

基于强化学习的医疗问题诉求分类

吴昊, 黄德根, 林晓惠

(大连理工大学 计算机科学与技术学院, 辽宁, 大连 116024)

摘要: 医学问题诉求分类属于文本分类, 是自然语言处理中的基础任务。该文提出一种基于强化学习的方法对医学问题进行诉求分类。该方法首先通过强化学习自动识别出医学问题中的关键词, 并且对医学问题中的关键词和非关键词赋予不同的值构成一个向量; 其次, 利用该向量作为 attention 机制的权重向量, 对 Bi-LSTM 模型生成的隐含层状态序列加权求和得到问题表示; 最后通过 softmax 分类器对问题表示进行分类。实验结果表明, 该方法比基于 Bi-LSTM 模型的分类结果准确率提高 1.49%。

关键词: 强化学习; Bi-LSTM; attention 机制

Medical Question Appeal Classification Based on Reinforcement Learning

Wu Hao, Huang Degen, Lin Xiaohui

(School of Computer Science and Technology, Dalian University of Technology, Dalian, Liaoning 116024, China)

Abstract : Medical problem appeal classification belongs to text classification and it is the basic task in natural language processing. This paper presents a method based on reinforcement learning to medical question appeal classification. Firstly, the key words in medical problems are automatically identified by reinforcement learning, and the key words and non-key words in medical problems are assigned different values to constitute a vector. Secondly, the vector as the weight vector of the attention mechanism to the weighted sum of the hidden layer state sequences generated by the Bi-LSTM model constitute question representation. At last, the question representation is classified by softmax classifier. Experimental results show that the accuracy of this method is 1.49% higher than that based on bi-lstm model.

Key words: reinforcement learning; Bi-LSTM; attention mechanism

0 引言

医学问题诉求分类属于文本分类的一种, 主要针对互联网中患者提出的问题, 判断其提出问题的意图, 根据对患者问题样本的分析及专业医药顾问的建议, 我们将诉求主要分为 6 大类:

① “怎么治”: 患者明确提出自己所患何种疾病, 询问有效治疗手段 (干预措施)。

② “什么病”: 患者询问所患的是什么疾病或区分所患的具体疾病名称。

③ “为什么”: 患者询问某种疾病或现象的原理。

④ “会怎么样”: 患者询问某种疾病或药物可能导致、引起的疾病负担或不良反应。

⑤ “注意事项”: 患者主体大多为特定人群, 询问某种疾病或药物需要注意的行为、饮食或其他事项。

⑥ “医学无关问题”: 在互联网中爬取的问题, 有些不属于医学问题, 但其参杂在医学问题中, 我们把其单独归为“医学无关问题”。

医学问题诉求分类任务可以了解普通大众对疾病的诊治需求, 推进智慧医疗的进程; 同时, 也可为药企了解市场需求、研发新药提供新的窗口, 更好地为大众服务。因此, 准确地判断互联网医学问题的诉求具有重要意义。

相比于其一般的文本分类, 互联网医学问题诉求分类具有以下特点:

① 互联网提问平台对问题的格式、表述方式没有统一的规定, 因此互联网医学问题存在表述灵活的特点, 这也增加了问题诉求分类的难度;

收稿日期: 定稿日期:

基金项目: 国家自然科学基金资助项目 (61672127)

