

DOI:10.16781/j.0258-879x.2020.10.1129

· 论 著 ·

一种基于注意力机制的电子病历中药物词向量转化方法

王青华¹, 邵劲松², 张远鹏¹, 姜磊^{3,4}, 白鹤鸣⁵, 黄勋⁶, 徐沪济³, 王理^{1,4,5*}

1. 南通大学医学院医学信息系, 南通 226001
2. 南通大学公共卫生学院公共卫生与预防医学系, 南通 226001
3. 海军军医大学(第二军医大学)长征医院风湿免疫科, 上海 200003
4. 清华大学精准医学研究院, 北京 100084
5. 南通大学智能信息技术中心, 南通 226001
6. 南通大学信息科学技术学院通讯工程系, 南通 226001

[摘要] **目的** 提出一种基于注意力机制的药物词向量生成模型 Drug2vec, 对药物信息做向量化表示, 并与 Word2vec 和 Med2vec 模型比较向量转化效果。**方法** 使用注意力机制捕获医疗实体对中心词的作用, 提出 Drug2vec 模型, 将非结构化电子病历中的医疗实体转化为向量。使用包含 14 219 例系统性红斑狼疮(SLE)患者和 963 个药物实体的数据集测试 Drug2vec 模型生成词向量的效果, 并且与广泛应用的语言概念空间向量转化模型 Word2vec 和 Med2vec 进行对比。**结果** 在 SLE 患者数据集中, Drug2vec 模型产生的药物词向量准确度优于 Word2vec 和 Med2vec 模型。药物词向量相似度排序结果显示 Drug2vec 模型产生的向量结果符合临床医师的用药顺序。**结论** Drug2vec 模型可以更精确地利用周围实体修正中心药物实体, 从而产生更准确的药物向量。

[关键词] 语言概念词向量; 注意力机制; 系统性红斑狼疮; 电子健康病历

[中图分类号] R 197.324 **[文献标志码]** A **[文章编号]** 0258-879X(2020)10-1129-07

A drug word vector conversion method in electronic medical record based on attention mechanism

WANG Qing-hua¹, SHAO Jin-song², ZHANG Yuan-peng¹, JIANG Lei^{3,4}, BAI He-ming⁵, HUANG Xun⁶, XU Hu-ji³, WANG Li^{1,4,5*}

1. Department of Medical Informatics, School of Medicine, Nantong University, Nantong 226001, Jiangsu, China
2. Department of Public Health and Preventive Medicine, School of Public Health, Nantong University, Nantong 226001, Jiangsu, China
3. Department of Rheumatology and Immunology, Changzheng Hospital, Naval Medical University (Second Military Medical University), Shanghai 200003, China
4. Institute of Precision Medicine, Tsinghua University, Beijing 100084, China
5. Intelligence Information Technology Center, Nantong University, Nantong 226001, Jiangsu, China
6. Department of Communication Engineering, School of Information Science and Technology, Nantong University, Nantong 226001, Jiangsu, China

[Abstract] **Objective** To propose a drug word vector conversion model based on attention mechanism named Drug2vec for generating vectorized representation of drug information, and to compare the vector conversion effect with Word2vec and Med2vec. **Methods** Using the attention mechanism to capture the roles of medical entities on the central word, we proposed a Drug2vec model to convert medical entities in unstructured electronic medical records into vectors. Using the systemic lupus erythematosus (SLE) dataset of 14 219 patients and 963 drug entities, we tested the effect of the drug vectors generated by Drug2vec and compared it with the widely used language concept space vector conversion models Word2vec and Med2vec. **Results** In the SLE dataset, the accuracy of drug vectors generated by Drug2vec was higher than those of Word2vec and Med2vec models. The rank results of the similarity of drugs showed that the drug vectors generated by Drug2vec were

[收稿日期] 2019-10-30 **[接受日期]** 2020-02-21

[基金项目] 国家自然科学基金(81873915), 国家重点研发计划(2018YFC0116902), 南通市科技项目(CP12016003). Supported by National Natural Science Foundation of China (81873915), National Key Research and Development Plan (2018YFC0116902), and Science and Technology Program of Nantong (CP12016003).

