



## 层次化双注意力神经网络模型的情感分析研究

曾碧卿, 韩旭丽, 王盛玉, 周武, 杨恒

引用本文:

曾碧卿, 韩旭丽, 王盛玉, 等. 层次化双注意力神经网络模型的情感分析研究[J]. 智能系统学报, 2020, 15(3): 460–467.

ZENG Biqing, HAN Xuli, WANG Shengyu, et al. Hierarchical double-attention neural networks for sentiment classification[J]. *CAAI Transactions on Intelligent Systems*, 2020, 15(3): 460–467.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.201812017>

## 您可能感兴趣的其他文章

### 深度学习的双人交互行为识别与预测算法研究

Human interaction recognition and prediction algorithm based on deep learning

智能系统学报. 2020, 15(3): 484–490 <https://dx.doi.org/10.11992/tis.201812029>

### 一种高效的稀疏卷积神经网络加速器的设计与实现

Design and implementation of an efficient accelerator for sparse convolutional neural network

智能系统学报. 2020, 15(2): 323–333 <https://dx.doi.org/10.11992/tis.201902007>

### 注意力机制和Faster RCNN相结合的绝缘子识别

Insulator recognition based on attention mechanism and Faster RCNN

智能系统学报. 2020, 15(1): 92–98 <https://dx.doi.org/10.11992/tis.201907023>

### 基于双向消息链路卷积网络的显著性物体检测

Salient object detection based on bidirectional message link convolution neural network

智能系统学报. 2019, 14(6): 1152–1162 <https://dx.doi.org/10.11992/tis.201812003>

### 基于改进卷积神经网络的多标记分类算法

A multi-label classification algorithm based on an improved convolutional neural network

智能系统学报. 2019, 14(3): 566–574 <https://dx.doi.org/10.11992/tis.201804056>

### 基于卷积特征和贝叶斯分类器的人脸识别

Face recognition based on convolution feature and Bayes classifier

智能系统学报. 2018, 13(5): 769–775 <https://dx.doi.org/10.11992/tis.201706052>

微信公众平台



关注微信公众号, 获取更多资讯信息

DOI: 10.11992/tis.201812017

# 层次化双注意力神经网络模型的情感分析研究

曾碧卿<sup>1</sup>, 韩旭丽<sup>2</sup>, 王盛玉<sup>2</sup>, 周武<sup>2</sup>, 杨恒<sup>2</sup>

(1. 华南师范大学 软件学院, 广东 佛山 528225; 2. 华南师范大学 计算机学院, 广东 广州 510631)

**摘 要:** 在篇章级的情感分类中由于篇章级文本较长, 特征提取较普通句子级分析相对较难, 大多方法使用层次化的模型进行篇章文本的情感分析, 但目前的层次化模型多以循环神经网络和注意力机制为主, 单一的循环神经网络结构提取的特征不够明显。本文针对篇章级的情感分类任务, 提出一种层次化双注意力神经网络模型。首先对卷积神经网络进行改进, 构建词注意力卷积神经网络。然后模型从两个层次依次提取篇章特征, 第一层次使注意力卷积神经网络发现每个句子中的重要词汇, 提取句子的词特征, 构建句子特征向量; 第二层次以循环神经网络获取整个篇章的语义表示, 全局注意力机制发现篇章中每个句子的重要性, 分配以不同的权重, 最后构建篇章的整体语义表示。在 IMDB、YELP 2013、YELP 2014 数据集上的实验表明, 模型较当前最好的模型更具优越性。

**关键词:** 情感分析; 注意力机制; 卷积神经网络; 情感分类; 循环神经网络; 词向量; 深度学习; 特征选取  
**中图分类号:** TP391   **文献标志码:** A   **文章编号:** 1673-4785(2020)03-0460-08

中文引用格式: 曾碧卿, 韩旭丽, 王盛玉, 等. 层次化双注意力神经网络模型的情感分析研究 [J]. 智能系统学报, 2020, 15(3): 460-467.

英文引用格式: ZENG Biqing, HAN Xuli, WANG Shengyu, et al. Hierarchical double-attention neural networks for sentiment classification[J]. CAAI transactions on intelligent systems, 2020, 15(3): 460-467.

