市高新区(西区)

西源大道2006号

(72) 发明人 陈波 朱舜文 曾俊涛 陈圩钦

邓媛丹 王庆先

(74) 专利代理机构 成都先导云创知识产权代理

事务所(普通合伙) 51321

专利代理师 李坤

(51) Int. Cl.

G06F 40/216 (2020.01)

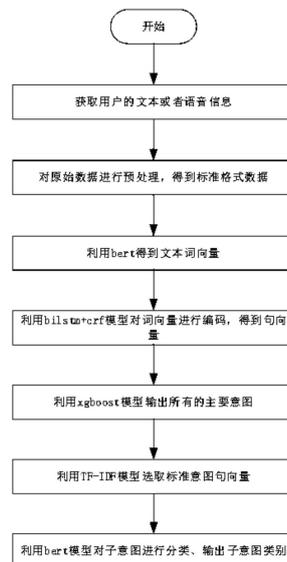
权利要求书1页 说明书5页 附图4页

(54) 发明名称

一种基于bert+ilstm+crf与xgboost模型的多意图识别方法和系统

(57) 摘要

本发明属于自然语言理解领域技术,具体涉及一种基于bert+ilstm+crf与xgboost模型的多意图识别方法和系统;本技术方案中,使用bert处理预处理完成的数据集,得到动态的词向量,不同于以往的使用word2vec或glove模型获取的词向量。Bert模型输出的词向量具有动态特性,能够解决一词多义的问题。词向量再经过biilstm+crf转换为句向量,biilstm+crf模型能够同时处理距离较远的上下文文本信息,通过近邻标签的关系得到最优的句向量预测序列。在主意图识别方面使用Xgboost模型,该模型的识别精度较高且更加灵活,因此用在主意图。得到所有的主意图过后,我们利用TF-IDF模型选取标准意图,以此为意图判断依据。将经由bert+biilstm+crf模型处理过后的句向量输入到新的bert模型当中,最终输出子意图。



1.一种基于bert+bilstm+crf与xgboost模型的多意图识别方法,其特征在于,包括如下步骤:

步骤1:通过得到用户的交互文本或者语音信息,构建数据集;

步骤2:对所述数据集进行预处理,得到标准格式数据;

步骤3:通过bert+bilstm+crf模型将所述标准格式数据转换为特征句向量;

步骤4:通过Xgboost模型训练相应的特征句向量模型来进行意图识别,识别用户交互意图并输出所有的主意图;

步骤5:利用TF-IDF模型计算同一主意图的所有文本数据中的意图对该类意图的贡献,确定标准意图,其他意图则为子意图,将标准意图的句向量作为标准句向量;

步骤6:通过bert模型将各子意图进行分类并输出个子意图类别。

2.根据权利要求1所述的一种基于bert+bilstm+crf与xgboost模型的多意图识别方法,其特征在于,所述步骤2中的预处理包括对数据集进行去停用词和标注。

3.根据权利要求1所述的一种基于bert+bilstm+crf与xgboost模型的多意图识别方法,其特征在于,所述特征句向量包括词向量、词性向量和命名实体向量。

4.根据权利要求3所述的一种基于bert+bilstm+crf与xgboost模型的多意图识别方法,其特征在于,所述步骤3中具体包括;利用了bert模型将标准格式数据转换为词向量,再由bilstm+crf模型计算出词性向量和命名实体向量。

5.根据权利要求1所述的一种基于bert+bilstm+crf与xgboost模型的多意图识别方法,其特征在于,所述bert+bilstm+crf模型的公式如下:

$$(1) \text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V)$$

$$(2) \text{Attention}(Q, K, V) = \text{Softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V$$

$$(3) \text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_n) W^O$$

式中:Q、K、V为字向量矩阵, d_k 是嵌入层维度,而MultiHead自注意力机制则是通过对Q、K、V进行投影,将自注意力机制结果拼接,如公式(2)和公式(3)。