[作者简介] 王青华, 硕士生. E-mail: 17160035@yjs.ntu.edu.cn

*通信作者(Corresponding author). Tel: 0513-85051891, E-mail: wangli@ntu.edu.cn

consistent with the clinician's medication order. **Conclusion** Drug2vec model can more accurately modify central drug entities using contextual entities, producing more precise drug vectors.

[Key words] language concept word vector; attention mechanism; systemic lupus erythematosus; electronic health records

[Acad J Sec Mil Med Univ, 2020, 41(10): 1129-1135]

随着电子病历的普及,电子病历系统累积的数据不断增加。电子病历中存储着大量医师记录的描述性信息,而非结构化文本,我们需要将这些复杂的非结构化文本信息转化为数值信息,才能对其进行有效二次利用,挖掘出潜在的医疗规律^[1]。近年来,机器学习和深度学习在医疗领域中的应用越来越多,如疾病预测和患者分类等^[2-3]。

有研究者针对非结构化文本空间向量转化的问题提出了一些解决方案,如 Google 在 2013 年提出了 Word2vec 模型,该模型采用连续词袋(continuous bag-of-words, CBOW)和 Skip-gram 2 种方式,在无监督的情况下可以将文本数据转化为空间词向量^[4]。Word2vec 模型只利用上下文单词向量之和生成中心词的向量,或者利用中心词生成上下文单词向量,是一个针对通用常识文本的词向量转化模型,并没有针对医疗领域的特殊场景进行设计。因此,Word2vec 模型不能准确地表达医疗场景下各医疗实体之间的关系。Glove 是 Pennington 等^[5]在 2014 年提出的一个基于全局内一定窗口词对共现矩阵的词向量转化模型,该模型结合了中心词的上下文信息和全局词共现矩阵信息生成词向量,理论上可以产生更准确的词向量,但其因使用了全局信息而导致内存耗费巨大,而且它的实际表现并不尽如人意,因此并不及 Word2vec 模型普及。Choi 等^[6-7]在 2016 年提出了 Med2vec 模型,这是一个专门为医疗实体设计的模型。但 Med2vec 也仅使用了类似 Skip-gram 的结构,同样没有很好地捕获各医疗实体之间的时序关系,所以转化出的向量准确性还有待提高。Cao 等^[8]从中文文字笔画出发,提出了一种基于中文笔画信息的向量转化模型 cw2vec,它借助笔画信息的 n-grams 抓住语义信息和汉字形态信息从而产生词向量。cw2vec 的核心也是采用 Skip-gram 方式利用中心词预测上下文,不同的是它使用笔画的 n-grams 表示上下文,更适合中文信息的抽取,但它只能处理汉

字,不能表示句子中的其他符号(如英文和标点符号等)。Zhao 等^[9]将 PubMed 和 DrugBank 中的药物语料作为训练样本输入 Skip-gram 模型,并使用向量和频率的关系作为评估指标调整 Skip-gram 模型的参数,最终获得了较准确的药物向量。使用药物领域的语料和向量、频率之间的关系虽然可一定程度提高向量的准确性,但并没有从根本上改变生成向量的方式。

