

基于注意力机制的生物医学文本分类模型

李启行, 廖薇

上海工程技术大学电子电气工程学院, 上海 201620

【摘要】对生物医学文本进行准确分类,是促进医院信息化发展的一个重要途径。本研究提出一种基于注意力机制的双层次文本分类模型,用于对生物医学文本进行有效分类。该模型结合卷积神经网络与循环神经网络的优势,对用户输入的疾疾病文本进行特征提取。首先,在第一层次通过Bi-GRU通道与Bi-LSTM通道提取文本中的上下文关联信息,同时,为增强模型的特征提取能力,在该层次引入注意力机制。其次,将两个通道提取到的时序特征进行特征拼接,并将拼接后的结果传入第二层次,从而进一步提取文本的局部特征,最后利用分类器输出最终的分类结果。对生物医学文本进行分类性能评估,结果表明,与基线模型相比,该模型的分类准确率可达91.45%,具有显著的分类性能。

【关键词】生物医学文本;注意力机制;卷积神经网络;循环神经网络;文本分类

【中图分类号】R318;TP391

【文献标志码】A

【文章编号】1005-202X(2022)04-0518-06

Biomedical text classification model based on attention mechanism

LI Qihang, LIAO Wei

School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China

Abstract: The accurate classification of biomedical texts is an important way to promote the development of hospital information. Herein a two-level text classification model based on attentional mechanism is proposed to effectively classify biomedical texts. The model combines the advantages of convolutional neural network and recurrent neural network to extract features from the disease texts input by users. The context association information in the text is firstly extracted through Bi-GRU channel and Bi-LSTM channel at the first level, and meanwhile, in order to enhance the feature extraction ability of the model, attention mechanism is introduced to this level. After the time series features extracted from the two channels are spliced-together, the spliced-result is input to the second level for further extracting the local features from the text. The final classification results are output by the classifier. The evaluation of the performances in classifying biomedical texts show that compared with baseline models, the proposed model achieves a classification accuracy of 91.45%, with a significant classification performance.

Keywords: biomedical text; attention mechanism; convolutional neural network; recurrent neural network; text classification

前言

近年来,随着互联网技术的不断发展,如何让数据更好的为人类服务变得越来越重要。医疗机构每天都会涌现出大量的医疗数据资源,仅仅依靠人力已经远远无法对这些数字资源进行有效分类,因此,实现生物医学文本的自动分类已经成为目前医疗健康领域重要的一部分,其中最基本的一个应用就是

网络问诊。网络问诊主要是患者通过互联网的方式来查询自身疾病并找寻答案,这个过程涉及到科室选择。如何将患者的疾病文本进行准确的分类,进而为患者正确匹配就诊科室的类别,是目前的一个研究热点。

在对生物医学文本进行分类时所用到的分类模型主要有两大类,第一类是采用支持向量机(Support Vector Machine, SVM)^[1-3]、隐含狄利克雷分布(Latent Dirichlet Allocation, LDA)^[4-5]和朴素贝叶斯^[6]等传统的分类方法,这种方法有两个主要缺点,一是在训练时特征空间稀疏,并且特征的维数较高,这就导致模型的性能偏低,另一个缺点就是在文本特征的提取过程中更偏向于人工提取,整个过程费时费力^[7-8]。第二类就是由卷积神经网络(Convolutional

【收稿日期】2021-11-05

【基金项目】国家自然科学基金(62001282)

【作者简介】李启行,硕士研究生,研究方向:自然语言处理, E-mail: 599510114@qq.com

【通信作者】廖薇,博士,副教授,研究方向:生物医疗与自然语言处理, E-mail: liaowei54@126.com

Neural Networks, CNN)^[9]与循环神经网络(Recurrent Neural Network, RNN)^[10]构成的基于深度学习的分类模型,相对于第一类分类模型,该类模型具有很大的应用优势。