6.一种基于bert+bilstm+crf与xgboost模型的多意图识别系统,用于执行权利要求1-5中任一所述的方法,其特征在于,包括语音接收模块,该模块的功能是用于识别用户语音;语音转换文本模块,用于接收语音之后的文本转换功能;意图识别模块,用于基于bert+bilstm+crf所述的标准意图选取,多意图识别;交互模块,用于识别用户意图后执行意图。

7.一种电子设备,其特征在于,所述电子设备包括:壳体、处理器、存储器、电路板和电源电路,其中,电路板安置在壳体围成的空间内部,处理器和存储器设置在电路板上;电源电路,用于为上述电子设备的各个电路或器件供电;存储器用于存储可执行程序代码;处理器通过读取存储器中存储的可执行程序代码来运行与可执行程序代码对应的程序,用于执行权利要求1-5中任一所述的方法。

8.一种计算机可读存储介质,其特征在于,所述计算机可读存储介质存储有一个或者多个程序,所述一个或者多个程序可被一个或者多个处理器执行,以实现权利要求1-5中任一所述的方法。

一种基于bert+bi lstm+crf与xgboost模型的多意图识别方法和系统

技术领域

[0001] 本发明属于自然语言理解领域技术,具体涉及一种基于bert+bi lstm+crf与xgboost模型的多意图识别方法和系统。

背景技术

[0002] 意图识别主要指在人与机器交互动作中,机器对用户所发出的语音或者文本进行自然语言理解,判断出用户的真实意图,为用户提供准确的服务。

[0003] 目前,意图识别大部分应用在应对用户单一意图的分类或者匹配用途。单一用途,顾名思义,指用户所发出的文本或者语音有且仅有一个意图,而在另一些情况下,用户会发出的语音或者文本交互含有多个意图时,在处理这类交互时,单一意图识别就会遇到困难。

[0004] 为了达到实现多意图识别,目前主流方式是将用户发出的指令拆分。但是拆分方式仅仅是通过指令的表层,如标点符号拆分、动词划分等方式。但是如果用户输入的是语音指令,或者单个动词不能完全表达用户的意图。

发明内容

[0005] 针对现有技术存在的传统的单一意图识别无法满足用户多意图语言指令下的要求,而传统的多意图识别拆分句子的方案并不能从根本上将一个多意图句子拆分程多个单意图句子的问题,本发明的目的在于提供一种基于bert+bi lstm+crf与xgboost模型的多意图识别方法,通过对句子分析,解析出句子里所包含的多意图信息;本发明一方面使用bert+bi lstm+crf、Xgboost模型进行主意图识别。本发明另一方面使用bert+bi lstm进行意图分类,能够有效的将意图划分到与之对应的类别。

[0006] 本发明采用的技术方案如下:

[0007] 一种基于bert+bi lstm+crf与xgboost模型的多意图识别方法,其特征在于,包括如下步骤:

[0008] 步骤1:通过得到用户的交互文本或者语音信息,构建数据集;

[0009] 步骤2:对所述数据集进行预处理,得到标准格式数据;

[0010] 步骤3:通过bert+bi lstm+crf模型将所述标准格式数据转换为特征句向量;

[0011] 步骤4:通过Xgboost模型训练相应的特征句向量模型来进行意图识别,识别用户交互意图并输出所有的主意图;

[0012] 步骤5:利用TF-IDF模型计算同一主意图的所有文本数据中的意图对该类意图的贡献,确定标准意图,其他意图则为子意图,将标准意图的句向量作为标准句向量;

[0013] 步骤6:通过bert模型将各子意图进行分类并输出个子意图类别。

[0014] 采用上述技术方案,使用bert处理预处理完成的数据集,得到动态的词向量,不同于以往的使用word2vec或glove模型获取的词向量。Bert模型输出的词向量具有动态特性,能够解决一词多义的问题。词向量再经过bi lstm+crf转换为句向量,bi lstm+crf模型能够

同时处理距离较远的上下文文本信息,通过近邻标签的关系得到最优的句向量预测序列。在主意图识别方面使用Xgboost模型,该模型的识别精度较高且更加灵活,因此用在主意图。得到所有的主意图过后,我们利用TF-IDF模型选取标准意图,以此为意图判断依据。将经由bert+bilstm+crf模型处理过后的句向量输入到新的bert模型当中,最终输出子意图。