目前常用的深度学习模型中,循环神经网络(recurrent neural network, RNN)和卷积神经网络(convolutional neural network, CNN)有不错的特征提取能力,一直被研究人员广泛应用于文本和图像的特征提取^[10-11]。例如,RNN 中的长短记忆神经网络(long short-term memory, LSTM)常被用在类似生成词向量的特征提取任务中^[12-13]。LSTM 的缺点是其多个隐藏层的神经元值需要一层一层传递,不能进行大规模的并行运算,降低了计算能力。CNN 通常被用来进行图像的分类和识别等任务^[14],由于它可以通过卷积和池化策略达到提取特征的目的,近些年来也被用于自然语言处理任务。Zheng 等^[15]在 2017 年提出了一种基于 CNN 的词向量转化模型,该模型使用 CNN 获取上下文特征信息,而非上下文相加的信息。CNN 虽然可以做并行计算,但是不能捕获自然语言处理任务中长距离的特征关系,如在机器翻译中 CNN 只能捕获与要翻译的词距离相近的词之间的关系,不能得到距离更远的词对它的影响,从而导致翻译不准确^[16]。

注意力机制是一种抽取特征的重要方法,在神经机器翻译的编码解码过程中,基于注意力机制的 Transformer 结构可以准确地捕获长距离的依赖关系,产生更准确的翻译内容^[17]。在医疗领域,患者的每个诊疗事件都包含了多个医疗实体,如诊断、药物和检查等。在生成一个中心词医疗实体向量时,周围的医疗实体对中心词的作用不同,如某个诊断对词向量生成的影响较明显,而另外一个诊断

和词向量生成的相关性较弱。本研究使用注意力机制捕获这种不同的作用, 提出了 Drug2vec 模型, 将非结构化电子病历中的医疗实体转化为向量。

系统性红斑狼疮(systemic lupus erythematosus, SLE) 是一种累及全身多个器官和系统的自身免疫性疾病, 常伴随多种并发症^[8]。本研究从 13 个 SLE 数据中心搜集了 57 367 份电子病历数据, 包含了 14 439 例患者 26 583 次诊疗。本研究在这个数据集上测试了 Drug2vec 模型生成词向量的效果, 并且与广泛应用的语言概念空间向量转化模型 Word2vec 和 Med2vec 进行对比。通过临床医师的解读, Drug2vec 模型转化的向量相似度结果更具临床价值。

1 资料和方法

1.1 实验数据 SLE 病情的发展一般呈现由轻到重的趋势, 逐步进展至累及全身多个器官和系统, 大多数患者有多次就诊经历。本研究中使用的数据来自中国 13 家三甲医院 14 439 例 SLE 患者的住

院记录, 时间跨度为 2001 年 10 月 28 日至 2017 年 3 月 31 日。通过使用自然语言处理的方法从这些文本中提取患者的人口统计学信息(年龄、性别和婚姻状况)和药物等信息, 构建临床活动表型数据。排除信息不合格病例 220 例, 最终使用的数据集中包含了 14 219 例患者和 10 469 个医疗实体, 其中药物实体 963 个、实验室指标实体 35 个、症状实体 8 365 个、诊断实体 1 106 个。

1.2 实验方法 Word2vec 模型的 Skip-gram 结构中, 主要通过中心词产生上下文的词向量, 中心词对上下文的每个词更新相同的值; CBOW 结构中, 通过对上下文向量求和产生中心词的词向量。这 2 种方式在修正词向量时都将上下文词视为同等重要, 通过训练大规模数据获取准确的向量值。Word2vec 模型的基本结构如图 1 所示, 其中 $e(t)$ 表示中心词, $e(t-2)$ 、 $e(t-1)$ 、 $e(t+1)$ 、 $e(t+2)$ 均表示其上下文词(Skip-gram 结构中相同颜色表示相同的值, CBOW 结构中不同颜色表示不同的值)。