本文旨在解决对生物医学文本进行科室分类问题,输入为一系列疾病文本的描述句子,输出为该疾病文本对应的科室类别。为缓解现有文本分类模型架构的局限性,提出一种基于注意力机制的双层次文本分类模型(Dual layer R-CNN, DR-CNN),用于解决医院中预检分诊的科室分类问题。在DR-CNN中,通过层次化的方式,结合词嵌入、双向长短期记忆网络(Bi-directional Long Short-Term Memory, Bi-LSTM)、双向门控循环单元网络(Bi-directional Gate Recurrent Unit, Bi-GRU)、注意力机制以及CNN的优势,来获取生物医学文本中的全局序列信息与局部特征,既考虑了文本的前向和后向上下文依赖关系,又进一步关注了文本中的重要特征。

1 研究基础

文本分类技术是自然语言处理领域的核心内容之一^[11],主要任务是将文本快速准确地分配到对应的标签。目前,随着深度学习方法的不断发展,深度学习模型在文本分类上的应用取得了很好的进展^[12-14],在意图识别、问答系统以及舆情分析等多个领域都有着广泛的应用^[15-17]。

CNN与RNN在对疾病文本进行文本分类的相关研究中较为常见,这是由于CNN可以提取疾病文本中的局部特征,RNN可以提取疾病文本中的全局序列信息。尽管RNN适用于对文本特征的提取,但是当文本数据中存在长期依赖关系时,会出现梯度爆炸或消失的问题^[18],为解决这一问题,长短期记忆网络(LSTM)和门控循环单元网络(GRU)被引入。同时考虑到上下文,Bi-LSTM和Bi-GRU被提出,通过组合前向与后向隐藏层,可以更好地解决顺序建模问题^[19]。

刘勘等^[20]构建一种基于CNN的文本分类模型,并将该模型用于辅助解决医院科室分配问题,结果表明了单一深度学习模型的有效性,解决了医院分诊的实际问题。张强强等^[21]提出了一种基于CNN和SVM的疾病症状分类模型,模型结合了CNN与SVM的优势,最终通过与传统特征提取方法的对比实验得出,基于深度学习的分类模型具有更高的准确率。陈德鑫等^[22]通过构建基于混合深度学习模型CNN-Bi-LSTM,来探究模型对医疗数据的实体抽取结果,并用对比实验证明了使用混合模型比单模型取得更好的效果。虽然Bi-LSTM和Bi-GRU在文本分类中

得到了广泛的应用,但它们不能集中提取文本中上下文信息中的重要部分,同时增加了文本特征的维度,使得模型难以优化。

2 DR-CNN模型的构建

2.1 模型结构

本文所提的DR-CNN模型结构如图1所示,以疾病文本作为输入,科室类别作为输出,实现端对端的中文文本分类任务。整个模型可以分为以下4个部分。(1)词嵌入层:将分词后的疾病文本向量化表示;(2)第一层次:该部分包括Bi-LSTM+Attention与Bi-GRU+Attention两个模块,获取文本序列的上下文语义信息,同时引入注意力机制来关注文本中的重要特征;(3)第二层次:该层次由CNN构成,获取文本序列的局部特征,通过最大值池化降低特征空间的维数;(4)Softmax层:对融合后的特征进行分类,获取分类结果。

