



(12) 发明专利

(10) 授权公告号 CN 113420552 B

(45) 授权公告日 2023. 09. 15

(21) 申请号 202110788532.6

(22) 申请日 2021.07.13

(65) 同一申请的已公布的文献号

申请公布号 CN 113420552 A

(43) 申请公布日 2021.09.21

(73) 专利权人 华中师范大学

地址 430000 湖北省武汉市珞喻路152号

(72) 发明人 赵卫中

(74) 专利代理机构 成都明涛智创专利代理有限

公司 51289

专利代理师 刘佩玲

(51) Int. Cl.

G06F 40/284 (2020.01)

G06F 40/205 (2020.01)

G06N 20/00 (2019.01)

(56) 对比文件

CN 110347807 A, 2019.10.18

CN 111797241 A, 2020.10.20

CN 112069811 A, 2020.12.11

US 2012150531 A1, 2012.06.14

胡扬; 闫宏飞; 陈翀. 面向金融知识图谱的实体和关系联合抽取算法. 重庆理工大学学报(自然科学). 2020, (第05期), 全文.

审查员 姜涛

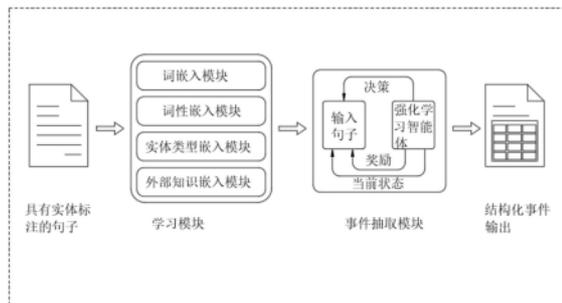
权利要求书1页 说明书5页 附图1页

(54) 发明名称

一种基于强化学习的生物医学多事件抽取方法

(57) 摘要

本发明公开了一种基于强化学习的生物医学多事件抽取方法, 将具有实体标注的句子通过词嵌入模块导入学习模块中, 并对词进行预先训练, 本发明的有益效果是: 通过利用强化学习框架建模生物医学事件抽取的两个子任务, 将两个任务分别作为主要任务和辅助任务, 并通过强化学习的奖励机制建模了两个任务之间的交互关系, 提高了事件抽取效率; 通过利用强化学习框架, 将已抽取的事件映射为向量保存在环境信息中, 可以很好的帮助后续的事件抽取任务, 从而提高模型在多事件抽取任务中的表现; 通过在学习模块中引入外部知识库, 有效地获取了这些特殊词语的外部知识库信息, 从而加强了文本的表示, 有利于后续的事件抽取任务。



1. 一种基于强化学习的生物学多事件抽取方法,其特征在于,包括以下步骤:

S1、将具有实体标注的句子通过词嵌入模块导入学习模块中,并对词进行预先训练,然后通过词性嵌入模块对句子中的每一个单词进行词性标注,并将其映射为向量,通过实体类型嵌入模块对标注后单词 X_t 的实体类型查找,并对其进行向量分配,然后将获取到的外部知识映射为一个向量导入外部知识嵌入模块中,通过学习模块将输入的句子 X 转换成一个新的词序列 $(W_1, \dots, W_2, \dots, W_n)$,并输入到事件抽取模块中;

S2、通过强化学习智能体依次扫描输入句子中的单词,并根据RL中的随机策略将事件类型分配给相应的单词,当强化学习agent识别出一个触发词后,将这个单词的词向量和当前的环境信息联接作为当前的状态 S_t ,并将作为agent的输入,然后agent根据当前的状态信息,输出决策行动 a_t ,表示对于当前触发词,模型预测出的事件类型;

S3、将表示行动的向量 a_t 和表示当前触发词的向量 w_t 与句子中的每个词向量联接得到一个新的句子表示 $X = (W_1, \dots, W_2, \dots, W_n)$,将其作为一个BiLSTM-CRF模块的输入,完成给定触发词的元素检测;