## Hierarchical double-attention neural networks for sentiment classification

ZENG Biqing<sup>1</sup>, HAN Xuli<sup>2</sup>, WANG Shengyu<sup>2</sup>, ZHOU Wu<sup>2</sup>, YANG Heng<sup>2</sup>

(1. College of Software, South China Normal University, Foshan 528225, China; 2. College of Computer, South China Normal University, Guangzhou 510631, China)

**Abstract:** In sentiment classification, feature extraction in the document level is more difficult than the analysis in the common sentence level because of the length of the text. Most methods apply a hierarchical model to the sentiment analysis of text in the document level. However, most existing hierarchical methods mainly focus on a recurrent neural network (RNN) and attention mechanism, and the feature extracted by a single RNN is unclear. To solve the sentiment classification problem in the document level, we propose a hierarchical double-attention neural network model. In the first step, we improve a convolutional neural network (CNN), construct a word attention CNN, and then extract the features of the chapter from two levels. In the first level, the attention CNN can identify important words and phrases from every sentence, extract the word feature of the sentence, and construct the feature vector of the sentence. In the second level, the semantic meaning of the document is derived by the RNN. The global attention mechanism can find the importance of every sentence in the document, attribute different weights to them, and construct the whole semantic representation of the document. The experiment results on IMDB, YELP 2013, and YELP 2014 datasets show that our model achieves a more significant improvement than all state-of-the-art methods.

**Keywords:** sentiment analysis; attention mechanism; convolutional neural network (CNN); sentiment classification; recurrent neural network (RNN); word vector; deep learning; feature selection

收稿日期: 2018-12-15.

基金项目: 国家自然科学基金项目 (61772211, 61503143).

通信作者: 曾碧卿. E-mail: zengbiqing0528@163.com.

情感分析的目的是根据用户或者产品中的数据进行分析其极性特征, 得出用户的行为分析结果。目前, 情感分析已经得到工业界和学术界的

广泛关注。

篇章的情感分类是情感分析任务中的一项基础任务,假设每个篇章只有一个情感目标的情况下,分析篇章的情感极性得到情感倾向性。Pang等<sup>[1]</sup>采用了传统统计学的方法在情感分析上进行应用,对文本中词的性质特征进行有效的标注,该方法在情感分析的实验中取得良好的实验效果,由此证明了标记词性在情感分析上具有积极的意义。Lu等<sup>[2]</sup>利用词典 WordNet 来发现情感词汇。Wang等<sup>[3]</sup>将机器学习方法 SVM 与朴素贝叶斯方法进行结合,在情感分类任务中得到良好的实验效果。Kiritchenko等<sup>[4]</sup>使用 tweets 中的标签和表情来构建情感特征。

神经网络模型在自然语言处理的各个领域都取得较好效果,如命名实体识别<sup>[5-6]</sup>、机器阅读<sup>[7-8]</sup>和问答系统<sup>[9]</sup>。在篇章情感分析中, Kim<sup>[10]</sup>提出用词向量训练方法表示文本信息,在 CNN 模型中成功地进行文本的篇章分类。Kalchbrenner等<sup>[11]</sup>构建卷积神经网络模型,采用 MAX-Pooling 进行提取情感特征并取得良好的实验效果。Zhang等<sup>[12]</sup>使用字符作为最小单元,以卷积神经网络进行文本分类,在多个数据上取得较好效果。

层次化神经网络模型和注意力机制结合在情感分析中取得较好效果。Tang等<sup>[13]</sup>建立了一种层次化模型,使用 CNN 和 LSTM 从词和句子两个角度获取篇章的特征表示。Yang等<sup>[14]</sup>将注意力机制与层次化模型结合,利用注意力机制分别关注句子中的关键词和篇章中的关键句子。这些方法多以循环神经网络模型为基础,结合注意力构建层次化的模型,或者仅从词角度分析篇章。

在篇章级的文本情感分析中,理想的层次化情感分析模型应当在层次的第一层中将重点情感特征进行明显表示,以帮助模型挖掘到重要的情感特征。在第二层发现并整合篇章的重要句子,得到篇章整体语义。CNN 具有较强的局部特征提取能力,但 CNN 无法自动判别输入文本中哪些局部特征词较为重要,故本文首先结合注意力和卷积神经网络构建词注意力的卷积神经网络,发现重点词汇。因此,本文构建一个双注意力的层次化情感分析模型。采用双注意力机制,模型使用两个注意力机制,分别关注重要的词汇和句子。在提取句子中的关键词时,使用局部注意力机制,获取篇章关键句子时,使用全局注意力机制。CNN 有助于获取局部特征,局部注意力可使模型更有效地关注重点词汇,提取句子的局部特

征。第二层中,GRU 神经网络可提取到整个篇章的语义,全局注意力可关注到篇章中的重点句子,得到篇章的整体语义信息<sup>[15-16]</sup>。

本文主要贡献如下:

1) 利用注意力机制与卷积神经网络进行建模,构建层化的情感分类模型,改进的 CNN 可有效发现句子的重点词汇,形成句子表示;

2) 本文的模型在多个数据集上取得当前最好实验效果;

3) 本文从局部特征和全局语义两个角度提取特征,两个注意力机制分别关注不同的侧面,多元的特征更有助于篇章的表达。

## 1 相关工作

篇章级别的情感分类是情感分析的一项基础工作<sup>[3]</sup>, Pang等<sup>[1]</sup>将情感分类问题视为同篇章分类相似的任务,采用监督学习算法解决该任务。Wan<sup>[17]</sup>使用英文数据作为训练集,训练模型解决了中文情感分类问题,用于解决交叉语言情感分析的难题。Zagibalov等<sup>[18]</sup>提出一种能自动提取商品评论中情感词的方法,该方法是一种无监督的机器学习算法,无需任何的标注数据集。