② 医学问题中存在大量的医学专有词汇，如药品名称，疾病名称等，这些词汇不能被现有的分词工具很好的分词，进而影响问题的诉求分类结果。

③ 医学问题中一般会有关键性的词语标志问题的诉求，正确使用关键词所提供的信息可以提高诉求分类结果。

文本分类作为自然语言处理中常见的任务之一，其核心问题是文本表示。传统的文本表示利用词袋原理(Bag-of-Words)，将词无序地表示到一个高维的向量空间模型(Vector Space Model)中，丢弃了词序、语法等文本结构信息。近年来，随着神经网络在各个领域中发挥出自己的优势，使用神经网络构造文本表示成为主流方法。Socher等^[1, 2]人提出使用递归神经网络构建文本表示，该方法通过人工构建的句法树，从句法树的叶子节点递归到根节点，最后在根节点生成的向量为文本表示向量。该方法被证实在构建句子级语义时较为有效。然而，递归神经网络需要按照一个树形结构来构建句子的语义，其性能依赖于构建句法树的精度。循环神经网络(Recurrent Neural Network, RNN)可以使用无结构数据构建文本表示^[3, 4]，该模型逐词处理整个文档，并把所有上下文的语义保存到一个固定大小的隐藏层中。RNN模型可以很好地捕捉上下文信息，但是其在生成文本表示时存在语义偏执，影响语义表示的精度，因此研究人员提出使用卷积神经网络(Convolutional Neural Network, CNN)^[5, 6]来避免这个问题。卷积神经网络不存在语义偏执问题，但是由于传统的CNN模型使用比较简单的卷积核(如：固定窗口)和池化层(如：Max-Pooling)，无法很好的刻画文本语义特征。因此研究人员对CNN模型进行改进，文献[7]提出采用多种卷积核提取多粒度特征，配合多种 pooling 策略，利用输入句子的多个视角，最大限度地利用句子信息构造句子表示，该方法在句子相似度任务中取得不错的结果。文献[8]提出了循环卷积网络(Recurrent Convolutional Neural Network, RCNN)构造文档表示，并将其用到文本分类中。该模型使用双向循环结构对上下文建模，在引入较少噪声的前提下，可以捕获尽可能远的上下文信息，从而保留长距离的词序信息，最后使用最大池化技术抽取文本的重要特征；该模型结合 RNN 模型和 CNN 模型各自的优点，既能很好的刻画上下文信息，又

能无语义偏执的描述文本语义。

长短期记忆网络(Long Short Term Memory Network, LSTM)^[9]的提出解决了 RNN 模型梯度爆炸和梯度消失的问题，但是其不能完全消除语义偏执问题，因此研究人员提出多种改进方法进一步解决该问题。文献[10]提出句子状态的长短记忆网络(Sentence-State Long Short Term Memory Network, S-LSTM)构造句子表示进行句子分类。该模型每一步都对所有词的隐含状态进行建模。模型将整个句子表示成一个状态，该状态由每个词的状态以及一个全局句子级状态组成，并通过循环操作进行更新。同时在每次循环操作内部，各个词也会与上下文词语进行信息交互，以获取局部的信息。文献[11]将标准的 LSTM 模型推广到树状网络拓扑中，并证明他在句子表示方面优于顺序的 LSTM 模型。标准 LSTM 模型当前步的状态是通过前一步的状态生成，而 Tree-LSTM 模型的当前节点是以其孩子作为输入进行得到。可以说标准 LSTM 模型是 Tree-LSTM 模型的一个特例，该模型生成的句子表示在句子分类和句子相似度计算方面都有不错的表现。

文本中的结构特征和关键词信息在文本表示生成的过程中至关重要。但是这类语料需要人工标注得到，不仅对专业知识的要求非常高，而且标注量一般不大，现有的标注语料很难应用到自然语言处理的各个领域。基于强化学习的文本表示模型^[12, 13]可以根据文本分类语料自动学习到文本中的重要信息。Tianyang Zhang 等^[12]人提出了利用强化学习分别学习到句子中的短语结构和无用词，并利用这些句子信息通过 LSTM 模型生成句子表示，该方法在句子分类任务中取得不错的结果。在医学问题诉求分类任务中，医学问题句子中的关键词对医学问题的诉求分类有着指导性的作用，因此抽取出句子中的关键词并把其应用到问题表示中，可以提高问题表示的质量。

本文提出一种基于强化学习的医学问题诉求分类模型，该模型可以根据医学问题诉求分类的标注结果自动识别出医学问题中的关键词，并对医学问题中的关键词和非关键词赋予不同的值构成权重向量，利用 attention 机制，对 Bi-LSTM 模型生成的隐含层状态序列使用该权重向量加权求和得到问题表示，对得到的问题表示，使用 softmax 分类器进行对其分类。

1 基于强化学习的医学问题诉求分类模型

1.1 模型整体框架

基于强化学习的医学问题分类模型由三部分构成：policy network 模块，问题表示模块和分类网络模块；policy network 模块负责对问题表示模块产生动作序列，并对不同的动作赋予不同的值并构成权重向量；问题表示模块基于 Bi-LSTM 模型，为 policy network 模块产生状态输入，利用 attention 机制，通过 policy network 模块产生的权重向量对 Bi-LSTM 模型输出的隐含层状态序列加权求和生成问题表示；分类网络模块对问题表示模块产生的问题表示进行分类，并为 policy network 模块产生延迟奖励。基于强化学习的医学问题分类模型基础框架如图 1 所示，下面对三部分具体介绍。