[0015] 进一步,所述步骤2中的预处理包括对数据集进行去停用词和标注。

[0016] 进一步,所述特征句向量包括词向量、词性向量和命名实体向量。

[0017] 进一步,所述步骤3中具体包括;利用了bert模型将标准格式数据转换为词向量,再由bilstm+crf模型计算出词性向量和命名实体向量,通过Xgboost模型输出用户所有的主要意图。

[0018] 进一步,所述bert+bilstm+crf模型的公式如下:

[0019] (1) $head_i = Attention(QW_i^Q, KW_i^K, VW_i^V)$

[0020] (2) $Attention(Q, K, V) = Soft\ max\left(\frac{QK^T}{\sqrt{d_k}}\right)V$

[0021] (3) $MultiHead(Q, K, V) = Concat(head_1, \dots, head_n)W^O$

[0022] 式中:Q、K、V为字向量矩阵, d_k 是嵌入层维度,而MultiHead自注意力机制则是通过对Q、K、V进行投影,将自注意力机制结果拼接,如公式(2)和公式(3)。

[0023] 一种基于bert+bilstm+crf与xgboost模型的多意图识别系统,用于执行上述多意图识别方法,包括语音接收模块,该模块的功能是用于识别用户语音;语音转换文本模块,用于接收语音之后的文本转换功能;意图识别模块,用于基于bert+bilstm+crf所述的标准意图选取,多意图识别;交互模块,用于识别用户意图后执行意图。

[0024] 一种电子设备,所述电子设备包括:壳体、处理器、存储器、电路板和电源电路,其中,电路板安置在壳体围成的空间内部,处理器和存储器设置在电路板上;电源电路,用于为上述电子设备的各个电路或器件供电;存储器用于存储可执行程序代码;处理器通过读取存储器中存储的可执行程序代码来运行与可执行程序代码对应的程序,用于执行上述多意图识别方法。

[0025] 一种计算机可读存储介质,所述计算机可读存储介质存储有一个或者多个程序,所述一个或者多个程序可被一个或者多个处理器执行,以实现上述多意图识别方法。

[0026] 综上所述,由于采用了上述技术方案,本发明的有益效果是:

[0027] 1) 使用bert+bilstm+crf模型训练中得到的词向量能够动态的表示一次多意,能够充分利用上下文语义信息,较传统的bert模型能够通过近邻标签关系得到最优化的句向量序列;

[0028] 2) 使用Xgboost模型精确地输出用户的所有的主要意图;

[0029] 3) 使用TF-IDF模型作为标准意图选择模型;四、使用bert模型对子意图进行分类提高分类准确率。

附图说明

[0030] 本发明将通过例子并参照附图的方式说明,其中:

[0031] 图1是本发明中一种基于bert+bilstm+crf与Xgboost模型多意图识别方法流程

图；

[0032] 图2是本发明具体实施方式提供的bert模型的结构示意图；

[0033] 图3是本发明具体实施方式提供的bert+bilstm+crf融合模型的结构示意图；

[0034] 图4是本发明具体实施方式提供的一种电子设备的结构示意图。

具体实施方式

[0035] 为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附图，对本申请实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅是本申请一部分实施例，而不是全部的实施例。通常在此处附图中描述和示出的本申请实施例的组件可以各种不同的配置来布置和设计。因此，以下对在附图中提供的本申请的实施例的详细描述并非旨在限制要求保护的本申请的范围，而是仅仅表示本申请的选定实施例。基于本申请的实施例，本领域技术人员在没有做出创造性劳动的前提下所获得的所有其他实施例，都属于本申请保护的范围。