但是,上述方法多依赖于特征工程,在构建特征工程过程中需要耗费大量时间和财力,并且不能很好地提取到评论的语法和语义信息。词嵌入<sup>[19-21]</sup>原理在于,利用一个计算好的词典向量表将文本映射成相应的实值向量。Tang等<sup>[22]</sup>使用一种改进的基于情感信息辅助的词嵌入方法提高情感分析的准确性。其他的传统神经网络模型在情感分析任务中应用得非常广泛,如 RNN 和 CNN 常被用来设计模型,并且承担了大量的情感分析任务。Kim<sup>[10]</sup>以 CNN 为基础构建了一个多通道 CNN 模型,以此提取多方面的情感特征。Johnson等<sup>[23]</sup>用单词的独热表示作为输入特征,使模型学习到更深层的语义特征,从而进行情感分析。Socher等<sup>[24]</sup>在 RNN 上进行构建情感分类模型并取得良好的实验效果。Tang等<sup>[13]</sup>提出将卷积神经网络与循环神经网络进行结合的基于层次化的篇章分类模型。该模型首先利用 CNN 或 RNN 学习句子的特征表示。然后,使用一种带有门控机制的 RNN 构建篇章的特征表示。

注意力机制是一种有效神经网络机制。Xu等<sup>[25]</sup>利用注意力机制进行图像分析,在图像分类上改善了分类的性能。注意力机制在其他领域也有良好的应用,例如在机器翻译领域获得了应用研究,其常被用来优化 encode-decode 模型以解决



机器翻译效果不佳等问题<sup>[26-27]</sup>。Zhou 等<sup>[28]</sup>结合 LSTM 和注意力机制用于情感分类。Allamanis 等<sup>[29]</sup>在文本摘要领域使用 attention 机制,都取得较好效果。Yin 等<sup>[30]</sup>将卷积神经网络与注意力机制融合,提出一种新的模型以用于情感分析。这些模型方法充分证明了注意力机制在卷积神经网络中的实验有效性。Wang 等<sup>[31]</sup>提出将卷积神经网络与多个注意力进行结合并在情感分类任务中取得了良好的实验效果。Chen 等<sup>[32]</sup>使用层次化的模型,借助注意力机制将用户和产品信息加入

到模型,取得较好的情感分类效果。

## 2 层次化的双注意力神经网络模型

如图 1 所示,该图是层次化的双注意力神经网络模型,模型进行篇章特征提取后,得到最终的情感特征,从而进行分析情感极性。首先对篇章中的每个句子通过词注意力卷积神经网络获取句子的词特征,进而构建句子的特征表示;然后使用全局注意力机制的 LSTM 提取句子向量的整体篇章表示。