S4、模型预测的结果与训练集的真实标注将用于计算一个奖励值,用来评估当前模型采取的决策的好坏,同时,预测的结果还被映射为一个向量 L_t 并保存在环境信息中,最后对结构化事件进行输出;

所述学习模块由词嵌入模块、词性嵌入模块、实体类型嵌入模块和外部知识嵌入模块组成;

所述词嵌入模块用于通过查找预先训练好的词,并将其嵌入模型,同时对初始化词嵌入向量,并在整个框架的训练过程中进行微调;

所述词性嵌入模块用于通过斯坦福标注工具对给定句子中的每个单词进行词性标注,并将其映射为向量,同时在模型训练过程中进行微调;

所述实体类型嵌入模块用于根据实体类型查找嵌入表,并对每个单词分 X_t 配一个随机初始化的向量,并在整个框架的训练过程中进行微调;

所述外部知识嵌入模块采用基因本体和蛋白质本体作为外部知识库,并针对生物学事件进行任务抽取,同时将获取到的外部知识映射为一个向量,并将该向量知识嵌入到外部知识库中。

2. 根据权利要求1所述的一种基于强化学习的生物学多事件抽取方法,其特征在于:所述强化学习智能体内部包含有强化学习框架,且所述强化学习框架用于对整个事件抽取过程进行建模。

3. 根据权利要求2所述的一种基于强化学习的生物学多事件抽取方法,其特征在于:所述强化学习框架由行动决策 A 、状态 S 、策略 π 和奖励 R 四个部分组成。

一种基于强化学习的生物医学多事件抽取方法

技术领域

[0001] 本发明涉及生物医学技术领域,具体为一种基于强化学习的生物医学多事件抽取方法。

背景技术

[0002] 近年来,生物医学领域的发展日新月异,相关科研文献增长的速度远远超过我们人类的阅读能力,研究者们通常需要花费大量的时间、精力去查找和阅读文献,以了解跟踪生物医学领域的最新研究进展,因此,如何自动从海量的生物医学文献中挖掘出对生物医学研究者有用的信息,并以结构化的方式进行展示成为了一个值得研究的问题。生物医学事件抽取是将非结构化的生物医学文本数据转化为生物医学事件的结构化表示,广泛应用于生物医学领域的知识获取和知识图谱构建等应用中,然而,由于生物医学事件中生物实体之间关系的复杂性,一种有效的事件提取方法在生物医学领域仍然亟待开发。

[0003] 生物医学事件抽取的概念源自于事件抽取,通用领域中通常会将事件抽取划分为触发词识别和元素检测两个任务,传统的方法都是按顺序执行这两个任务,综合目前国内外研究,事件抽取的主要研究方法分为基于规则、基于统计机器学习和基于深度学习方法三大类,基于规则的方法在特定领域内能取得较高的性能,但移植性较差,相对于基于规则的方法,统计机器学习与领域无关,无需太多领域专家的指导,系统移植性较好,但是在统计机器学习方法中,提取的特征的质量往往决定着系统的性能,而复杂的特征工程又是极其复杂和耗时的,因此自动提取特征的端到端的深度学习已成为事件抽取的主流研究方法,然而目前的这些方法仍然具有两个主要的缺点:1、对于生物医学领域的专有名词、缩写等词表示信息不够充分,无法很好的建模生物医学文本;2、一个句子中存在多个事件在生物医学领域是一个很普遍的现象,现有的方法在处理这种多事件问题时效率低下。

发明内容

[0004] 本发明的目的在于提供一种基于强化学习的生物医学多事件抽取方法,以解决上述背景技术中提出的问题。

[0005] 为实现上述目的,本发明提供如下技术方案一种基于强化学习的生物医学多事件抽取方法,包括以下步骤:

[0006] S1、将具有实体标注的句子通过词嵌入模块导入学习模块中,并对词进行预先训练,然后通过词性嵌入模块对句子中的每一个单词进行词性标注,并将其映射为向量,通过实体类型嵌入模块对标注后单词 X_i 的实体类型查找,并对其进行向量分配,然后将获取到的外部知识映射为一个向量导入外部知识嵌入模块中,通过学习模块将输入的句子 X 转换成一个新的词序列 $(W_1, \dots, W_2, \dots, W_n)$,并输入到事件抽取模块中;