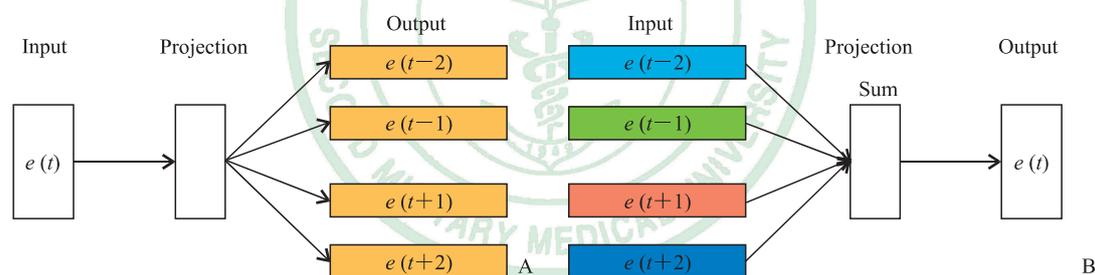


图 1 Word2vec 模型的基本结构

Fig 1 Basic structure of Word2vec model

A: Skip-gram; B: CBOW. $e(t)$: Target word; $e(t-2)$, $e(t-1)$, $e(t+1)$, $e(t+2)$: Contexts of the target word. The same color indicates the same value in the Skip-gram model, and the different colors indicate different values in the CBOW model. CBOW: Continuous bag-of-words

Med2vec 模型分为 2 层: 第一层用于捕获医疗实体之间的关系, 第二层用于捕获诊疗事件序列之间的关系。该模型本质上采用的是 Skip-gram 结构, 仍然认为周围词是同等重要的。Med2vec 模型的基本结构如图 2 所示, 输入 x_t 表示患者 1 次就诊, $|C|$ 表示医疗实体总数, 在 1 次就诊 x_t 中, 每个医疗实体用 0 和 1 表示; 输出的 x_{t-2} 、 x_{t-1} 、 x_{t+1} 、 x_{t+2} 表示就诊 x_t 的上下文信息; 黑色粗箭头表示神经网络中信息的传递方向。

在自然语言中, 每个词对中心词的作用是不

同的, 因此 Drug2vec 模型中, 上下文的每个词对中心词的更新值也不同, 其基本结构如图 3 所示, 其中 $e(t)$ 表示中心词向量, $e(t-k)$ 、 \dots 、 $e(t-2)$ 、 $e(t-1)$ 、 $e(t+1)$ 、 $e(t+2)$ 、 \dots 、 $e(t+n)$ 分别表示上下文相关的实体向量。Event 代表患者 1 次诊疗事件, 中心词前面有 k 个向量, 后面有 n 个向量, 连接线的颜色深浅和粗细表示每个词对中心词的不同作用, 颜色越深、连接线越粗表示它对中心词的影响越显著。

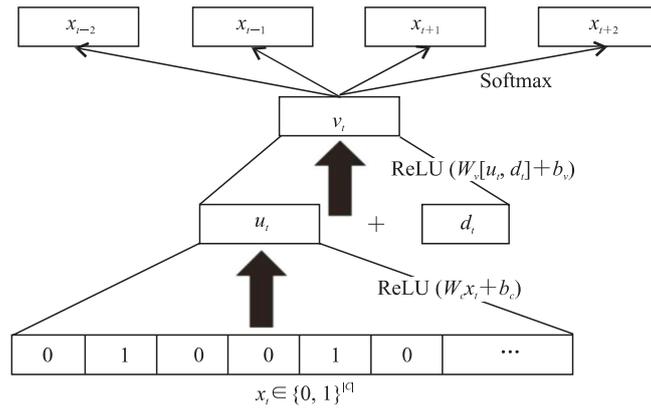


图2 Med2vec 模型的基本结构

Fig 2 Basic structure of Med2vec model

The thick black arrows indicate the direction of information transmission in neural network. x_t : A binary vector comprised of several medical codes; $x_{t-2}, x_{t-1}, x_{t+1}, x_{t+2}$: Contexts of x_t ; ReLU: Rectified linear unit; W_c, W_v : Weight of network; b_c, b_v : Biases of network; d_t : A vector of demographic information; $|C|$: The number of all medical codes