1.2 Police Network 模块

Police network 根据问题表示模块中 Bi-LSTM 模型产生的每一步隐含层状态作为输入生成动作序列，动作序列指导问题表示模型产生问题表示，分类网络模块根据问题表示对问题分类，根据问题分类结果产生延迟奖励。policy network 模块采用随机策略 $\pi(a_t|s_t; \theta)$ 和延迟奖励指导 policy

network 进行学习。以下具体介绍 policy network 模块的 state、action、policy、reward 和目标函数。

state: state 表示是通过当前词和当前词的上下文信息作为输入产生的隐含层状态，在本实验中，state 表示是通过 Bi-LSTM 模型产生，当前词的前向隐含层状态和后向隐含层状态拼接产生 policy network 所需的 state。

action: 根据实验需要，本文设计 action 集合为 {common, keyword}，如果一个词的动作作为“keyword”，则表示当前词为问题的关键词，“common”则表示当前词为问题的非关键词。

policy: 本实验采用随机策略，其公式表示如下：

$$\pi(a_t|s_t; \theta) = \sigma(W * s_t + b) \quad (1)$$

其中， a_t ， s_t 分别表示问题中第 t 个词语对应的 action 和 state， $\pi(a_t|s_t; \theta)$ 表示在状态 s_t 下选择动作 a_t 的概率， σ 表示 sigmoid 函数， $\theta = \{W, b\}$ 表示模型的参数。

reward: 问题表示模型在动作序列的指导下生成问题表示，分类网络根据问题表示对问题进行分类，即 $P(y|X)$ ， y 为问题标签， X 为问题表示。Reward 根据 $P(y|X)$ 获得，本实验中的奖励函数如公式 2 所示：

$$R_L = \log P(c_g|X) - \gamma L'/L \quad (2)$$

其中， c_g 表示问题标准标签， L' 为问题中关键词的个数， L 为问题包含的词语个数， γ 为参数。

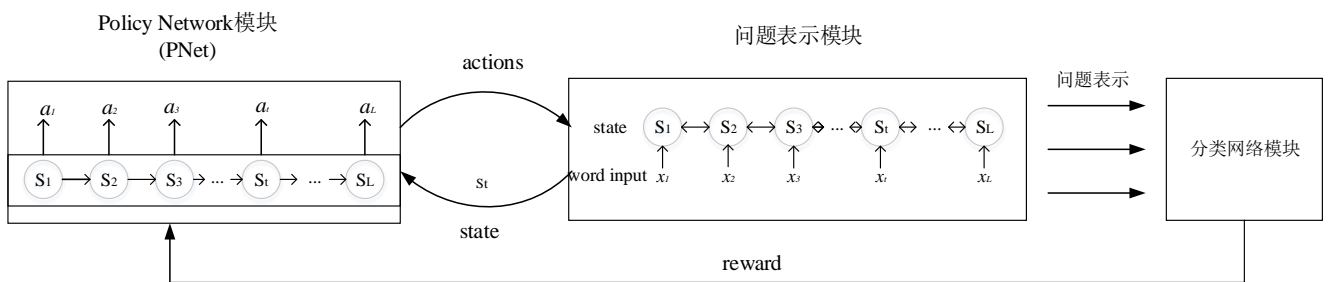


图 1 强化学习基础框架

目标函数: 本文采用 REINFORCE 算法^[14]和 policy gradient 算法^[15]对 Policy Network 的参数进行优化，目的是使得到的奖励最大化：

$$\begin{aligned} J(\theta) &= E(s_t, a_t) \sim P_\theta(s_t, a_t) r(s_1 a_1 \dots s_L, a_L) \\ &= \sum_{s_1 a_1 \dots s_L, a_L} P_\theta(s_1 a_1 \dots s_L, a_L) R_L \\ &= \sum_{s_1 a_1 \dots s_L, a_L} p(s_1) \prod_t \pi_\theta(a_t|s_t) p(s_{t+1}|s_t) R_L \end{aligned}$$