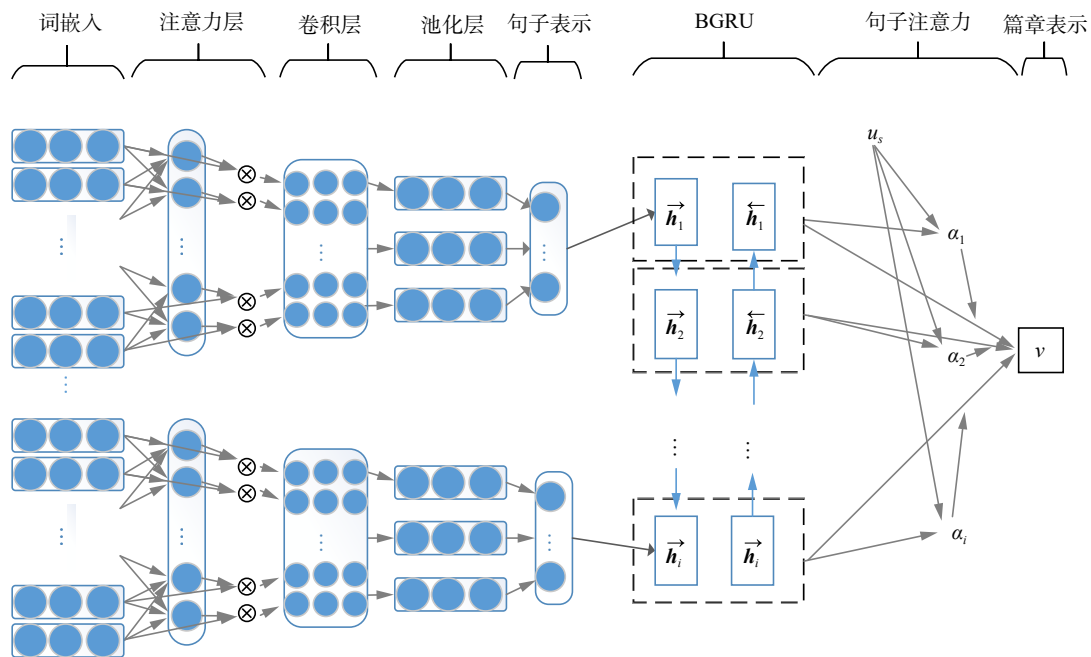


图 1 层次化双注意力神经网络模型

Fig. 1 Architecture of hierarchical double-attention networks

### 2.1 词特征提取

针对在情感分析应用模型中卷积神经网络模型存在的缺点,本文提出一种词特征提取的词注意力卷积神经网络模型,具体包括如下几部分:

1) 词嵌入层:词注意力神经网络模型的第一层为词嵌入层,主要的作用是将词进行序列化表示后得到文本的词嵌入表示。

2) 词注意力层:针对上一层的输出,对词嵌入层进行重点词提取,由此在词嵌入后进行添加词注意力机制,作用是减少在训练过程中噪声对实验的影响。

3) 卷积层:在词注意力层之后加入卷积层,本文使用的卷积核的窗口大小为 1、3、4 和 5。用于提取更多的情感特征。

在情感分析中,通过词注意力层和卷积层,使模型在训练过程中具有文本词特征的鉴别能力。其卷积层是为了选出对情感分析作用帮助更大的

特征,而注意力机制的作用是为了将重要的特征进行突出。

**词嵌入层:**文本进行词嵌入映射表示得到连续低维的实值向量,假设文本表示为  $S = \{w_1, w_2, \dots, w_{n-1}, w_n\}$ ,其中  $w_i$  是文本中第  $i$  个词语或者短语。本文利用 one-hot 算法对文本进行向量表示。本文设置的向量维度用  $d$  行表示:

$$x_i = w_i e, x_i \in \mathbf{R}^d \quad (1)$$

式中:  $e \in \mathbf{R}^{|\mathbf{V}| \times d}$  是词嵌入矩阵;  $|\mathbf{V}|$  表示整个数据集中词汇的数量;  $d$  代表词向量的维度。将文本进行词嵌入表示  $X = \{x_1, x_2, \dots, x_n\}$ 。

**词注意力层:**在词嵌入层表示后,让单词表示中的情感特征更为突出,然后本文借助于  $n$ -grams 的语言模型,并借助该中心词的上下文的  $n$  个词作为该中心词的语义表示。本文以大小  $L = [p_i - D, p_i + D] = 1 + 2D$  的词表示该中心词的语义表示。用  $\mathbf{W}_{att} \in \mathbf{R}^{L \times d}$  作为滑动窗口的矩阵参数。

其中,  $p_i$  表示该中心词,  $D$  表示选取的上下文的长度。如图2所示, 由此计算每个词表示  $x_i$  的特征值权重  $\alpha_i$ , 可以帮助文本中情感分析的重要程度值更加明显。

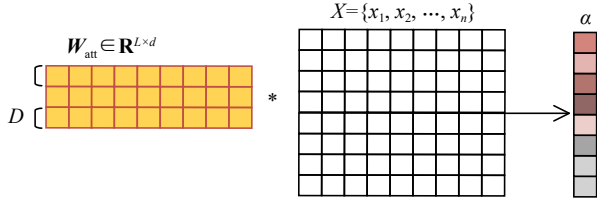


图2 词注意力层

Fig. 2 Local attention

$$X_i^{att} = \{x_{i-D}, \dots, x_i, \dots, x_{i+D}\} \quad (2)$$

$$\alpha_i = h(X_i^{att} W_{att} + b_{att}) \quad (3)$$

式中:  $X_i^{att}$  表示以  $x_i$  为中心, 将上下文的大小设置为  $L$ ;  $h(\cdot)$  代表 sigmoid 激活函数; 用  $b_{att}$  代表偏置项。如图2所示, 在词注意力层中运用滑动窗口机制, 进一步探索词嵌入  $X$  的深层特征值, 并为每个词的局部特征  $X$  分配权重  $\alpha$ :

$$\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_n\} \quad (4)$$

式中  $n$  表示文章中句子的长度。用  $\alpha$  表示句子中词或短语的重要程度值, 将词特征与权重相乘, 更新每个词的特征表示, 得到  $X_{att}$ :

$$\begin{aligned} x_i^{att} &= \alpha_i x_i \\ x_i^{att} &\in \mathbb{R}^d \\ X_{att} &= \{x_1^{att}, x_2^{att}, \dots, x_i^{att}, x_{i+1}^{att}, \dots, x_n^{att}\} \end{aligned} \quad (5)$$