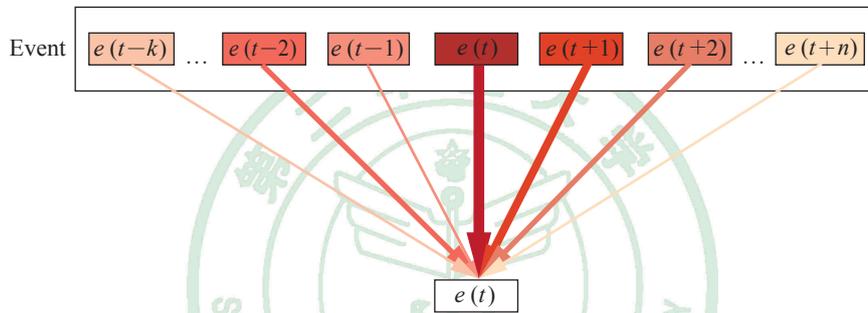


图3 Drug2vec 模型的基本结构

Fig 3 Basic structure of Drug2vec model

$e(t)$: Target word; $e(t-k), \dots, e(t-2), e(t-1), e(t+1), e(t+2), \dots, e(t+n)$: Contexts of the target word; Event: A diagnosis and treatment event of patients. The darker the color and the thicker the connection line, the more significant the influence of the context-related words on the target word

Drug2vec 模型使用注意力机制捕获患者各医疗实体之间的关系。其中注意力机制使用的是多头点乘注意力, 如下公式:

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

其中 Q, K, V 分别表示查询向量、键向量和值向量, d_k 表示 Q, K 或 V 的维度, 除以 d_k 的平方根是防止 QK^T 过大 Softmax 函数进入饱和区而使梯度过小, T 为向量矩阵计算过程中的转置符号。使用多个注意力头计算注意力时可以并行运算, 加快计算速度。每一层都使用了残差连接以防止梯度消散。

为了捕获患者诊疗事件中各个医疗实体的相互关联信息, 需要使用一个合适的损失函数。在 Word2vec 模型的 Skip-gram 结构中使用中心词生成上下文的词向量, 输入为中心词的词向量, 预测输出为上下文的词向量。在 Drug2vec 模型的输入序列中, 患者 1 次诊疗中的医疗实体是没有前后顺序的,

所以没有的固定的输入和输出之分。经过迭代训练得到向量矩阵 W , W 中的第 i 行即医疗实体集合中的第 i 个向量。因此, 通过不断优化向量矩阵 W 得到最终的医疗实体向量。Drug2vec 模型的损失函数使用 log 似然函数, 其公式如下:

$$\frac{1}{T} \sum_{i=1}^T \sum_{i: e_i \in E} \sum_{j: e_j \in E, j \neq i} \log P(e_j | e_i),$$

$$P(e_j | e_i) = \frac{\text{Exp}(W[i, :]^T W[j, :])}{\sum_{k=1}^{\text{all}} \text{Exp}(W[k, :]^T W[i, :])}$$

其中 T 表示某例患者的诊疗事件总数; t 为计数变量, 从 1 到 T ; e_i, e_j 分别表示 1 次诊疗事件中的第 i 个和第 j 个医疗实体; E_t 表示某例患者的第 t 个诊疗事件; P 表示概率; W 表示向量矩阵; all 表示所有唯一性的医疗实体的总数; k 为计数变量, 从 1 到 all 。

2 结果

所有模型的训练是在 1 台配置了 2 个英伟达 TESLA P100 显卡的 CentOS 服务器上进行的, 显卡的显存都为 16 G。Drug2vec 模型使用 TensorFlow1.8.0 深度学习框架实现和训练^[19]。Drug2vec 模型注意力机制中的注意力头个数是 8 个, 词向量的维度是 512。为了保持一致, Word2vec 和 Med2vec 模型的词向量维度也均是

512。Word2vec 模型使用的是 gensim3.6.0 工具包, 其中窗口大小为 5, 最小词频为 5。

本实验指定了诊断这个分类中 SLE 作为中心词, 通过计算药物向量和中心词的余弦相似值确定和 SLE 相关度较高的药物实体; 同时筛选数据集中出现频率较高的药物实体, 并记录其频数。与 SLE 相关度最高和数据集中出现频率最高的前 10 个药物实体见表 1。