2.2 文本表示

词向量的生成方式一般分为两种,一种是随机

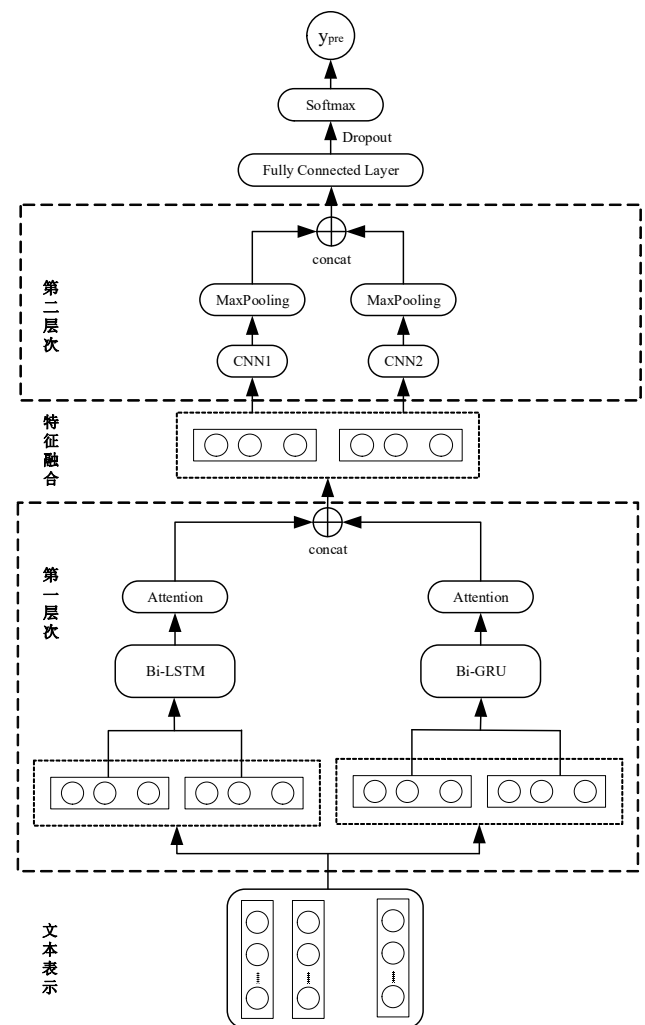


图1 DR-CNN文本分类模型

Figure 1 DR-CNN text classification model

生成的,另一种是通过神经语言模型学习到的,后者生成的词向量包含了不同词语之间的语义联系,可以进一步增强文本表示的表达能力。本文使用Google开源推出的Word2Vec工具包进行词向量训练,Word2Vec中包含CBOW和Skip-gram两个模型,这两个神经语言模型均属于浅层神经网络,在对生物医学文本数据库中的文本使用结巴分词工具处理后,使用Word2Vec中的Skip-gram模型进行词向量的预训练。Skip-gram模型简单来说就是通过目标词来预测上下文,同时可以学习到词语与词语之间的联系。假设一个句子 S 的长度为 N ,整个文本可以向量化表示为下式:

$$S = [w_1, w_2, w_3, \dots, w_N] \quad (1)$$

2.3 第一层次

模型的第一层次采用Bi-LSTM和Bi-GRU两个并行通道对任意长度的文本序列进行处理,提取出前后两个方向的依赖关系。我们使用GRU和LSTM同时学习文本的长序列特征和短序列特征,并在两个并行通道中引入注意力机制来强化模型的特征提取能力。

LSTM是一种特殊的RNN,主要为了解决RNN面临的梯度爆炸或消失等问题。每个LSTM单元由3个非线性门组成,包括遗忘门、输入门和输出门,这些门有各自的功能,遗忘门决定从单元中需要遗忘哪些信息,输入门决定将哪些信息更新为单元状态,输出门决定整个单元的哪些部分被输出。整个过程中的节点状态由式(2)~(7)决定:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (3)$$

$$\tilde{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (4)$$

$$c_t = f_t \otimes c_{t-1} + i_t \otimes \tilde{c}_t \quad (5)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (6)$$

$$h_t = o_t \otimes \tanh(c_t) \quad (7)$$

其中, σ 表示Sigmoid函数, f_t 、 i_t 、 o_t 分别表示 t 时刻的遗忘门、输入门与输出门的节点操作, W 为权重矩阵, b 为偏置向量, h_{t-1} 表示前一步产生的状态, c_{t-1} 表示上一步输出的单元状态, c_t 是当前单元的状态, x_t 和 h_t 为时间 t 的输入向量和隐藏状态向量,运算符 \otimes 表示逐元素乘积。

GRU是LSTM的一个变体,它只有更新门和重置门两个门,其中更新门结合了LSTM的遗忘门与输入门,GRU相对于LSTM来说更简单,更新和重置的过程见式(8)~(9):

$$r_t = \sigma(W_r h_{t-1} + U_r x_t + b_r) \quad (8)$$

$$z_t = \sigma(W_z h_{t-1} + U_z x_t + b_z) \quad (9)$$

其中, σ 为Sigmoid函数, W 和 U 表示权重矩阵, b 表示偏置向量, x_t 为第 t 层的输入, h_{t-1} 表示上一时刻的状态。重置门决定什么时候应该忽略当前的隐藏状态,更新门决定应该传递到当前状态的信息量。候选状态 \tilde{h}_t 和当前状态 h_t 可以表示为:

$$h_t = (1 - z_t) \otimes h_{t-1} + z_t \otimes \tilde{h}_t \quad (10)$$

$$\tilde{h}_t = \tanh(W_h (h_{t-1} \otimes r_t) + U_h x_t) \quad (11)$$

其中,运算符 \otimes 表示逐元素乘积, z_t 为更新门获取的信息, \tilde{h}_t 表示当前候选状态,其他同上。

标准的LSTM与GRU网络均为单向的神经网络,这就表明当前的状态只能按从前往后的顺序来进行输出,但对于文本数据来说,当前时刻的输出与过去和未来的整体信息均存在联系,因而本文采用包含前向隐藏层与后向隐藏层的Bi-LSTM与Bi-GRU来处理两个方向的序列,双向网络结构图如图2所示。Bi-LSTM前向隐藏层的输出为 \overrightarrow{h}_{ul} ,后向隐藏层的输出为 \overleftarrow{h}_{ul} ,Bi-GRU前向隐藏层的输出为 \overrightarrow{h}_{ig} ,后向隐藏层的输出为 \overleftarrow{h}_{ig} ,则Bi-LSTM的输出 H_{ul} 与Bi-GRU的输出 H_{ig} 可以表示为:

$$H_{ul} = [\overrightarrow{h}_{ul}, \overleftarrow{h}_{ul}] \quad (12)$$

$$H_{ig} = [\overrightarrow{h}_{ig}, \overleftarrow{h}_{ig}] \quad (13)$$

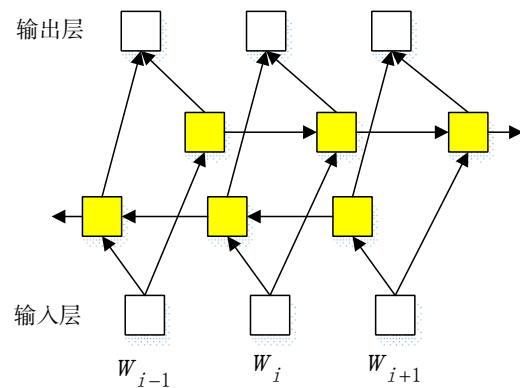


图2 双向网络结构图

Figure 2 Structure diagram of bidirectional neural network

注意力机制可以被用来为不同的词赋予不同权重,因为文本中不同的词语对文本所表达的意思有着不同的贡献,不同的权重可以体现出每个词语对整个句子语义的贡献程度。通过注意力机制来捕获句子中词语之间的句法特征或者语义特征,可以进一步提高模型对文本的理解。对Bi-LSTM和Bi-GRU两个并行层添加注意力机制的数学计算公式如下:

$$u_{iLSTM} = \tanh(W_{sL}H_{iL} + b_{sL}) \quad (14)$$

$$u_{iGRU} = \tanh(W_{sG}H_{iG} + b_{sG}) \quad (15)$$

$$\alpha_{iLSTM} = \text{Soft max}(u_{iLSTM}^T, u_{sL}) \quad (16)$$

$$\alpha_{iGRU} = \text{Soft max}(u_{iGRU}^T, u_{sG}) \quad (17)$$

$$T_{LSTM} = \sum_i \alpha_{iLSTM} H_{iL} \quad (18)$$

$$T_{GRU} = \sum_i \alpha_{iGRU} H_{iG} \quad (19)$$

其中, W_s 为权重矩阵, u_i 为 H_i 的注意力隐层表示, α_i 是 u_i 通过 Softmax 转换后的归一化权重, 然后将权重 α_i 与隐藏层的输出 H_i 进行点乘与累加操作, 最后将 Bi-LSTM 和 Bi-GRU 两个并行层的输出进行合并, 合并过程如下式, 得到第一层次的输出 T :

$$T = \text{tf.concat}([T_{LSTM}, T_{GRU}], \text{axis} = 1) \quad (20)$$

其中, tf.concat 的作用是对 T_{LSTM} 与 T_{GRU} 进行拼接操作, $\text{axis} = 1$ 表示对 T_{LSTM} 与 T_{GRU} 进行横向拼接, 不改变词向量的维度, 只改变词向量的个数, 因而大大增加了模型捕获文本特征的能力。