[0007] S2、通过强化学习智能体依次扫描输入句子中的单词,并根据RL中的随机策略将事件类型分配给相应的单词,当强化学习agent识别出一个触发词后,将这个词的词向量和当前的环境信息联接作为当前的状态 S_t ,并将作为agent的输入,然后agent根据当前的状

态信息,输出决策行动 a_t ,表示对于当前触发词,模型预测出的事件类型;

[0008] S3、将表示行动的向量 a_t 和表示当前触发词的向量 w_t 与句子中的每个词向量联接得到一个新的句子表示 $X=(W_1, \dots, W_2, \dots, W_n)$,将其作为一个BiLSTM-CRF模块的输入,完成给定触发词的元素检测;

[0009] S4、模型预测的结果与训练集的真实标注将用于计算一个奖励值,用来评估当前模型采取的决策的好坏,同时,预测的结果还被映射为一个向量 L_t 并保存在环境信息中,最后对结构化事件进行输出;

[0010] 所述学习模块由词嵌入模块、词性嵌入模块、实体类型嵌入模块和外部知识嵌入模块组成;

[0011] 所述词嵌入模块用于通过查找预先训练好的词,并将其嵌入模型,同时对初始化词嵌入向量,并在整个框架的训练过程中进行微调;

[0012] 所述词性嵌入模块用于通过斯坦福标注工具对给定句子中的每个单词进行词性标注,并将其映射为向量,同时在模型训练过程中进行微调;

[0013] 所述实体类型嵌入模块用于根据实体类型查找嵌入表,并对每个单词分 X_t 配一个随机初始化的向量,并在整个框架的训练过程中进行微调;

[0014] 所述外部知识嵌入模块采用基因本体和蛋白质本体作为外部知识库,并针对生物医学事件进行任务抽取,同时将获取到的外部知识映射为一个向量,并将该向量知识嵌入到外部知识库中。

[0015] 作为优选,所述强化学习智能体内部包含有强化学习框架,且所述强化学习框架用于对整个事件抽取过程进行建模。

[0016] 作为优选,所述强化学习框架由行动决策 A 、状态 S 、策略 π 和奖励 R 四个部分组成。

[0017] 与现有技术相比,本发明的有益效果是:

[0018] 1.通过利用强化学习框架建模生物医学事件抽取的两个子任务,将两个任务分别作为主要任务和辅助任务,并通过强化学习的奖励机制建模了两个任务之间的交互关系,提高了事件抽取效率;

[0019] 2.通过利用强化学习框架,将已抽取的事件映射为向量保存在环境信息中,可以很好的帮助后续的事件抽取任务,从而提高模型在多事件抽取任务中的表现;

[0020] 3.通过在学习模块中引入外部知识库,有效地获取了这些特殊词语的外部知识库信息,从而加强了文本的表示,有利于后续的事件抽取任务。

附图说明

[0021] 为了更清楚地说明本发明具体实施方式或现有技术中的技术方案,下面将对具体实施方式或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图是本发明的一些实施方式,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0022] 图1为本发明生物医学多事件抽取方法的流程图。

具体实施方式

[0023] 下面对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例

仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0024] 本发明提供一种技术方案:一种基于强化学习的生物学多事件抽取方法,包括以下步骤:

[0025] S1、将具有实体标注的句子通过词嵌入模块导入学习模块中,并对词进行预先训练,然后通过词性嵌入模块对句子中的每一个单词进行词性标注,并将其映射为向量,通过实体类型嵌入模块对标注后单词 x_t 的实体类型查找,并对其进行向量分配,然后将获取到的外部知识映射为一个向量导入外部知识嵌入模块中,通过学习模块将输入的句子 X 转换成一个新的词序列 $(w_1, \dots, w_2, \dots, w_n)$,并输入到事件抽取模块中;