**卷积层** 传统的情感分析处理自然语言的方法一般采用 1-gram、2-gram 和  $n$ -gram 语言模型提取文本的情感特征。Pang 等<sup>[1]</sup>曾使用该方法进行情感分析, 并取得良好的实验效果。卷积神经网络模型根据上下文大小提取局部特征值, 卷积核参数  $\varpi \in \mathbb{R}^{h \times d}$  都将以  $X_i$  为中心, 在每一次的卷积操作中, 提取卷积上下文窗口大小作为  $D$  的局部特征, 并将其视为  $n$ -grams 特征提取。

本文中, 取卷积核大小为 3, 则文本中心词对应的上下范围为 1, 并遵循马尔可夫原则, 即卷积核大小变为 5 时, 中心词对应上下文范围扩大为 2。当卷积核大小取 3 时, 其上下文大小为 1, 进行卷积操作时, 无法取到第一个词的上文以及最后一个词的下文表示。即首尾两个词导致有  $[h/2]$  个词无法提取到对应的  $n$ -grams 特征, 将产生文本信息提取不足的问题。为解决此问题, 本文采取首尾填充操作, 即在卷积层的输入末尾进行大小为  $h-1$  的 0 向量补全。

$$X = X_{att} \oplus \underbrace{\dots \oplus x_0 \oplus x_0}_{h-1} \quad (6)$$

式中  $\oplus$  表示串联操作。

卷积层的作用在于提取每个词特征信息和词的上下文信息。在词注意力之后得到的文本特征信息  $X_{att}$  作为卷积的输入。同时在卷积层中运用多个不同大小的卷积核进行提取文本特征, 每个卷积核取一定的数量。不同之处在于, 增加了大小为 1 的卷积核, 用于对每个词进行特征提取。将卷积层输出的特征图作为最大池化层的输入, 得到篇章的特征表示  $s_i$ :

$$s_i = \text{CNN}(X_{att}) \quad (7)$$

式中  $i$  代表篇章中的第几个句子。

## 2.2 篇章特征提取

**篇章分析** 假设输入句子层次词注意卷积神经网络获取到句子向量  $s_i$ , 将向量作为双向 GRU 神经网络的输入, 双向 GRU 的输出串联作为特征表示:

$$h_i = [\vec{h}_i, \overleftarrow{h}_i] = \text{Bi-GRU}(s_i) \quad (8)$$

**篇章全局注意力** 通过词注意力卷积神经网络获取到每个句子的特征表示, 经过双向 GRU 进一步编码提取每个句子的信息, 将双向 GRU 的输出作为全局注意力层的输入, 得到每个句子对篇章表示的贡献, 排除掉无用的句子并获取到篇章的整体语义信息。具体计算如下:

$$u_i = \tanh(W_g h_i + b_g) \quad (9)$$

$$\alpha_i = \frac{\exp(u_i)}{\sum_i \exp(u_i)} \quad (10)$$

$$v = \sum_i \alpha_i h_i \quad (11)$$

其中双向 GRU 的输出经过一个多层感知机, 使用  $\tanh$  非线性激活函数, 得到每个编码句子的权重值, 归一化 softmax 函数将权重分布在  $[0, 1]$  之间; 将句子编码为向量  $h_i$  并与对应权重  $\alpha_i$  相乘求和最终得到特征表示  $v$ 。

## 2.3 篇章分类

篇章特征向量由提取词和句子特征后得到。向量  $v$  首先通过多层感知机映射到情感类别空间  $C$ , 再使用 softmax 函数得到情感类别分布  $P_C$ 。

$$P_C = \text{softmax}(W_C v + b_C) \quad (12)$$

式中  $P_C$  代表篇章在类别  $C$  上的预测概率。

实验中将预测的类别分布同真实分布进行对比, 训练过程中使用交叉熵作为模型的损失函数。以  $P_C^g$  作为真实类别整体概率分布,  $P_C(d)$  是篇章的预测概率分布, 损失函数为

$$\text{loss} = - \sum_{d \in D} \sum_{c=1}^C P_C^g(d) \ln(P_C(d)) \quad (13)$$

式中  $D$  表示训练的样本集。

### 3 实验数据与实验设置

#### 3.1 数据集

为验证本文提出模型的有效性,实验部分在以下3个公开数据集上进行,所用数据集来自于Chen等<sup>[32]</sup>的数据集:

IMDB: 用户在不同商品上的评论数据集。用于对评论进行情感倾向性分析,每条评论分为10个情感强度。

YELP: Yelp 比赛 2013 年和 2014 年的数据集,每条评论分为5个强度。

数据集详细信息如表1所示。

表1 数据集  
Table 1 Datasets

数据集	类别	文档数	文档平均句子数	句子平均单词数
IMDB	10	84 919	16.08	24.54
YELP 2013	5	78 966	10.89	17.38
YELP 2014	5	231 163	11.41	17.26

