

融合双重权重机制和图卷积神经网络的微博细粒度情感分类



李浩 张兰 杨兵 杨海潇 寇勇奇 王飞 康雁

云南大学软件学院 昆明 650504

(lihao707@ynu.edu.cn)

摘要 利用深度学习模型和注意力机制对微博文本进行细粒度情感分类,已成为研究的热点,但是现有注意力机制只考虑单词对单词的影响,对单词本身的多种维度特性(如词义、词性、语义等特征信息)缺乏有效的融合。为了解决这个问题,文中提出了一种双重权重机制 WDWM(Word and Dimension Weight Mechanism),并将其与基于解析依赖树的 GCN 模型相结合,通过选择每条微博中含有关键信息的单词,来抽取单词的重要维度特性,对单词的多种维度特性进行有效融合,从而捕获更加丰富的特征信息。在针对微博细粒度情感分类的实验中,融合双重权重机制和图卷积神经网络的微博细粒度情感分类模型(WDWM-GCN)的 F 测度达到了 84.02%,比 2020 年提出的最新的算法高出 1.7%,这进一步证明,WDWM-GCN 能对单词的多维度特性进行有效的融合,能够捕获丰富的特征信息。在对搜狗新闻数据集进行分类的实验中,BERT 模型在加入 WDWM 后,其分类效果得到了进一步提升,这充分证明 WDWM 对所提分类模型有明显的改进效果。

关键词: 细粒度情感分类;双重权重机制;注意力机制;图卷积神经网络

中图法分类号 TP391

Fine-grained Sentiment Classification of Chinese Microblogs Combining Dual Weight Mechanism and Graph Convolutional Neural Network

LI Hao,ZHANG Lan,YANG Bing,YANG Hai-xiao,KOU Yong-qi,WANG Fei and KANG Yan

School of Software,Yunnan University,Kunming 650504,China

Abstract Using deep learning models and attention mechanisms to classify fine-grained emotions of Chinese microblogs has become a research hotspot. However, the existing attention mechanisms consider the impact of words on words, and lack effective integration of the various dimensional characteristics of the words themselves (such as word meaning, part of speech, semantics and other characteristic information). In order to solve this problem, the paper proposes a dual weight mechanism WDWM (word and dimension weight mechanism), and combines it with the GCN model based on the analytical dependency tree, so that it can not only select the words that contain key information in each microblog, but also extract the important dimensional characteristics of the word and effectively integrate multiple dimensional characteristics of words, so as to capture more rich feature information. The F measure of fine-grained sentiment classification of Chinese microblogs combining dual weight mechanism and graph convolutional neural network(WDWM-GCN) reaches 84.02%, which is 1.7% higher than the latest algorithm proposed by WWW in 2020, which further proves that WDWM-GCN can effectively integrate the multi-dimensional characteristics of words and capture rich feature information. In the experiment on the classification of Sogou news data set, after the BERT model is added to the WDWM mechanism, the classification effect is further improved, which fully proves that the WDWM has a significant improvement on the text classification model.

Keywords Fine-grained emotion classification, Dual weight mechanism, Attention mechanism, Graph convolutional neural network

到稿日期:2020-12-07 返修日期:2021-06-08 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:国家自然科学基金(61762092);云南省软件工程重点实验室开放基金项目(2020SE303);云南省重大科技专项(202002AB080001);《材料基因工程-基于 Metcloud 的集成计算功能模块计算软件开发》(2019CLJY06);云南省稀贵金属材料基因工程-稀贵金属材料高通量集成计算和数据分析技术研发及示范应用(2019ZE001-1,202002AB080001)

This work was supported by the National Natural Science Foundation of China(61762092),Open Fund Project of Key Laboratory of Software Engineering in Yunnan Province(2020SE303),Major Science and Technology Projects in Yunnan Province(202002AB080001),Material Genetic Engineering-Calculation Software Development of Integrated Calculation Function Module Based on Metcloud(2019CLJY06) and Gene Engineering of Rare and Precious Metal Materials in Yunnan Province-R & D and Demonstration Application of High-throughput Integrated Computing and Data Analysis Technology for Rare and Precious Metal Materials(2019ZE001-1,202002AB080001).