[0026] S2、通过强化学习智能体依次扫描输入句子中的单词,并根据RL中的随机策略将事件类型分配给相应的单词,当强化学习agent识别出一个触发词后,将这个单词的词向量和当前的环境信息联接作为当前的状态 S_t ,并将作为agent的输入,然后agent根据当前的状态信息,输出决策行动 a_t ,表示对于当前触发词,模型预测出的事件类型;

[0027] S3、将表示行动的向量 a_t 和表示当前触发词的向量 w_t 与句子中的每个词向量联接得到一个新的句子表示 $X = (w_1, \dots, w_2, \dots, w_n)$,将其作为一个BiLSTM-CRF模块的输入,完成给定触发词的元素检测;

[0028] S4、模型预测的结果与训练集的真实标注将用于计算一个奖励值,用来评估当前模型采取的决策的好坏,同时,预测的结果还被映射为一个向量 L_t 并保存在环境信息中,最后对结构化事件进行输出;

[0029] 所述学习模块由词嵌入模块、词性嵌入模块、实体类型嵌入模块和外部知识嵌入模块组成;

[0030] 所述词嵌入模块用于通过查找预先训练好的词,并将其嵌入模型,同时对初始化词嵌入向量,并在整个框架的训练过程中进行微调;

[0031] 所述词性嵌入模块用于通过斯坦福标注工具对给定句子中的每个单词进行词性标注,并将其映射为向量,同时在模型训练过程中进行微调;

[0032] 所述实体类型嵌入模块用于根据实体类型查找嵌入表,并对每个单词 x_t 配一个随机初始化的向量,并在整个框架的训练过程中进行微调;

[0033] 所述外部知识嵌入模块采用基因本体和蛋白质本体作为外部知识库,并针对生物学事件进行任务抽取,同时将获取到的外部知识映射为一个向量,并将该向量知识嵌入到外部知识库中。

[0034] 其中,所述强化学习智能体内部包含有强化学习框架,且所述强化学习框架用于对整个事件抽取过程进行建模。

[0035] 其中,所述强化学习框架由行动决策 A 、状态 S 、策略 π 和奖励 R 四个部分组成。

[0036] 具体的,当模型的输入是一个生物学文本中的句子 X ,它由 N 个单词组成,即 $X = (x_1, \dots, x_2, \dots, x_n)$,在表示学习模块中,每个单词 x_t 被映射为一个向量 w_t ,其由词嵌入、词性嵌入、实体类型嵌入和外部知识嵌入四部分信息组成:对于每个单词 x_t ,通过查找预先训练好的词嵌入模型初始化词嵌入向量,并在整个框架的训练过程中进行微调,在本发明中采用预训练模型BioBERT作为词嵌入模型;对于给定的句子,使用斯坦福标注工具给句中的

每个单词进行词性标注,并将其映射为向量,在模型训练过程中进行微调;由于实体标注能够提供某些线索来识别事件触发词和元素,因此实体类型(包括类型“None”)的嵌入也被用来丰富单词的语义信息,对于每个单词 x_t ,通过根据实体类型查找嵌入表来分配一个随机初始化向量,并在整个框架的训练过程中进行微调;针对生物学事件抽取任务,主要使用了基因本体和蛋白质本体作为外部知识库,可以为跨所有物种的基因或蛋白质的属性提供更详细的描述,为了更方便地获取这些外部知识库提供的额外信息,还使用QuickGOAPI和ProteinsAPI来从相应的知识库中查询生物学实体的属性,对于每个实体,主要选择实体类型和本体定义两种类型的属性,同时还将获取到的外部知识映射为一个向量作为这个词的外部知识嵌入,通过学习模块将输入的句子 x 转换成一个新的词表示序列 $(W_1, \dots, W_2, \dots, W_n)$,并输入到事件抽取模块中用于事件抽取,同时还可以学习信息更丰富的词性表示,