#### 3.2 实验参数与数据处理

本数据集使用Keras进行数据预处理,分词任务由Keras提供的tokenizer实现。将数据集划分为训练、验证和测试集,划分比例分别为8:1:1。实验时对词向量进行降维处理,训练词向量维度为200维。卷积神经网络参数设置如表2所示。本文设置上下文范围大小 $D=2$ ,滑动窗口数量为1,GRU的输出维度设为100,双向得到向量为200维,全连接层隐藏单元200个。

表2 卷积层参数设置  
Table 2 Hyperparameter of CNN

参数	属性
激活函数	Recified linear Units
滑动窗口大小	1,3,4,5
滑动窗口数量	100
Dropout	0.5
Minibatch	64
池化	最大池化

实验模型指定batch\_size大小为32,同时对词向量进行动态微处理。本文使用Adadelata算法对模型参数进行了优化调整。为了提升实验效果和模型准确率,训练过程中将含有句子数相近的篇章形成一个batch,用以加速训练过程。最后,模型使用准确率作为评判标准。

#### 3.3 对比方法

为了测定模型效果,本实验对比了不同的模

型实验结果,如表3所示,有SVM(support vector machine)、NN(neural network)等。

表3 情感分析模型的实验结果  
Table 3 Results of the sentiment analysis of the models

模型	IMDB	YELP 2013	YELP 2014
Majority	19.6	41.1	39.2
Trigram	39.9	56.9	57.7
TextFeature	40.2	55.6	57.2
AvgWordvec	30.4	52.6	53
SSWE	31.2	54.9	55.7
Paragraph Vector	34.1	55.4	56.4
RNTN+RNN	40	57.4	58.2
NSC	44.3	62.7	63.7
NSC+LA	48.7	63.1	63
Cnn2Cnn	48.6	63.2	62.8
Cnn2Rnn	49.4	64.1	64.2

对比方法如下:

Majority: 将训练集中占多数的情感标签视为测试集的情感标签。

Trigram: 以一元词、二元词和三元词为特征训练SVM分类器<sup>[33]</sup>。

TextFeature: 提取文本的词特征、字符特征和词性特征训练SVM分类器。

AvgWordvec: 将文档的词向量平均化得到文档表示,以此训练SVM分类器。

SSWE: 使用特别训练的情感词向量生成特征,SVM为分类器。

NSC: 双层LSTM分别获取词特征和句子特征,构建篇章表示<sup>[32]</sup>。

NSC+LA: 双层LSTM结合全局注意力机制提取篇章的特征表示<sup>[14]</sup>。

RNTN+RNN: 使用递归神经网络提取句子的特征表示,并用循环神经网络得到篇章的特征表示。

Paragraph Vector: PVD<sup>[34]</sup>模型用于情感分类。

Cnn2Cnn: 词特征提取和句子特征提取都使用词注意卷积神经网络。

Cnn2Rnn: 词特征提取使用词注意卷积神经网络,句子特征提取使用GRU和全局注意力。

#### 3.4 实验结果与分析

本文的实验结果如表3所示,在3个数据集上分别得到模型的准确率,将模型的对比实验分为3组:第1组以传统特征提取结合分类器进行



情感分类;第2组使用层次化的循环神经网络作为基础,结合全局注意力机制,得到篇章的情感类别;第3组是本文提取出的模型,以词注意力卷积神经网络模型提取句子的词特征,构建句子表示,并以层次化的模型得到篇章情感倾向。

从第1组实验可以看出,传统方法和简单基于词向量的方法较其余两组的情感分类效果差距较大,说明在长文本的篇章中使用层次化的模型更有助于提取篇章的特征。

第2组的层次化循环神经网络,使用GRU或LSTM构建层次化的模型,模型在各个数据集上都有较大幅度的提升,NSC较Paragraph Vector提升都在8%左右。注意力机制有助于模型关注突出的情感特征,从而,提升模型效果,NSC+LA改进了NSC,取得了更优效果。

卷积神经网络和循环神经网络在处理文本时各有优缺点,单一的循环神经网络结合注意力机制,无法做到模型特征提取的多样性,从第3组实验可以看出,Cnn2Cnn模型与NSC+LA模型取得相近的结果,注意力机制虽不同,但CNN更关注于篇章的局部特征,对整体语义表征不够。

Cnn2Rnn取得了较之前更好的效果,在IMDB、YELP 2013和YELP 2014分别取得49.4%、64.1%和64.2%的准确率,较传统的第1组实验的方法提升为10%左右,较双层LSTM或GRU在IMDB、YELP 2013、YELP 2014提升5%、2%、1%左右,较单一模型的层次化注意力机制模型NSC+LA和Cnn2Cnn,各数据集提升在1%左右。

词注意力卷积神经网络作为第1层,提取每个句子中的关键词,发现句子的局部重要信息,如情感词、语气词、转折词等,以此构建句子的特征表示;全局注意力机制发现篇章中重要的句子,分配以高权重,循环神经网络模型提取整个篇章句子的语义表达,避免CNN的局部提取而丢失信息。