$$= \sum_{s_1 a_1 \dots s_L, a_L} \prod_t \pi_\theta(a_t|s_t) R_L \quad (3)$$

在上述公式中，第 $t+1$ 步的状态完全依靠第 t 步的状态生成，因此 $p(s_1)$ 和 $p(s_{t+1}|s_t)$ 等于 1。

本文使用梯度更新 policy 网络，具体公式如下：

$$\nabla_\theta J(\theta) = \sum_{t=1}^L R_L \nabla_\theta \log \pi_\theta(a_t|s_t) \quad (4)$$

1.3 问题表示模块

问题表示模块是一个基于 attention 的 Bi-LSTM 模型，Bi-LSTM 模型^[16]负责生成状态序列，attention 机制对生成的状态序列进行加权求和生成问题表示，本文中 attention 机制的权重向量根据 policy network 模块产生的动作序列生成。问题表示模块的具体结构如图 2 所示。

1.3.1 LSTM 模型

LSTM 模型是对 RNN 模型的一种改进，于 1997 年被提出，目的是解决 RNN 模型在反向梯度传播过程中产生的梯度爆炸和梯度消逝以及缓解 RNN 模型存在的语义偏差问题。LSTM 模型可以

通过输入门、遗忘门和输出门等控制机制有选择的保留上下文信息。LSTM 模型的主要结构可以表示为：

i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \tag{5}

f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \tag{6}

o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \tag{7}

c_t = f_t * c_{t-1} + i_t * \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \tag{8}

h_t = o_t * \tanh(c_t) \tag{9}

其中，i_t, f_t, o_t 分别表示在 t 时刻的输入门、遗忘门和输出门，W 为门的权重矩阵，b 为偏置向量，c 为记忆单元状态，σ 为 sigmoid 激活函数，* 为点乘运算，tanh 是双曲正切激活函数。通过门控记忆单元，LSTM 可以有控制地保留长距离信息，并将当前输入中的非重点内容遗忘，结合了“长时记忆”与“短时记忆”，更加适用于问题表示。

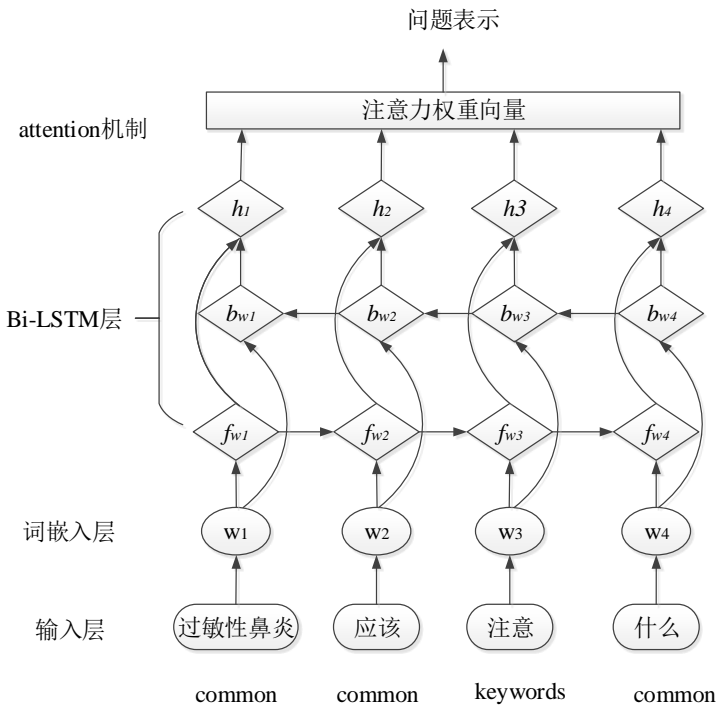


图2 基于attention机制的Bi-LSTM模型

1.3.2 基于 attention 机制的 Bi-LSTM 模型

原始的 LSTM 模型是从前往后依次传递，生成的文本表示依然存在语义偏执问题，因此采用 Bi-LSTM 模型更好的刻画文本特征。基于 attention 机制的 Bi-LSTM 模型是 Bi-LSTM 模型分别采用从前往后依次递进和从后往前依次递进两种方法计算得到两组不同的隐含层，然后利用 attention 机制进行对隐含加权求和得到问题表示^[17]。具体过程如下：