表 1 与系统性红斑狼疮相关度和出现频率最高的前 10 个药物实体

Tab 1 Top 10 medication entities with highest correlation or frequency to systemic lupus erythematosus

Rank	Drug2vec		Word2vec	
	Cosine similarity	Medication	Cosine similarity	Medication
1	0.92	Hydroxychloroquine sulfate tablets	0.68	Fluvoxamine maleate tablets
2	0.90	Prednisone acetate injection	0.67	Dexamethasone
3	0.89	Methylprednisolone sodium succinate injection	0.66	Thiamazole tablets
4	0.88	Prednisone acetate tablets	0.66	Salvianolate injection
5	0.87	Calcium carbonate D3 tablets	0.65	Colquhounia root tablets
6	0.86	Cyclophosphamide injection	0.65	<i>Tripterygium wilfordii</i> Hook. f. tablets
7	0.81	Omeprazole enteric capsules	0.65	Prednisone
8	0.78	Calcitriol capsules	0.65	Omeprazole injection
9	0.69	Alfacalcidol soft capsules	0.64	Breviscapinun injection
10	0.69	Aspirin enteric-coated tablets	0.63	Cefoperazone sodium and tazobactam

Rank	Med2vec		Frequency	Medication
	Cosine similarity	Medication		
1	0.29	Cyclophosphamide injection	16 436	Hydroxychloroquine sulfate tablets
2	0.24	Oxacillin sodium injection	16 558	Methylprednisolone sodium succinate injection
3	0.23	Rebamipide tablets	12 629	Cyclophosphamide injection
4	0.22	Oxytocin injection	12 095	Prednisone acetate injection
5	0.21	Heart saving pill	8 191	Calcium carbonate D3 tablets
6	0.20	Compound reserpine tablets	7 386	Prednisone acetate tablets
7	0.20	Prednisone acetate tablets	6 722	Calcitriol capsules
8	0.17	Ferrous sulfate tablets	5 658	Omeprazole enteric capsules
9	0.17	Recombinant human insulin injection	3 564	Alfacalcidol soft capsules
10	0.16	Caltrate D	3 459	Aspirin enteric-coated tablets

3 讨论

SLE 患者一般首选糖皮质激素治疗, 如注射用醋酸泼尼松, 但大部分患者病程中需要使用免疫调节药物 (如硫酸羟氯喹等), 若合并器官损伤和并发症还需使用更强效的免疫抑制剂 (如环磷酰胺和

吗替麦考酚酯片等)^[20-21]。由于糖皮质激素可能导致胃黏膜破坏和骨质疏松等, 因此 SLE 患者往往还使用奥美拉唑肠溶胶囊和骨化三醇等辅助药物。

根据本实验数据, Drug2vec 模型中前 10 个与 SLE 最相关的药物符合医师的用药顺序 (从常用的糖皮质激素、免疫抑制剂到一些辅助药物), 其中

相关度最高的药物硫酸羟氯喹片是最常用的免疫调节药物。其次为注射用醋酸泼尼松、注射用甲泼尼龙琥珀酸钠、醋酸泼尼松片,三者均为常用的治疗SLE的糖皮质激素类药物。相关度排在第5位的药物是碳酸钙D3咀嚼片,作为一种补钙剂,其与相关度排在其后的骨化三醇胶丸和阿法骨化醇软胶囊常被用于治疗骨质疏松和骨软化症等疾病。奥美拉唑肠溶胶囊被用于治疗胃溃疡等疾病,也被用于治疗SLE相关器官损伤。