[0036] 在本申请实施例的描述中，需要说明的是，术语“上”、“下”、“左”、“右”、“竖直”、“水平”、“内”、“外”等指示的方位或位置关系为基于附图所示的方位或位置关系，或者是该发明产品使用时惯常摆放的方位或位置关系，仅是为了便于描述本申请和简化描述，而不是指示或暗示所指的装置或元件必须具有特定的方位、以特定的方位构造和操作，因此不能理解为对本申请的限制。此外，术语“第一”、“第二”、“第三”等仅用于区分描述，而不能理解为指示或暗示相对重要性。

[0037] 下面结合图1~图4对本发明作详细说明。

[0038] 如图1所示，为本发明一种基于bert+bilstm+crf与xgboost模型的多意图识别方法的实施例，包括：

[0039] 一、通过得到用户的交互文本或者语音信息，构建数据集；

[0040] 二、对数据集进行预处理，包括去停用词、标注等，最后得到标准格式数据；

[0041] 三、使用bert模型将数据转换为词向量，再由bilstm+crf模型计算出句向量、命名实体向量等；首先通过分词处理得到分词文本序列；然后对分词序列的部分词进行全词Mask，再为序列的开头添加一个特殊标记[CLS]，句子间用标记[SEP]分隔。此时序列的每个词的输出Embedding由3部分组成：Token Embedding、Segment Embedding和Position Embedding。将序列向量输入到双向Transformer进行特征提取，最后得到含有丰富语义特征的序列向量。

[0042] 四、使用Xgboost模型训练相应的句向量模型来进行意图识别，Xgboost模型将输入的句向量进行样本特征划分，再通过权重机制进行用户交互意图的重要性评估，输出所有的主意图；

[0043] 五、选取标准句向量，利用TF-IDF模型计算同一主意图的所有文本数据中的意图对该类意图的贡献，确定标准意图；TF-IDF是一种统计方法，用于评估一字词对于一个文件集或语料库中的一份文件的重要程度。字词的重要性随着他在文件中出现的次数成正比增加，但同时也会随着它在语料库中出现的频率成反比下降。TF-IDF的主要思想是：如果某个单词在一篇文章中出现的频率TF高，并且在其他文章中很少出现，则认为此词或短语具有很好的类别区分能力，适合用来分类。

[0044] TF是词频(Term Frequency)

$$[0045] \quad tf_{ij} = \frac{n_{i,j}}{\sum_k n_{k,j}}$$

$$[0046] \quad TF_w = \frac{\text{在某一类中词条}w\text{出现的次数}}{\text{该类中所有的词条数目}}$$

[0047] 其中, $n_{i,j}$ 是该词在文章 d_j 中出现的次数,分母则是文章 d_j 中所有词汇出现的次数总和。

[0048] IDF是逆向文件频率(Inverse Document Frequency),某一特定词语的IDF,可以由总文章数除以包含该词语的文件数目,再将得到的商取对数得到。

$$[0049] \quad idf_i = \log \frac{|D|}{|\{j: t_i \in d_j\}|}$$

[0050] 如果包含词条 t 的文档越少, IDF越大,说明词条具有很好的类别区分能力。

[0051] 其中 $|D|$ 是语料库中的文章总数。 $|\{j: t_i \in d_j\}|$ 表示包含词语 t_i 的文章数目如果该词语不再语料库中,就会导致分母为0,因此一般情况下使用 $1+|\{j: t_i \in d_j\}|$

[0052] 则有 $TF-IDF = TF * IDF$

[0053] 六、使用bert模型将各子意图进行分类、输出个子意图类别。

[0054] 图2为bert模型,主要将标准文本语料输入其中得到最终的词向量,对于bert而言最关键的部分是transformers结构,而该编码器的关键部分则是自注意力机制,通过同一个句子内词之间的相关度来调控权重系数矩阵最终获得词向量。

$$[0055] \quad head_1 = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \quad (1)$$

[0056] 其中, Q, K, V 为字向量矩阵, d_k 是嵌入层维度,而MultiHead自注意力机制则是通过对 Q, K, V 进行投影,将自注意力机制结果拼接,如公式(2)和公式(3)

$$[0057] \quad \text{Attention}(Q, K, V) = \text{Soft max} \left(\frac{QK^T}{\sqrt{d_k}} \right) V \quad (2)$$

$$[0058] \quad \text{MultiHed}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_n) W^0 \quad (3)$$