首先定义问题序列为 $S = (w_1, w_2, \dots, w_n)$ ，其中 n 表示问题中包含 n 个词， w_i 表示问题中第 i 个词的词向量。

其次，使用 S 作为 Bi-LSTM 模型的输入生成每个词语对应的隐含层状态，单步公式如下：

f_{wt} = \overrightarrow{LSTM}(w_t, f_{wt-1}) \tag{10}

b_{wt} = \overleftarrow{LSTM}(w_t, b_{wt+1}) \tag{11}

拼接每一步的 f_{wt} 和 b_{wt} 组成隐含状态 h_t ，所有的隐含层状态构成隐含层状态序列 $H =$

(h_1, h_2, \dots, h_n) 。

最后使用 attention 机制对 H 中的隐含层状态加权求和，一般的权重向量a由以下公式获得：

$$a = \text{softmax}(w_{s2} \tanh(W_{s2} H^T)) \quad (12)$$

其中， W_{s2} 为权重矩阵， w_{s2} 为参数向量。

本实验中，权重向量 a 由 Police Network 模块产生的动作序列确定，如果动作为“keyword”则该动作赋予大的权重，否则统一赋予较小的权重。

1.4 分类网络模块

分类网络模块是根据问题表示模块生成的问题表示采用 softmax 分类器对问题表示进行分类，具体公式如公式 13 所示：

$$P(y|X) = \text{softmax}(W_s h_l + b_s) \quad (13)$$

其中 $W_s \in \mathbb{R}$ 和 $b_s \in \mathbb{R}$ 是分类网络的参数。

为了训练分类网络，本文采用交叉熵作为损失函数,具体损失函数如公式 14 所示：

$$Loss = - \sum_{y=1}^K \hat{p}(y, X) \log P(y|X) \quad (14)$$

其中，K 为分类类别的个数， $\hat{p}(y, X)$ 为标准标注结果的 one-hot 分布， $P(y|X)$ 为 softmax 分类器的分类结果。

2 实验结果与分析

2.1 实验数据

实验数据主要来源于百度知道、360 问答、搜狗问问等主流中文问答平台网站。通过网络爬虫抓取平台上的问题文本，得到每个问题的链接、问题关键词和问题标题等信息。实验先后收集医学问题数据约 40000 条，经过去重、筛除不符合要求的医学问题以后选取部分语料进行标注，最终标注语料为 8926 句，8056 个问题作为训练集，870 个问题作为测试集。在语料标注过程中，我们人工标注总结每个问题中的关键词，并对问题中的关键词根据问题分类结果进行分类。

2.2 语料前期处理

分词：由于医学问题中存在大量的医学专有

词汇，现有的分词工具无法很好的对其分词，因此我们整理医学词汇词典辅助分词。医学词汇词典主要包括：药品名称、症状名称和疾病名称。其中词典中包含 17513 个药品名称、1101 个症状名称、4538 个疾病名称。

词向量：本文使用网络爬虫获得的约 40000 条问题进行训练词向量，首先对这些问题进行分词，然后使用 word2vec 模型^[18]对其训练获得词向量。

2.3 参数设置

本文中采用 word2vec 模型生成的词向量参与训练，词向量的维度为 70 维；采用 Adam 优化算法对模型进行优化；dropout 层参数设置为 0.5；学习率设置为 0.0005，batchsize 设置为 10，epoch 大小为 50。

2.4 实验评价指标

本文的实验结果采用准确率（accuracy）作为评价指标，其公式具体如下

$$\text{准确率}(\text{accuracy}) = \frac{\text{分类正确的问题数}}{\text{问题总数}} \times 100\% \quad (15)$$

2.5 实验结果及分析

在医学问题诉求分类任务中，问题表示的好坏直接影响诉求分类结果。问题中的关键词在问题分类时有着重要作用，合理的把关键词信息应用到问题表示中可以提高问题表示的质量。本文中根据问题表示构造方法的不同选择不同的方法将关键词信息加入到问题表示中。