2.4 第二层次

该层次由 CNN 模型构成, CNN 模型主要由输入层、卷积层以及池化层组成, 其中卷积层为 CNN 模型的核心, 用于提取局部特征, 本模型采用两个不同尺寸的卷积核作为卷积层, 分别为 CNN1 与 CNN2。池化层位于卷积层之后, 可以对高维特征进行降维操作, 本模型采用最大值池化(Max-pooling)来提取特征值。取卷积核 $W = [w_0, w_1, \dots, w_{\beta-1}]$, 卷积过程的数学计算公式如下:

$$C_t = g(W \times X_{t:t+\beta-1} + b) \quad (21)$$

其中, g 为非线性激活函数, 本模型采用 ReLU 激活函数, b 为偏置项, t 表示卷积核滑动窗口的位置, $X_{t:t+\beta-1}$ 为窗口从 t 移动到 $t + \beta - 1$ 所包含的词向量矩阵。

将第一层次输出的结果 T 进行输入, 使用两个不同尺寸的卷积核进行特征提取, 并通过最大值池化的方法提取出对应的文本特征 C_1 与 C_2 , 最后将 C_1 与 C_2 进行拼接操作, 可以得到输出向量 C :

$$C = \text{tf.concat}([C_1, C_2], \text{axis} = 1) \quad (22)$$

2.5 Softmax 层

全连接层将输入的向量特征重新拟合, 在全连接层之后引入 Dropout 机制, 防止训练过程中出现过拟合现象。接着将处理后的向量特征作为 Softmax 分类器的输入, 模型最终由 Softmax 分类器输出类别的概率分布。将 x 分类为类别 j 的概率公式如下:

$$P(y^{(i)} = j | x^{(i)}; \theta) = \frac{\exp(\theta_j^T x^{(i)})}{\sum_{n=1}^k \exp(\theta_n^T x^{(i)})} \quad (23)$$

3 性能评测与分析

3.1 实验数据

本文使用中文医疗对话数据集进行实验, 用来测试 DR-CNN 文本分类模型的性能, 该数据集来自网络问诊平台, 本文选用其中 8 个类别的样本进行实验, 科室类别如表 1 所示, 每个类别选取 5 000 条样本作为训练集, 500 条作为验证集, 1 000 条为测试集。

表 1 数据集类别统计

Table 1 Data set category statistics

类别	缩写	类别	缩写
肛肠科	K1	乳腺科	K5
妇产科	K2	肝病科	K6
儿科	K3	呼吸科	K7
泌尿科	K4	心血管科	K8

3.2 实验平台与参数设置

本文实验环境配置如下: 操作系统为 Windows10, 显卡为 Quadro M6000, CPU 使用的是 Intel Xeon Silver 4114, 开发框架、开发语言以及开发工具分别为 TensorFlow-gpu 1.12.0、Python3.6 与 Pycharm。

本文模型参数如表 2 所示, 其中 kernel_size 设为 (3, 4) 表示两个卷积核的尺寸分别为 3 与 4, 使用 Adam 作为模型优化器。

表 2 实验参数设置

Table 2 Experimental parameter setting

参数名称	参数值	参数名称	参数值
embedding_dim	100	num_filters	100
vocab_size	6 000	attention_size	100
seq_length	30	kernel_size	3, 4
num_classes	8	batch_size	64
hidden_dim	128	keep_prob	0.5

3.3 实验结果与分析

3.3.1 模型性能分析 分类模型的评价指标主要采用准确率(Acc)、精确率(P)、召回率(R)和 F1 值。为了验证 DR-CNN 模型在生物医学文本分类任务上的分类效果, 对 DR-CNN 模型及其变体 DR-CNN-A 模型进行测试, 两者的区别在于后者在第一层次内未引入注意力机制。测试时使两个模型的实验参数保持一致, 测试结果如图 3 所示, 纵坐标为这些类别在测试集上的精确率, 单位为%。