[0059] 图3为bert模型与bilstm+crf模型结合, bilstm模型使用前向与后向lstm网络合并的方式,对于每一时刻都包含了前向与后向的信息。crf模块能够有效处理bilstm所带来无法处理近邻标签之间关系的问题,输出最优句向量序列。

[0060] 我们用本发明实例与传统方式的多意图识别方式进行对比:

指标/技术	本技术方法	传统方法
Recall	84.2%	80.7%
Precision	85.6%	81.3%
f1-score	84.4%	81.1%

[0062] 本发明实施例另一方面还提供一种基于bert+bilstm+crf与xgboost模型的多意图识别系统,用于执行前述任一实施例所述的方法,包括语音接收模块,该模块的功能是用于识别用户语音;语音转换文本模块,用于接收语音之后的文本转换功能;意图识别模块,

用于基于bert+bilstm+crf所述的标准意图选取,多意图识别;交互模块,用于识别用户意图后执行意图。

[0063] 本发明实施例另一方面还提供一种电子设备,如图4所示,可以实现本发明图1所示实施例的流程,如图4所示,上述电子设备可以包括:壳体、处理器、存储器、电路板和电源电路,其中,电路板安置在壳体围成的空间内部,处理器和存储器设置在电路板上;电源电路,用于为上述电子设备的各个电路或器件供电;存储器用于存储可执行程序代码;处理器通过读取存储器中存储的可执行程序代码来运行与可执行程序代码对应的程序,用于执行前述任一实施例所述的方法。

[0064] 处理器对上述步骤的具体执行过程以及处理通过运行可执行程序代码来进一步执行的步骤,可以参见本发明图1所示实施例的描述,在此不再赘述。

[0065] 该电子设备以多种形式存在,包括但不限于:

[0066] (1) 移动通信设备:这类设备的特点是具备移动通信功能,并且以提供话音、数据通信为主要目标。这类终端包括:智能手机(例如iPhone)、多媒体手机、功能性手机,以及低端手机等。

[0067] (2) 超移动个人计算机设备:这类设备属于个人计算机的范畴,有计算和处理功能,一般也具备移动上网特性。这类终端包括:PDA、MID和UMPC设备等,例如iPad。

[0068] (3) 便携式娱乐设备:这类设备可以显示和播放多媒体内容。该类设备包括:音频、视频播放器(例如iPod),掌上游戏机,电子书,以及智能玩具和便携式车载导航设备。

[0069] (4) 服务器:提供计算服务的设备,服务器的构成包括处理器、硬盘、内存、系统总线等,服务器和通用的计算机架构类似,但是由于需要提供高可靠的服务,因此在处理能力、稳定性、可靠性、安全性、可扩展性、可管理性等方面要求较高。

[0070] (5) 其他具有数据交互功能的电子设备。

[0071] 本发明的实施例另一方面还提供一种计算机可读存储介质,所述计算机可读存储介质存储有一个或者多个程序,所述一个或者多个程序可被一个或者多个处理器执行,以实现前述基于知识图谱的多意图识别方法。

[0072] 本发明针对用户提问的输入文本,根据知识图谱自动进行多意图识别,不局限于标点、句式和句法分析,且对意图自动进行切分,并进行答案生成;用户意图不清时,可在自行推理后,通过自动反问实现意图收敛,有效提高问题识别率和准确率,极大提升问答机器人的灵活性,对话自然。

[0073] 最后应说明的是:以上各实施例仅用以说明本发明的技术方案,而非对其限制;尽管参照前述各实施例对本发明进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分或者全部技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本发明各实施例技术方案的范围,其均应涵盖在本发明的权利要求和说明书的范围当中。