Word2vec模型训练出的相关度前10的药物实体虽然都是与SLE相关的药物,但是它们和SLE相关性不高,不仅缺乏一些通用的SLE治疗药物,排序也较为混乱。如缺少注射用醋酸泼尼松等糖皮质激素类药物及使用频率最高的药物硫酸羟氯喹片。马来酸氟伏沙明片(排第1位)属于精神病治疗药物,甲巯咪唑片(排第3位)用于治疗甲状腺病,丹参多酚酸盐注射液(排第4位)用于治疗心血管病,火把花根片(排第5位)和注射用灯盏花素(排第9位)都不是通用的SLE治疗药物,它们和SLE的相关性都比较弱,但在Word2vec模型中却排在了最相关的前10个药物中。

Med2vec模型训练出的相关度前10的药物实体中出现了一些和SLE相关性较强的药物,如注射用环磷酰胺和醋酸泼尼松片。但该模型排序混乱,醋酸泼尼松片排在第7位,而排在它之前的药物如注射用苯唑西林钠注射液(排第2位)属于抗感染药物,瑞巴哌特片(排第3位)用于治疗胃溃疡和胃炎,缩宫素注射液(排第4位)一般用于孕妇引产和催产等,救心丸(排第5位)常用于治疗心绞痛和心肌梗死等疾病,这些药物和SLE相关性都很弱。

从本实验计算的相关度可见,Word2vec模型中排在第1位的药物马来酸氟伏沙明片与中心词SLE的相关度为0.68,Med2vec模型中排在第1位的药物注射用环磷酰胺与SLE的相关度仅为0.29,而在Drug2vec模型中硫酸羟氯喹片与SLE的相关度达到了0.92。Word2vec模型计算出的前10个药物向量相关度不高的原因在于,使用Skip-gram结构时,中心词对上下文词更新同样的值;使用CBOW结构时,将上下文词的向量值进行简单求和获得中心词的词向量,因此生成的词向量边界不易

区分,造成一些异质性的医疗概念模糊。Med2vec模型同样使用了Skip-gram结构训练诊疗事件序列,Word2vec模型中的窗口大小在最大窗口范围内随机变化,而Med2vec模型中的窗口大小固定,因此其词向量区分度也不高。Drug2vec模型基于注意力机制计算的方式解决了中心词对上下文词更新同样值的问题,使计算出的概念表示向量更加准确,所以最终的医疗实体向量有更高的区分度。

本实验中Drug2vec模型是建立在电子病历数据集的基础上,需要使用自然语言处理等技术从电子病历中抽取患者的表型信息,包括诊断、症状和药物等实体及患者诊疗的时序信息。患者的1次诊疗经历中多个医疗实体没有前后顺序信息,而1例患者的多次诊疗之间有时序信息,因此需将患者前后几次诊疗信息按照时间排列。最后将整理的数据集输入Drug2vec模型进行训练,得到实体向量之后通过计算向量之间的余弦相似值衡量实体与中心词之间的相关度。

本实验提出了一个新的药物实体向量生成模型Drug2vec,它是一个基于注意力机制的深度学习模型。相较于Word2vec和Med2vec,Drug2vec模型可以更精确地利用周边医疗实体修正中心词医疗实体向量。在一个真实世界的临床疾病数据集——SLE数据集中,我们对比测试了Drug2vec模型的性能,结果表明Drug2vec模型可以生成更精准的向量。未来,我们准备测试更多种类的医疗实体如诊断、实验室指标、症状和体征等,并将生成的向量应用到更多的场景,如SLE并发症预测和住院时间预测等。

[参考文献]

- [1] RAJKOMAR A, OREN E, CHEN K, DAI A M, HAJAJ N, HARDT M, et al. Scalable and accurate deep learning with electronic health records[J/OL]. NPJ Digit Med, 2018, 1: 18. doi: 10.1038/s41746-018-0029-1.
- [2] BAHL M, BARZILAY R, YEDIDIA A B, LOCASCIO N J, YU L, LEHMAN C D. High-risk breast lesions: a machine learning model to predict pathologic upgrade and reduce unnecessary surgical excision[J]. Radiology, 2018, 286: 810-818.
- [3] 李渊, 骆志刚, 管乃洋, 尹晓尧, 王兵, 伯晓晨, 等. 生物医学数据分析中的深度学习应用[J]. 生物化学与生物物理进展, 2016, 43: 472-483.