- 1. 组合函数构造问题表示：由于组合函数构造问题表示的方法是问题中的词向量通过组合函数直接得到，无法直接利用关键词信息构造问题表示。因此构造类别 one-hot 向量与问题表示向量相拼接，以此对问题表示加入关键词信息。类别 one-hot 向量是根据诉求分类类别确定，每个类别对应一个 one-hot 向量，当问题中出现某一类别的关键词时，其通过组合函数得到的问题表示后面拼接相应的类

- 别 one-hot 向量。
2. 基于 CNN 模型，S-LSTM 模型的问题表示：在实验中，通过半泛化的形式加入关键词信息。即在问题中每个关键词后面加入相应的类别标签，类别标签看作问题中独立的词参与问题表示的生成，以加强关键词在问题表示生成过程中的作用。类别标签向量是通过 word2vec 模型生成，即在训练词向量的问题语料中加入类别标签，将类别标签看作问题中独立的词进行训练得到类别标签向量。
3. 基于 Bi-LSTM 模型，LSTM 模型的问题表示：在实验中，赋予问题中的关键词和非关键词不同的权重构成权重向量，利用 attention 机制，对 Bi-LSTM 模型和 LSTM 模型生成的隐含层状态序列进行加权求和得到问题表示。

表 1 加入关键词信息和不加关键词信息的模型结果对比

	准确率（accuracy）×100%	
基础模型	无关键词信息	加入关键词信息
组合函数	81.26	87.36
CNN	85.86	87.24
LSTM	86.90	89.77
S-LSTM	87.59	89.31
Bi-LSTM	89.43*	91.03*

由表 1 可知，在不加关键词信息的问题表示模型中，Bi-LSTM 模型取得最好的结果，比结果第二高的 S-LSTM 模型高出 1.84%。基于 LSTM 模型、基于 CNN 模型和基于 S-LSTM 模型的到的分类结果明显好于通过组合函数构造问题表示的方法。本实验中采用的组合函数为相加，即对问题中的所有词向量相加得到问题向量。对比添加关键词信息和没有添加关键词信息构造问题表示的分类结果，前者的分类结果明显高于后者，分类结果最高的模型是加入关键词信息的 Bi-LSTM 模型，该模型通过问题中的关键词和非关键词构造的权重向量对 Bi-LSTM 模型的隐含层状态序列加权求和获得问题表示，其分类结果比 Bi-LSTM 模型提高 1.6%，比加入关键词信息的 S-LSTM 模型提高 1.72%。

由以上实验可知，在问题表示生成过程中加入关键词信息可以提高医学问题诉求分类的结果，但是面对庞大的互联网语料如果仅仅依靠人工提取关键词代价太大，因此本文提出使用强化学习自动识别问题中的关键词，根据是识别出的关键词构造问题表示。

表 2 不同的权重向量对 Bi-LSTM 模型隐含层状态加权求和得到的分类结果对比

模型	准确率（accuracy）×100%
Bi-LSTM + 人工关键词	91.03*
Bi-LSTM+attention	90.11
强化学习	90.92

由表 2 可知，基于强化学习的医学问题诉求分类模型得到的结果虽然没有加入人工总结关键词信息的 Bi-LSTM 模型高，但是仅仅低了 0.11%。另外，其结果比 Bi-LSTM+attention 模型的结果高了 0.81%，比基于 Bi-LSTM 模型高了 1.49%。基于强化学习的方法识别关键词的准确率虽然还没有达到人工总结的高度，但是其在医学问题诉求分类任务中得到的分类结果已经和加入人工总结关键词信息的分类结果非常接近。

表 3 不同模型中各诉求类别的分类结果对比

	准确率（accuracy）×100%			
类别	Bi-LSTM + 人工关键词	Bi-LSTM+attention	强化学习	测试总数
怎么治	95.3	96.4*	94.9	554
会怎么样	74.5	67.2	76.3*	55
什么病	87.9*	85.4	86.9	199
为什么	87.0	78.2	91.3*	23
注意事项	83.3	76.7	86.7*	30
医学无关问题	33.3*	11.1	33.3*	9