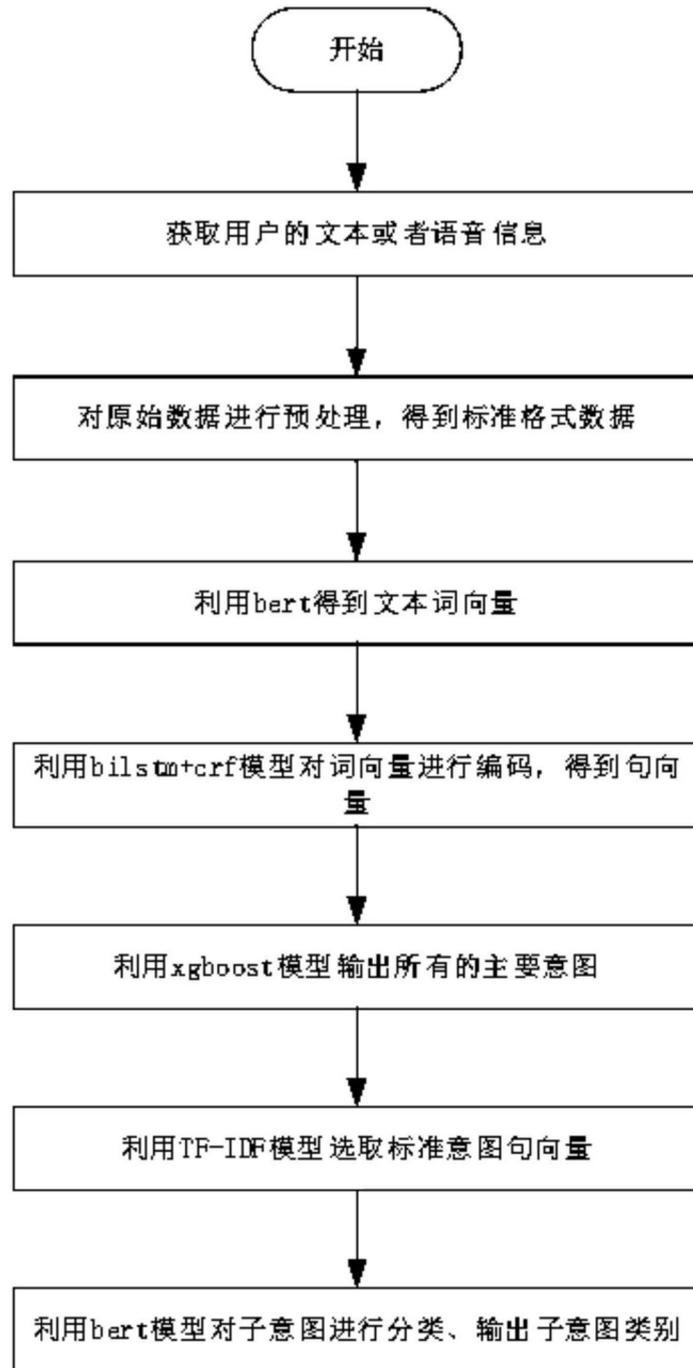


图1