## 4 结束语

本文提出一种层次化的双注意力神经网络模型,将其用于情感分析任务,取得较好效果。模型基于CNN和RNN,从局部注意力机制和全局注意力两个角度提取不同的特征,层次化篇章的注意力神经网络模型分别获得句子的重点词汇以及整个篇章的重点句子特征。模型在多个数据集上得到验证,说明在不同角度获取的特征更有助于表达篇章特征。虽然本文在篇章级文本中取得

较好实验结果,由于模型较为复杂,并且文本较长,导致模型训练时间上并非理想。在未来研究中,优化模型的复杂性将是下一个改进的方向。

## 参考文献:

- [1] PANG Bo, LEE L, VAITHYANATHAN S. Thumbs up?: sentiment classification using machine learning techniques[C]//Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing-Volume 10. Stroudsburg, USA, 2002: 79–86.
- [2] LU Yue, CASTELLANOS M, DAYAL U, et al. Automatic construction of a context-aware sentiment lexicon: an optimization approach[C]//Proceedings of the 20th International Conference on World Wide Web. Hyderabad, India, 2011: 347–356.
- [3] WANG Sida, MANNING C D. Baselines and bigrams: simple, good sentiment and topic classification[C]//Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers. Jeju Island, Korea, 2012: 90–94.
- [4] KIRITCHENKO S, ZHU Xiaodan, MOHAMMAD S M. Sentiment analysis of short informal texts[J]. *Journal of artificial intelligence research*, 2014, 50: 723–762.
- [5] LAMPLE G, BALLESTEROS M, SUBRAMANIAN S, et al. Neural architectures for named entity recognition[C]//Proceedings of 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. San Diego, USA, 2016: 260–270.
- [6] SHEN Dinghan, MIN M R, LI Yitong, et al. Adaptive convolutional filter generation for natural language understanding. [J]. arXiv: 1709.08294, 2017.
- [7] WANG Shuohang, JIANG Jing. Machine comprehension using match-LSTM and answer pointer[C]//Proceedings of International Conference on Learning Representations. Toulon, France, 2017: 1–15.
- [8] WANG Wenhui, YANG Nan, WEI Furu, et al. Gated self-matching networks for reading comprehension and question answering[C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. Vancouver, Canada, 2017: 189–198.
- [9] KUMAR A, IRSOY O, ONDRUSKA P, et al. Ask me anything: dynamic memory networks for natural language processing[C]//Proceedings of the 33rd International Conference on Machine Learning. New York, USA, 2016: 1378–1387.
- [10] KIM Y. Convolutional neural networks for sentence clas-

- sification[C]//Proceedings of 2014 Conference on Empirical Methods in Natural Language Processing. Doha, Qatar, 2014: 1746–1751.
- [11] KALCHBRENNER N, GREFFENSTETTE E, BLUNSOM P, et al. A convolutional neural network for modeling sentences[C]//Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics. Baltimore, USA, 2014: 655–665.
- [12] ZHANG Xiang, ZHAO Junbo, LECUN Y. Character-level convolutional networks for text classification[C]//Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal, Canada, 2015: 649–657.
- [13] TANG Duyu, QIN Bing, LIU Ting. Document modeling with gated recurrent neural network for sentiment classification[C]//Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Lisbon, Portugal, 2015: 1422–1432.
- [14] YANG Zichao, YANG Diyi, DYER C, et al. Hierarchical attention networks for document classification[C]//Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. San Diego, USA, 2016: 1480–1489.
- [15] ADI Y, KERMANY E, BELINKOV Y, et al. Fine-grained analysis of sentence embeddings using auxiliary prediction tasks[C]//Proceedings of International Conference on Learning Representations. Toulon, France, 2017: 1608–1622.
- [16] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of the 31st Conference on Neural Information Processing Systems. Long Beach, USA, 2017: 5998–6008.
- [17] WAN Xiaojun. Co-training for cross-lingual sentiment classification[C]//Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP. Suntec, Singapore, 2009: 235–243.
- [18] ZAGIBALOV T, CARROLL J. Automatic seed word selection for unsupervised sentiment classification of Chinese text[C]//Proceedings of the 22nd International Conference on Computational Linguistics. Manchester, United Kingdom, 2008: 1073–1080.
- [19] LIU Jiangming, ZHANG Yue. Attention modeling for targeted sentiment[C]//Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics. Valencia, Spain, 2017: 572–577.
- [20] ZHANG Lei, WANG Shuai, LIU Bing. Deep learning for sentiment analysis: a survey[J]. WIREs data mining and knowledge discovery, 2018, 8(4): e1253.
- [21] CHEN Peng, SUN Zhongqian, BING Lidong, et al. Recurrent attention network on memory for aspect sentiment analysis[C]//Proceedings of 2017 Conference on Empirical Methods in Natural Language Processing. Copenhagen, Denmark, 2017: 452–461.
- [22] TANG Duyu, WEI Furu, YANG Nan, et al. Learning sentiment-specific word embedding for twitter sentiment classification[C]//Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics. Baltimore, USA, 2014: 1555–1565.
- [23] JOHNSON R, ZHANG Tong. Effective use of word order for text categorization with convolutional neural networks[C]//Proceedings of 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Denver, USA, 2015: 103–112.
- [24] SOCHER R, PERELYGIN A, WU J, et al. Recursive deep models for semantic compositionality over a sentiment Treebank[C]//Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing. Seattle, USA, 2013: 1631–1642.
- [25] XU K, BA J, KIROUS R, et al. Show, attend and tell: neural image caption generation with visual attention[C]//Proceedings of the 32nd International Conference on Machine Learning. Lille, France, 2015: 2048–2057.
- [26] BAHDANAU D, CHO K, BENGIO Y. Neural machine translation by jointly learning to align and translate[C]//Proceedings of the 3rd International Conference on Learning Representations, 2014. San Diego, USA, 2015: 473–488.
- [27] LUONG T, PHAM H, MANNING C D, et al. Effective approaches to attention-based neural machine translation[C]//Proceedings of 2015 Conference on Empirical Methods in Natural Language Processing. Lisbon, Portugal, 2015: 1412–1421.
- [28] ZHOU Xinjie, WAN Xiaojun, XIAO Jianguo. Attention-based LSTM network for cross-lingual sentiment classification[C]//Proceedings of 2016 Conference on Empirical Methods in Natural Language Processing. Austin, USA, 2016: 247–256.
- [29] ALLAMANIS M, PENG Hao, SUTTON C A. A convolutional attention network for extreme summarization of source code[C]//Proceedings of 2016 International Conference on Machine Learning. New York, USA, 2016: 2091–2100.
- [30] YIN Wenpeng, SCHÜTZE H, XIANG Bing, et al. AB-