表 3 对不同模型中的每个诉求分类具体结果进行展示。由表 3 可知，基于强化学习的诉求分类模型在“怎么治”类别中的分类结果低于其他两个模型；但在“会怎么样”，“为什么”，“注意事项”三个类别分类中，基于强化学习的诉求分类模型的结果明显高于 Bi-LSTM+attention 模型的结果和略高于加入人工总结关键词信息的 Bi-LSTM 模型的结果。实验表明，基于强化学习的诉求分类模型对语料中的少数类有比较好的分

类结果。

3 总结与展望

本文使用多种方法将关键词信息加入问题表示中,实验证明,加入关键词信息的分类结果明显高于基础模型的分类结果。为了通过机器学习自动获取关键词信息,本文提出基于强化学习的医学问题诉求分类模型,该模型首先根据 policy network 产生动作序列,根据动作序列中动作的不同赋予不同的值构造权重向量;其次利用权重向量对 Bi-LSTM 模型生成的隐含层状态加权求和得到问题向量表示;最后利用 softmax 分类器对问题向量进行分类。该模型自动识别医学问题中的关键词并加强关键词在构造问题表示中的作用,在医学问题诉求分类任务中取得相对较高的结果。

本文主要针对单诉求的医学问题进行分类,但是有些医学问题包含多种诉求,利用单诉求分类模型无法解决,下一步我们将设计一种多诉求分类模型,通过强化学习识别一个问题中多种类别的关键词进行问题的多诉求分类。

参考文献

[1] Richard Socher, Alex Perelygin, Jean Wu, Jason Chuang, Christopher D. Manning, Andrew Ng, and Christopher Potts. Recursive deep models for semantic compositionality over a sentiment treebank. In Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing (EMNLP13), pages 1631–1642, 2013.

[2] Socher R , Huval B , Manning C D , et al. Semantic Compositionality through Recursive Matrix-Vector Spaces[C]// Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning. 2012.

[3] Cho K, Van Merriënboer B, Gulcehre C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[J]. arXiv preprint arXiv:1406.1078, 2014.

[4] Hill F, Cho K, Korhonen A. Learning distributed representations of sentences from unlabelled data[J]. arXiv preprint arXiv:1602.03483, 2016.

[5] Yin W, Kann K, Yu M, et al. Comparative study of cnn and rnn for natural language processing[J]. arXiv preprint arXiv:1702.01923, 2017.

[6] Wen Y, Zhang W, Luo R, et al. Learning text representation using recurrent convolutional neural network with highway layers[J]. arXiv preprint arXiv:1606.06905, 2016.

[7] He H, Gimpel K, Lin J. Multi-perspective sentence similarity modeling with convolutional neural networks. Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. 2015:

1576-1586.

[8] Siwei Lai, Liheng Xu, Kang Liu, and Jun Zhao. Recurrent Convolutional Neural Networks for Text Classification. In Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI'15), pages 2267–2273. 2015.

[9] Sepp Hochreiter, Jürgen Schmidhuber. Long short-term memory. Neural computation, 9(8):1735–1780, 1997.

[10] Zhang Y, Liu Q, Song L. Sentence-State LSTM for Text Representation[J]. 2018.

[11] Kai Sheng Tai, Richard Socher, and Christopher D. Manning. Improved semantic representations from tree-structured long short-term memory networks[C]. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics. 2015

[12] Zhang T, Huang M, Zhao L. Learning structured representation for text classification via reinforcement learning[C]//Thirty-Second AAAI Conference on Artificial Intelligence. 2018.

[13] Yogatama D, Blunsom P, Dyer C, et al. Learning to compose words into sentences with reinforcement learning[C]. ICLR :2017

[14] Williams R J. Simple statistical gradient-following algorithms for connectionist reinforcement learning[J]. Machine learning, 1992, 8(3-4): 229-256.

[15] Sutton R S, McAllester D A, Singh S P, et al. Policy gradient methods for reinforcement learning with function approximation[C]//Advances in neural information processing systems. 2000: 1057-1063.

[16] Graves A, Mohamed A R, Hinton G. Speech recognition with deep recurrent neural networks[J]. 2013,38(2003): 6645-6649.

[17] Lin Z, Feng M, Santos C N, et al. A structured self-attentive sentence Embedding[C]. International Conference on Learning Representations. 2017.

[18] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed representations of words and phrases and their compositionality. In Advances in neural information processing systems, pages 3111–3119, 2013.