通信作者:康雁(562530855@qq.com)

1 引言

随着 NLP 的快速发展,在信息检索、机器翻译、文本分类、问答系统、语音识别等多个方面的研究已取得突破性的进展。其中,对给定句子或篇章的情感倾向进行判别的文本情感分类得到了大量应用。各大购物平台和社交网站每天都会生成大量的网络数据,而这些文本数据存在巨大的研究价值和商业价值,例如:亚马逊通过对购物平台的用户评价进行情感分析,来提升商品的分析决策^[1];京东根据用户的评价反馈信息优化经营战略;政府通过对微博评论进行情感分析,来掌握网络舆论、监控网络舆情等^[2]。传统的文本情感分类本质上是指定一个整体的情感极性,如将微博文本表达的情感分为两类:消极和积极。而细粒度的文本情感分类旨在分辨出确切的情感类型,如将微博文本表达的情感细分为 7 类:幸福、悲伤、喜欢、愤怒、讨厌、恐惧和惊奇。对文本的情感进行细分,更有利于描述观点持有者复杂的情绪状态。

one-hot, LDA, Word2vec 等词向量表示方法的提出,为深度学习模型处理自然语言奠定了基础。众多研究者开始利用深度学习进行文本情感分类研究, CNN 的多层神经网络被用于提取深层次的特征^[3-6], RNN 也被广泛应用于情感分类,它能获取上下文语义信息,从而提高情感分类的准确度^[7-9]。LSTM 是对 RNN 模型的进一步优化,其能有效挖掘短文本包含的隐含特征,从而达到更精准地识别文本的情感极性的效果^[10-12]。

近年来,注意力机制在自然语言处理领域取得了巨大成功,在机器翻译、情绪分析、词性标记等多个自然语言处理任务中取得了显著成效^[13-15]。利用注意力机制进行文本情感分类成为了研究热点,部分工作者将注意力机制用于情感分类^[16-17],以捕获上下文单词对特征向量的影响,还有部分工作者将注意力机制与神经网络模型相结合^[18-19],以抽取抽象的深层次特征,但是上述方法并未关注词向量内部的维度特性,不能有效地融合单词的多种维度特性。

本文提出了一个新的模型(WDWM-GCN), WDWM 能选择含有关键信息的单词,并对词向量的每一个维度分配权重,关注词向量的重要维度特性,无论是深度学习模型还是注意力模型,在处理文本信息时往往会忽略句子的语法信息。基于解析依赖树的 GCN 模型将 GCN 与解析依赖树相结合,关注句子的语法信息。最后,有效地将基于解析依赖树的 GCN 模型与 WDWM 相结合,从而提高情感分类模型的准确度。

总而言之,本文的研究贡献如下:

(1) 本文提出了一种基于单词和维度双重权重机制 WDWM,通过选择核心单词并抽取单词的重要维度特性,来获取更加丰富的特征信息。

(2) 将 WDWM 与基于解析依赖树的 GCN 模型相结合,使其不仅能够关注句子的语法信息,还能够对单词的多种维度特性进行有效的融合。在 15664 条中国微博数据集上分别对情感二元模型和情感七元模型进行实验,结果表明了该方法的有效性。

(3) 在 19 万条搜狗新闻数据集上分别对多个不同的文本分类模型进行实验,将 WDWM 与多个不同的文本分类模型相结合,结果表明,WDWM 对不同的文本分类模型都有显著的改进效果。

2 相关工作

2.1 传统模型

传统基于情感词典的方法主要是基于情感知识构建情感词典,并将其作为工具来判断文本情感极性。Wang 等^[20]利用 40 万条微博数据构建新词词典,对已有情感资源进行拓展。Vilares 等^[21]引入了带有情感标签的 Twitter 语料库来解决 Twitter 上多语言极性分类问题。

传统的基于机器学习的方法通过人工方式来提取特征,然后利用朴素贝叶斯(Naive Bayes)、支持向量机、KNN 等算法进行情感分类^[22-24],显著提高了文本情感分类的效果。然而,传统的模型不仅耗费大量的人力,人工提取的特征覆盖面有限,而且任务完成效率和任务完成质量较低,因此传统方法不具有普适性。深度学习可以通过构建网络模型来对文本进行逐步分析、特征抽取并且自动学习、优化模型输出,以提高文本分类的正确性。