通过图3可以看出,DR-CNN模型在该项测试中取得了很好的分类效果。其中5个科室类别的精确率超过了95%,乳腺科与肝病科的分类精确率更是达到了97%以上,只有呼吸科的分类精确率处于85%以下,造成这种状况的原因可以归纳为两种,一是数据集中涵盖了许多专业词语,另一个就是数据集中包含了大量的非正式语言,这都使得模型不能充分的学习到相应的文本特征,从而影响了分类性能。此外,从图3中可以明显的发现DR-CNN模型比未引入注意力机制的DR-CNN-A模型具有更好的分类效果,其中妇产科类别的精确率更是提高了4%以上。以上结果不仅证明了注意力机制可以改善模型的性能,更是体现了本文模型具有良好的分类效果。

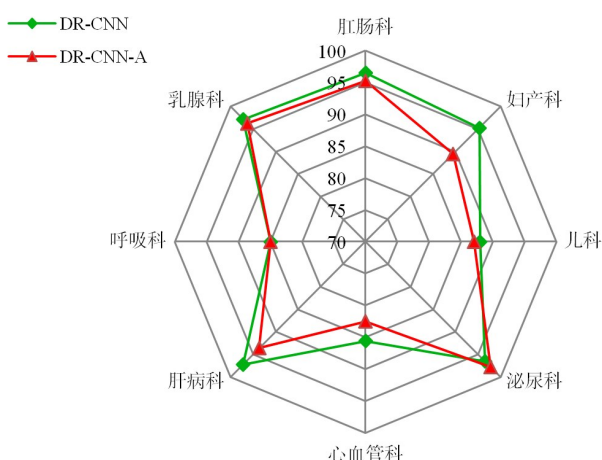


图3 性能测试结果(精确率,%)

Figure 3 Performance test results (Precision rate, %)

3.3.2 各层次对模型性能的影响 为了更清楚的探究模型中两个层次对模型性能的影响,本文在中文医疗对话数据集上设置对照实验,分别设置单独的层次一模型与层次二模型,测试结果如表3所示。

表3 各层次对模型性能的影响(%)

Table 3 Effect of each level on model performances (%)

模型	准确率	召回率	F1值
第一层次	90.76	90.76	90.83
第二层次	90.96	90.96	91.03
DR-CNN	91.45	91.45	91.55

从表3可以看出,使用第一层次和第二层次的单模型对疾病文本进行分类时模型的性能偏低,这表明同时使用两个层次对疾病文本进行特征提取时能够学习到更加充分的特征。因此,上述测试结果更能说明本文所提出的DR-CNN模型具有很好的分类可行性。

3.3.3 对比实验分析 为了进一步验证本文所提模型在文本分类方面的性能,本小节设置多个分类模型进行对比实验,实验均在中文医疗对话数据集上进行。本节设置的对比实验,不仅包括了基于深度学习的单模型,也与其他混合模型进行了对比。基线模型包括:CNN、LSTM、Bi-LSTM^[23]、Bi-LSTM+Att模型^[24]、CNN-LSTM模型^[25]、Bi-LSTM-CNN+Att模型^[26]。其中Bi-LSTM+Att模型在Bi-LSTM模型的基础上引入了注意力机制;CNN-LSTM模型采用单通道的形式将CNN与LSTM进行结合;Bi-LSTM-CNN+Att模型结合了Bi-LSTM、CNN以及注意力机制的优势对文本进行特征提取,参数与DR-CNN中对应参数保持一致。实验结果如表4所示。

表4 各模型准确率对比(%)

Table 4 Comparison of accuracy among different models (%)

模型	准确率	召回率	F1值
CNN	89.71	89.64	89.66
LSTM	88.56	88.41	88.45
Bi-LSTM	89.11	88.99	89.05
CNN-LSTM	89.67	89.56	89.62
Bi-LSTM+Att	90.26	90.18	90.22
Bi-LSTM-CNN+Att	90.81	90.81	90.87
DR-CNN	91.45	91.45	91.55