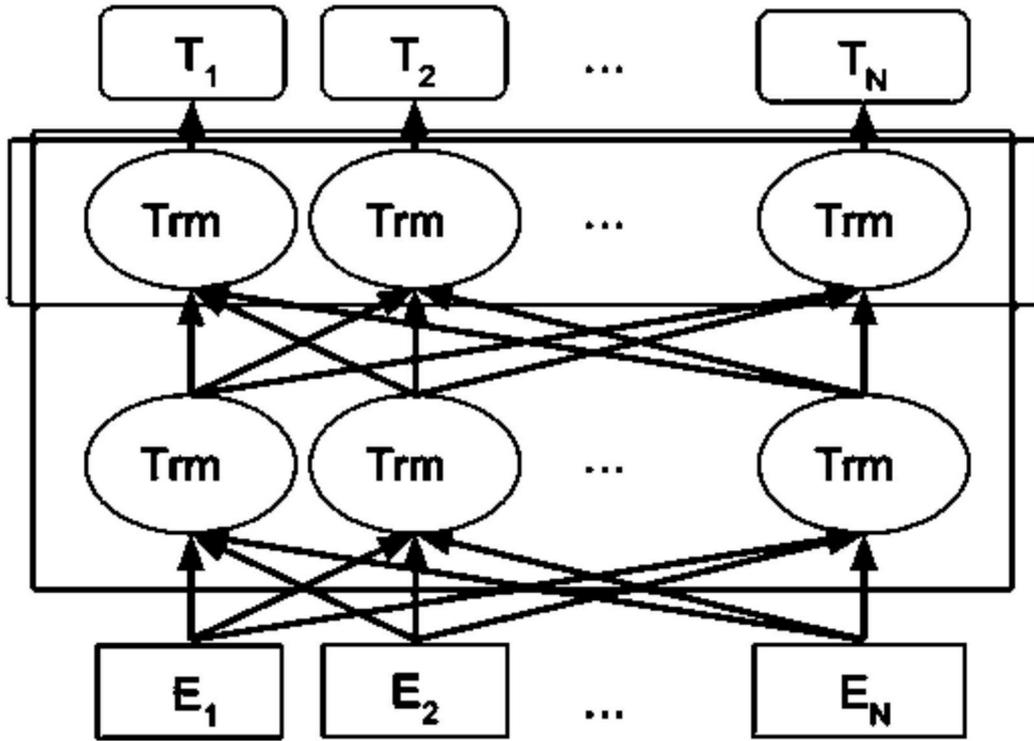


图2

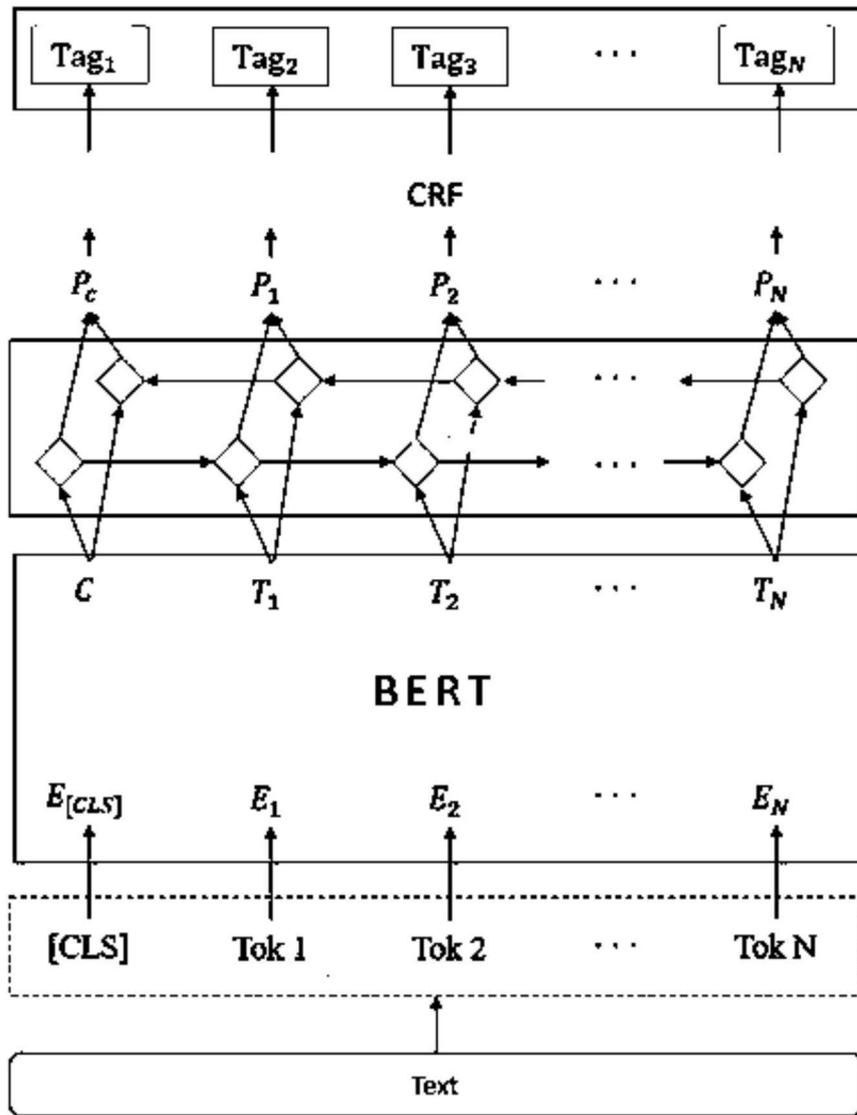