CNN: attention-based convolutional neural network for modeling sentence pairs[J]. *Transactions of the association for computational linguistics*, 2016, 4: 259–272.

[31] WANG Linlin, CAO Zhu, DE MELO G, et al. Relation classification via multi-level attention CNNs[C]//Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. Berlin, Germany, 2016: 1298–1307.

[32] CHEN Huimin, SUN Maosong, TU Cunchao, et al. Neural sentiment classification with user and product attention [C]//Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. Austin, USA, 2016: 1650–1659.

[33] FAN Rongen, CHANG Kaiwei, HSIEH C J, et al. LIBLINEAR: a library for large linear classification[J]. *Journal of machine learning research*, 2008, 9: 1871–1874.

[34] LE Q, MIKLOV T. Distributed representations of sentences and documents[C]//Proceedings of the 31st International Conference on Machine Learning. Beijing, China, 2014: 1188–1196.

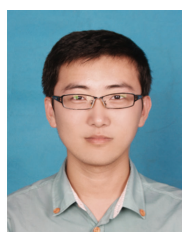
## 作者简介:



曾碧卿, 教授, 博士, 主要研究方向为认知计算和自然语言处理。获发明专利 6 项, 发表学术论文 100 余篇, 出版学术专著 2 部。



韩旭丽, 硕士研究生, 主要研究方向为自然语言处理、情感分析。发表学术论文 10 篇。



王盛玉, 硕士研究生, 主要研究方向为自然语言处理、情感分析。发表学术论文 6 篇。

## 2020 年度中国人工智能学会会士增选公示

根据《中国人工智能学会章程》、《中国人工智能学会会士产生与评定工作办法》，学会于 2020 年 6 月启动中国人工智能学会会士候选人提名工作。经学会会士提名、会士评定专家委员会办公室审查、会士评定专家委员会审议投票等程序，评选出王耀南、王恩东等 9 位会士人选。现将名单予以公示，公示期共 14 日（自 2020 年 8 月 25 日至 9 月 7 日）。

任何单位和个人如对公示的评审结果持有异议，请在 9 月 7 日前实名向中国人工智能学会会士评定专家委员会办公室提出，凡匿名异议不予受理。

地址：北京市海淀区西土城路 10 号北京邮电大学教一楼 122 室

邮编：100876

电话：010-62281360 13240274156

邮箱：zhhb@caai.cn

2020 年度中国人工智能学会会士增选名单

序号	姓名	工作单位
1	王耀南	湖南大学
2	王恩东	浪潮集团有限公司
3	陈纯	浙江大学
4	何积丰	华东师范大学
5	吾守尔·斯拉木	新疆大学
6	苗夺谦	同济大学
7	张亚勤	清华大学
8	张学工	清华大学
9	于剑	北京交通大学

中国人工智能学会  
2020 年 8 月 25 日