[0037] 由于强化学习智能体中包含有强化学习框架,可以通过强化学习框架来建模整个事件抽取过程,因为在事件抽取的过程中,触发词识别的结果会直接影响到后续元素检测任务的表现,因此在我们的模型中,触发词识别被作为主要任务,而元素检测作为其附属任务,在主任务中,强化学习智能体(Agent)依次扫描输入句子中的单词,并根据RL中的随机策略将事件类型(包括“None”)分配给相应的单词,如果当前单词被识别为事件触发词,则启动一个附属任务来完成已识别触发词的元素检测,并利用元素检测的结果来计算主任务中所采取行动的奖励,否则,如果预测的事件类型为“None”,Agent将跳转到下一个令牌以进行后续触发器识别,整个事件抽取的过程如下:

[0038] 当强化学习agent识别出一个触发词后,将这个单词的词向量和当前的环境信息(用于表示之前已经抽取到的事件信息)联接作为当前的状态 S_t ,并将作为agent的输入,然后agent根据当前的状态信息,输出决策行动 a_t ,表示对于当前触发词,模型预测出的事件类型,随后,将表示行动的向量 a_t 和表示当前触发词的向量 w_t 与句子中的每个词向量联接得到一个新的句子表示 $X = (W_1, \dots, W_2, \dots, W_n)$,将其作为一个BiLSTM-CRF模块的输入,完成给定触发词的元素检测,然后模型预测的结果与训练集的真实标注将用于计算一个奖励值,用来评估当前模型采取的决策的好坏,同时,预测的结果还被映射为一个向量 L_t 并保存在环境信息中,用于帮助后续的事件抽取;通过使用奖励机制和向量 L_t ,附属任务元素检测的结果能够反馈到触发词识别的过程中,在强化学习框架下,智能体agent被训练采取能获得更高的奖励的行动,这将在给定的输入句子中得出更好的生物学事件抽取结果,强化学习框架主要由:行动(决策) A 、状态 S 、策略 π 和奖励 R 四个部分组成,每个组成部分的细节描述如下:

[0039] 行动(决策) A :在 t 时刻,行动 a_t 是表示当前单词 w_t 所触发的生物学事件的类型,在本方法中,每个行动都是从 $A = \{\text{None}\}$ UT中选择的,其中 $\{\text{None}\}$ 表示当前词不是触发词, t 是事件类型的集合;

[0040] 状态 S :在 t 时刻,状态 $s_t \in S$ 是由当前单词的词表示 w_t ;最近一次采取的行动 a_t ($a_t \neq \text{None}$);表示上一次事件抽取结果的向量 L_t ;上一时刻的状态 s_{t-1} ;且 s_t 由 $s_t = f(w_s [w_t; a_t; L_t; s_{t-1}])$ 公式得到,其中 $f(\cdot)$ 是一个由MLP实现的非线性函数, w_s 是在模型训练过程中可以学习的权重矩阵;

[0041] 策略 P :用于触发词识别的随机策略 $\pi: S \rightarrow A$,定义了一个行动的概率分布: $a_t \sim \pi(a_t | s_t) = \text{softmax}(W_\pi s_t)$,其中 $\text{softmax}(\cdot)$ 的参数在模型的训练过程中被优化;

[0042] 奖励R:在采取了行动 a_t 后,环境首先给出了一个基础奖励 r_t^b ,通过判断预测的事件类型 a_t 是否在句子 x 中,来估计最终奖励,基础奖励的定义为:

$$[0043] \quad r_t^b = \begin{cases} 1, & \text{if } a_t \text{ in } X \\ 0, & \text{if } a_t = \text{None} \\ -1, & \text{if } a_t \text{ not in } X \end{cases}$$

$$[0044] \quad r_t^f = r_t^b \cdot \sum_{i=1}^N g(\hat{y}_i == y_i)$$

[0045] 如果 $a_t = \text{None}$,强化学习agent将跳到下一个标记,并在下一时刻采取一个新的主任务决策,同时保持环境信息中的 a_t 和 L_t 不改变,否则,agent将在确定的事件触发词下,为相应的元素检测启动一个附属任务,随后,模型预测的结果 \hat{Y} 和训练集提供的标准标注 Y 用于计算一个最终奖励,用于评估事件抽取结果的好坏,计算公式如下:

$$[0046] \quad r_t^f = r_t^b \cdot \sum_{i=1}^N g(\hat{y}_i == y_i)$$

[0047] 其中 $g(\cdot)$ 是指示函数,如果里面的断言为真时,输出1,否则输出0,为了得到最策略,采用一种策略梯度算法对模型进行训练,可以使预期的累积优