图3

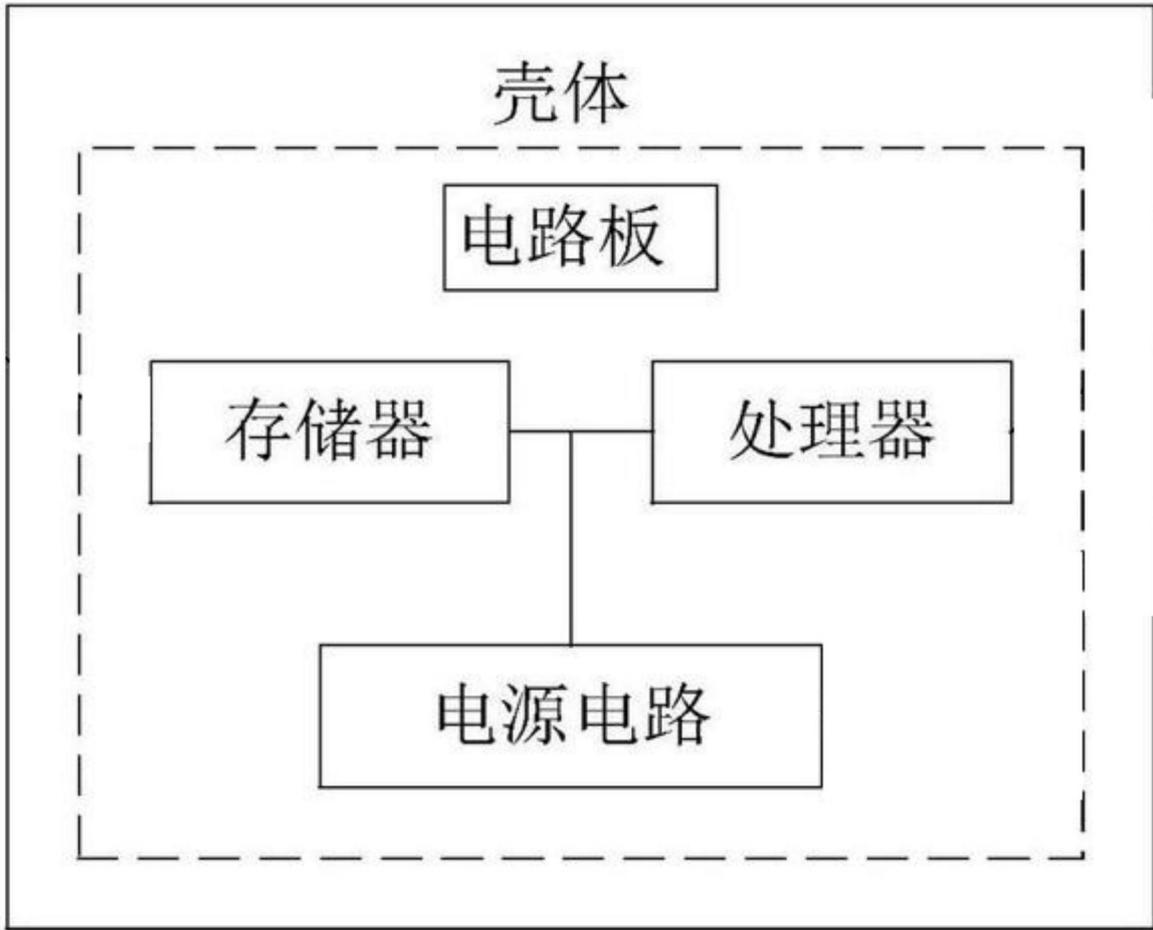


图4