从上述结果可以看出,本文所提模型与CNN、LSTM等基于深度学习的单模型相比,准确率分别提高了1.74%与2.83%,通过比较CNN-LSTM模型与单模型的结果,可以清晰的证明CNN与RNN结合后的深度学习神经网络,更能够充分提取文本中的关键特征。对比模型Bi-LSTM+Att与模型Bi-LSTM的测试结果,分类准确率、召回率以及F1值分别提高了1.15%、1.19%以及1.17%,证明注意力机制的引入可以大大改善模型的性能。将本模型DR-CNN的测试结果与模型Bi-LSTM-CNN+Att的测试结果进行对比,DR-CNN模型的性能比Bi-LSTM-CNN+Att模型提高了0.64%,由此证明了Bi-LSTM与Bi-GRU结合后可以显著地提升模型的特征提取能力。综上所述,通过在同一数据集下与基线模型的性能进行比较,进一步验证了本文所提出的DR-CNN模型在进行文本分类任务时有更好的性能。

4 结束语

本文旨在解决对生物医学文本进行科室分类问题,利用深度学习模型,进一步缩小预检分诊时分配对应科室的误差,输入的是一系列疾病文本的描述句子,输出为该疾病文本对应的科室类别。为解决现有文本分类模型架构的局限性,提出一种基于注意力机制的生物医学文本分类模型(DR-CNN)。该模型通过层次化的方式,结合词嵌入、Bi-LSTM、Bi-GRU、注意力机制以及CNN的优势,来获取生物医学文本中的上下文关联信息与局部特征。通过与其他文本分类模型进行对比实验,结果表明,本文所提模型取得了更好的分类效果。在接下来的研究中,可以从数据库的完善、文本分类模型的优化以及分诊系统的构建出发,进一步解决患者初检时能快速找到对应科室的实际问题。