[0048] 奖励最大化,在每个时间步长 t ,强化学习agent按照策略 π 对预测的行动轨迹进行采样,累积奖励可按如下方式计算:

$$[0049] \quad J(\theta_{\pi,t}) = \mathbb{E}_{s,a \sim \pi(a|s)} \left[\sum_{i=t}^N \gamma^{i-t} r_i^f \right]$$

[0050] 其中 $\theta_{\pi,t}$ 表示在 t 时刻的策略 π 的参数, γ 表示折损因子,然后使用REINFORCE算法,用策略分布中的行动采样轨迹近似计算公式中的期望值,并在训练过程中优化策略更新参数的梯度计算公式如下:

$$[0051] \quad \nabla J(\theta_{\pi,t}) \approx \sum_{i=t}^N [\gamma^{i-t} r_i^f - b] \nabla \log \pi(a|s)$$

[0052] 其中 b 表示阈值,它被设定为以前所有奖励的平均值。

[0053] 尽管已经示出和描述了本发明的实施例,对于本领域的普通技术人员而言,可以理解在不脱离本发明的原理和精神的情况下可以对这些实施例进行多种变化、修改、替换和变型,本发明的范围由所附权利要求及其等同物限定。

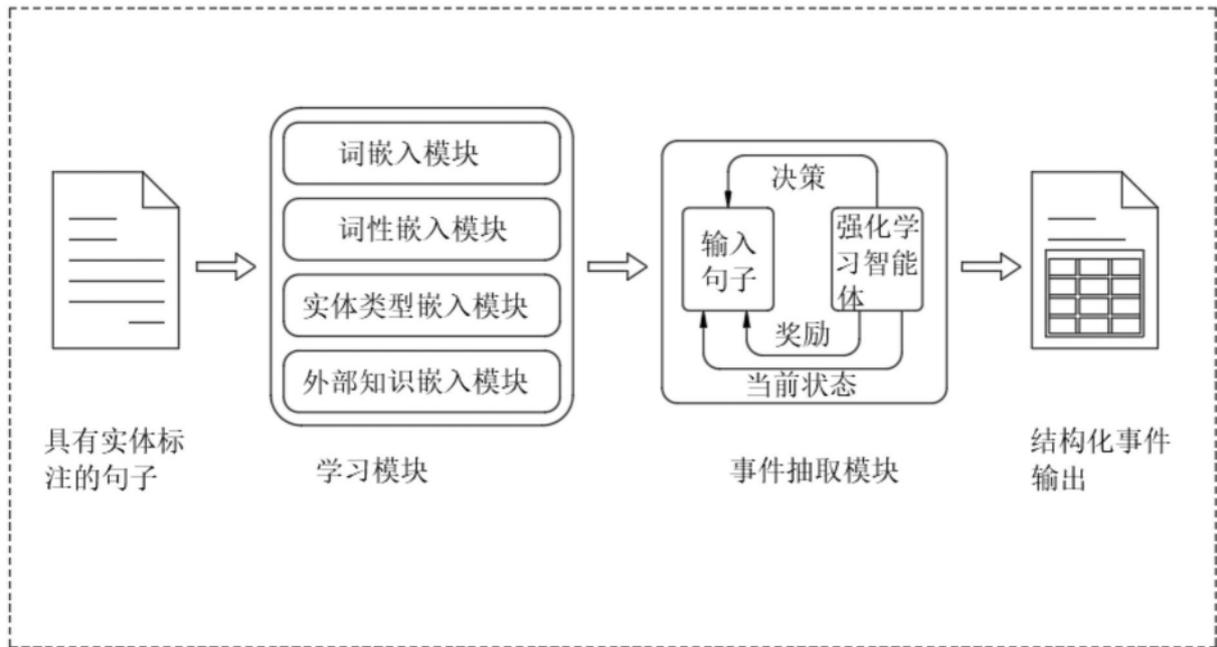


图1