- [4] MIKOLOV T, SUTSKEVER I, CHEN K, CORRADO G S, DEAN J. Distributed representations of words and phrases and their compositionality[C]//Advances in Neural Information Processing Systems 26 (NIPS 2013). Lake Tahoe: NIPS, 2013: 3111-3119.
- [5] PENNINGTON J, SOCHER R, MANNING C D. Glove: global vectors for word representation[C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing empirical. Doha: EMNLP, 2014: 1532-1543.
- [6] CHOI E, BAHADORI M T, SEARLES E, COFFEY C, THOMPSON M, BOST J, et al. Multi-layer representation learning for medical concepts[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco: KDD, 2016: 1495-1504.
- [7] CHOI E, SCHUETZ A, STEWART W F, SUN J. Using recurrent neural network models for early detection of heart failure onset[J]. J Am Med Inform Assoc, 2017, 24: 361-370.
- [8] CAO S, LU W, ZHOU J, LI X. cw2vec: learning Chinese word embeddings with stroke n-gram information[C]//The Thirty-Second AAAI Conference on Artificial Intelligence. New Orleans: AAAI, 2018. 5053-5061.
- [9] ZHAO M, MASINO A J, YANG C C. A framework for developing and evaluating word embeddings of drug-named entity[C]//Proceedings of the BioNLP 2018 workshop. Melbourne: BioNLP, 2018: 156-160.
- [10] WILLIAMS R J, ZIPSER D. A learning algorithm for continually running fully recurrent neural networks[J]. Neural Comput, 1989, 1: 270-280.
- [11] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Neural Information Processing Systems, 2012, 141: 1097-1105.
- [12] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Comput, 1997, 9: 1735-1780.
- [13] BAYTAS I M, XIAO C, ZHANG X S, WANG F, JAIN A K, ZHOU J. Patient subtyping via time-aware LSTM networks[C]//Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Nova Scotia: KDD, 2017: 65-74.
- [14] WEI Y, XIA W, LIN M, HUANG J, NI B, DONG J, et al. HCP: a flexible CNN framework for multi-label image classification[J]. IEEE Trans Pattern Anal Mach Intell, 2016, 38: 1901-1907.
- [15] ZHENG X, FENG J, CHEN Y, PENG H, ZHANG W. Learning context-specific word/character embeddings[C]//Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence. San Francisco: AAAI. 2017: 3393-3399.
- [16] ZHANG J, ZONG C. Deep neural networks in machine translation: an overview[J]. IEEE Intell Syst, 2015, 30: 16-25.
- [17] VASWANI A, SHAZEER N, PARMAR N, USZKOREIT J, JONES L, GOMEZ A N, et al. Attention is all you need[C]//The 31st Annual Conference on Neural Information Processing Systems. Long Beach: NIPS, 2017: 5998-6008.
- [18] 蒋明. 系统性红斑狼疮诊断治疗的最新进展[J]. 中国实用内科杂志, 2000, 20: 55-57.
- [19] ABADI M, BARHAM P, CHEN J, CHEN Z, DAVIS A, DEAN J, et al. TensorFlow: a system for large-scale machine learning[C]//Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation. Savannah: OSDI. 2016: 265-283.
- [20] 余碧娥, 冯树芳, 方丽, 陆肇曾. 硫酸羟氯喹治疗盘状红斑狼疮的分析[J]. 临床皮肤科杂志, 2001, 30: 310-311.
- [21] PUNTHAKEE Z, LEGAULT L, POLYCHRONAKOS C. Prednisolone in the treatment of adrenal insufficiency: a re-evaluation of relative potency[J]. J Pediatr, 2003, 143: 402-405.

[本文编辑] 杨亚红