【参考文献】

- [1] 雷炳业, 潘嘉瑜, 吴逢春, 等. 基于机器学习的神经精神疾病辅助诊断研究进展[J]. 中国医学物理学杂志, 2020, 37(2): 257-264.
LEI B Y, PAN J Y, WU F C, et al. Advances in auxiliary diagnosis of neuropsychiatric diseases based on machine learning[J]. Chinese Journal of Medical Physics, 2020, 37(2): 257-264.
- [2] 葛晓伟, 李凯霞, 程铭. 基于CNN-SVM的护理不良事件文本分类研究[J]. 计算机工程与科学, 2020, 42(1): 161-166.
GE X W, LI K X, CHENG M. Text classification of nursing adverse events based on CNN-SVM[J]. Computer Engineering & Science, 2020, 42(1): 161-166.
- [3] ZHANG M, AI X, HU Y. Chinese text classification system on regulatory information based on SVM[C]//IOP Conference Series: Earth and Environmental Science. IOP Publishing, 2019, 252(2): 022133.
- [4] CHEN Q, YAO L, YANG J. Short text classification based on LDA topic model[C]//2016 International Conference on Audio, Language and Image Processing (ICALIP). IEEE, 2016: 749-753.
- [5] 陈欢, 黄勃, 朱翌民, 等. 结合LDA与Self-Attention的短文本情感分类方法[J]. 计算机工程与应用, 2020, 56(18): 165-170.
CHEN H, HUANG B, ZHU Y M, et al. Short text emotion classification method combining LDA and Self-Attention[J]. Computer Engineering and Applications, 2020, 56(18): 165-170.
- [6] SARITAS M M, YASAR A. Performance analysis of ANN and naive bayes classification algorithm for data classification[J]. Int J Intell Syst, 2019, 7(2): 88-91.
- [7] WANG R J, WANG G. Web text categorization based on statistical merging algorithm in big data environment[J]. Int J Ambient Comput, 2019, 10(3): 17-32.
- [8] SALLOUM S A, AL-EMRAN M, MONEM A A, et al. Using text mining techniques for extracting information from research articles [M]//Intelligent natural language processing: Trends and Applications. Springer, Cham, 2018: 373-397.
- [9] KIM Y. Convolutional neural networks for sentence classification[J]. arXiv preprint, arXiv:1408.5882, 2014.
- [10] LI Y, WANG X, XU P. Chinese text classification model based on deep learning[J]. Future Internet, 2018, 10(11): 113.
- [11] 王海涛, 宋文, 王辉. 一种基于LSTM和CNN混合模型的文本分类方法[J]. 小型微型计算机系统, 2020, 41(6): 1163-1168.
WANG H T, SONG W, WANG H. Text classification method based on hybrid model of LSTM and CNN[J]. Journal of Chinese Computer Systems, 2020, 41(6): 1163-1168.
- [12] WANG R, LI Z, CAO J, et al. Convolutional recurrent neural networks for text classification[C]//2019 International Joint Conference on Neural Networks (IJCNN). IEEE, 2019: 1-6.
- [13] 王超凡, 琚生根, 孙界平, 等. 融入多尺度特征注意力的胶囊神经网络及其在文本分类中的应用[J]. 中文信息学报, 2022, 36(1): 65-74.
WANG C F, JU S G, SUN J P, et al. Capsule network with multi-scale feature attention for text classification[J]. Journal of Chinese Information Processing, 2022, 36(1): 65-74.
- [14] MINAEI S, KALCHBRENNER N, CAMBRIA E, et al. Deep learning--based text classification: a comprehensive review [J]. arXiv preprint, arXiv: 2004.03705v2, 2020.
- [15] SARROUTI M, EL ALAOUI S O. A machine learning-based method for question type classification in biomedical question answering[J]. Methods Inf Med, 2017, 56(3): 209-216.
- [16] 张民航, 蔡东风, 李绍鸣. 面向社交群问答对获取的深度学习模型[J]. 计算机应用与软件, 2022, 39(3): 172-179.
ZHANG M H, CAI D F, LI S M. Deep learning model for acquisition of Q & A pairs in social groups[J]. Computer Applications and Software, 2022, 39(3): 172-179.
- [17] EE H D, HU Y M, PENG H P, et al. Theme and sentiment analysis model of public opinion dissemination based on generative adversarial network[J]. Chaos Soliton Fract, 2019, 121(1): 160-167.
- [18] DU J, VONG C M, CHEN C L. Novel efficient RNN and LSTM-Like architectures: recurrent and gated broad learning systems and their applications for text classification[J]. IEEE Trans Cybern, 2021, 51(3): 1586-1597.
- [19] ZHANG J R, LIU F A, XU W Z, et al. Feature fusion text classification model combining CNN and BiGRU with multi-attention mechanism[J]. Future Internet, 2019, 11(11): 237.
- [20] 刘勤, 陈露. 面向医疗分诊的深度学习神经网络学习[J]. 数据分析与知识发现, 2019, 3(6): 99-108.
LIU K, CHEN L. Deep neural network learning for medical triage[J]. Data Analysis and Knowledge Discovery, 2019, 3(6): 99-108.
- [21] 张强强, 苏变萍, 李敏. 基于卷积神经网络的SVM疾病症状分类模型[J]. 传感器与微系统, 2020, 39(3): 64-67.
ZHANG Q Q, SU B P, LI M. Disease symptom classification model of SVM based on CNN[J]. Transducer and Microsystem Technologies, 2020, 39(3): 64-67.
- [22] 陈德鑫, 占袁圆, 杨兵, 等. 基于CNN-BiLSTM模型的在线医疗实体抽取研究[J]. 图书情报工作, 2019, 63(12): 105-113.
CHEN D X, ZHAN Y Y, YANG B, et al. Research on extraction of online medical entities based on mixed deep learning model[J]. Library and Information Service, 2019, 63(12): 105-113.
- [23] LÜ X, YOU X. Relation extraction toward patent domain based on keyword strategy and attention+BiLSTM model (short paper)[C]//International Conference on Collaborative Computing: Networking, Applications and Worksharing. Springer, Cham, 2019: 408-416.
- [24] 田园, 马文. 基于Attention-BiLSTM的电网设备故障文本分类[J]. 计算机应用, 2020, 40(S2): 24-29.
TIAN Y, MA W. Attention-BiLSTM-based fault text classification for power grid equipment[J]. Journal of Computer Applications, 2020, 40(S2): 24-29.
- [25] ZHANG J, LI Y, TIAN J, et al. LSTM-CNN hybrid model for text classification[C]//2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC). IEEE, 2018: 1675-1680.
- [26] JANG B, KIM M, HARERIMANA G, et al. Bi-LSTM model to increase accuracy in text classification: combining Word2Vec CNN and attention mechanism[J]. Appl Sci, 2020, 10(17): 5841.

(编辑:薛泽玲)