2.2 深度学习模型

近年来,深度学习在海量文本数据的智能理解上表现出了独特的优势,得到了越来越多研究者的青睐。同样,在文本情感分类任务中取得了显著成效。

CNN 可以捕捉到文本中的局部信息,提取复杂的文本情感特征。文献[4]将卷积神经网络与门控机制相结合,使模型更加准确和高效。Chen 等^[25]提出把单词的词向量、词性向量和位置向量信息结合,形成多个不同的通道作为 CNN 的输入,以提升分类的性能。文献[26]提出将 CNN 添加到 BERT 模型的特定任务的层中,可以获取文本中重要片段的信息。还有部分工作者提出使用循环神经网络(RNN)进行建模,文献[8]使用门控递归网络对 24 种细粒度的数据进行分类,并达到了最新的技术水平。但 RNN 存在梯度消失或梯度爆炸的现象,针对这一问题,文献[27]提出了一种基于 LSTM 的双向深度文本表示模型,从文本的两个方向来抓取词间关系,充分融合了文本的上下文信息,得到了更丰富的文本特征表示,从而提高了情感分类的准确率。还有部分工作者将两个或多个模型相结合,文献[28]提出了基于 3 层 LSTM 和 CNN 网络结构的 ResLCNN (Residual-LSTM-CNN)深度学习模型。ResLCNN 深度学习模型通过 LSTM 获取文本序列数据的长距离依赖特征, CNN 的卷积操作获取句子的局部特征,有效结合两者的优势,同时构建残差层,缓解深层模型梯度消失问题。

2.3 Attention 机制

Attention 机制最初被用于机器翻译^[13],并取得了巨大的成功,后来被广泛用于自然语言处理的各个任务。Yang 等^[29]提出了 Hierarchical Attention,在单词和句子级别分别设计了两个不同级别的注意力机制,但其将单词和句子分开考虑,忽略了语境信息。Luong 等^[30]提出了 Global Attention

和 Local Attention, Global Attention 在计算语义向量时会考虑编码器所有的隐藏状态。在 Local Attention 中,计算语义向量时只关注每个目标词的一部分编码器隐藏状态,无法捕获单词间长距离的依赖关系。文献[31]提出了 Self-Attention 和 Multi-head Attention,自注意力通过引入自我关注来捕捉数据或特征的内部相关性,多头注意力允许模型共同关注来自不同子空间的信息,以更好地学习特征,但缺乏足够的位置信息,容易出现词序错误的现象。注意力模型在文本情感分类任务中也取得了显著成效,文献[32]设计了一种新颖的注意权重生成方法,该方法将方面词与其上下文语义信息相结合,从而可以更好地利用方面词与其上下文文词之间的相关性。文献[33]提出了 BiGRU-Attention 与门控机制(Gated Mechanisms)相结合的文本情感分类模型,使用 BiGRU 对文本进行深层次的特征提取,然后,结合门控机制和注意力机制,进一步提取与方面词相关的上下文情感特征信息。

上述注意力都是通过关注上下文单词对词向量表达产生的影响,来优化分类器的分类效果,但是每个词向量包含多个维度特性,如词义、词性、语义等。目前存在的注意力机制对每个单词本身的多种维度特性缺乏有效的融合,对不同维度特性的重要程度也缺乏考虑。因此,本文提出了一种双重权重机制 WDWM,通过对每个单词分配不同的权重,来选择关键单词,以获取核心信息;通过对单词的每一维度分配权重,抽取词向量的重要维度特性,有效融合多种维度特性,来捕获更加丰富的特征信息,提高微博情感分类的效果。

2.4 GCN 模型

GCN 最早由 Kip 等^[34]提出,通过谱图卷积(Spectral Graph Convolutions,SGC)的局部一阶近似,来确定卷积网络的结构。目前 GCN 被广泛应用于人工智能的 3 大领域:计算机视觉、自然语言处理和推荐系统。而文本分类是自然语言处理的一个重要任务,常见的文本分类主要基于传统的 CNN,LSTM 及 transform,bert 等方法,用于处理排列整齐的矩阵特征。但对于社交网络、信息网络等一系列很多不规则的以图的形式存在的数据,传统的模型无法处理该类数据,提取特征进行进一步的学习,因此考虑将 GCN 应用于文本分类任务。目前 GCN 在文本分类任务的应用上已经取得了一些研究成果,文献[35]将 GCN 与依赖解析树相结合,通过融入句子的语法结构,使分类器的性能得到显著提高,但是该模型忽略了单词的多维度特性,因此本文将基于解析依赖树的 GCN 模型生成的特征向量通过 WDWM,使得 GCN 能够关注句子的语法信息,抽取单词的重要维度特性,对词向量的多维度特性进行有效融合,提高分类的准确率。

3 WDWM-GCN 模型

3.1 WDWM-GCN 模型的概述

WDWM-GCN 模型将双重权重机制 WDWM 与基于解析依赖树的 GCN 模型相结合来预测给定句子的情感极性,图 1 给出了所提模型的整体架构。

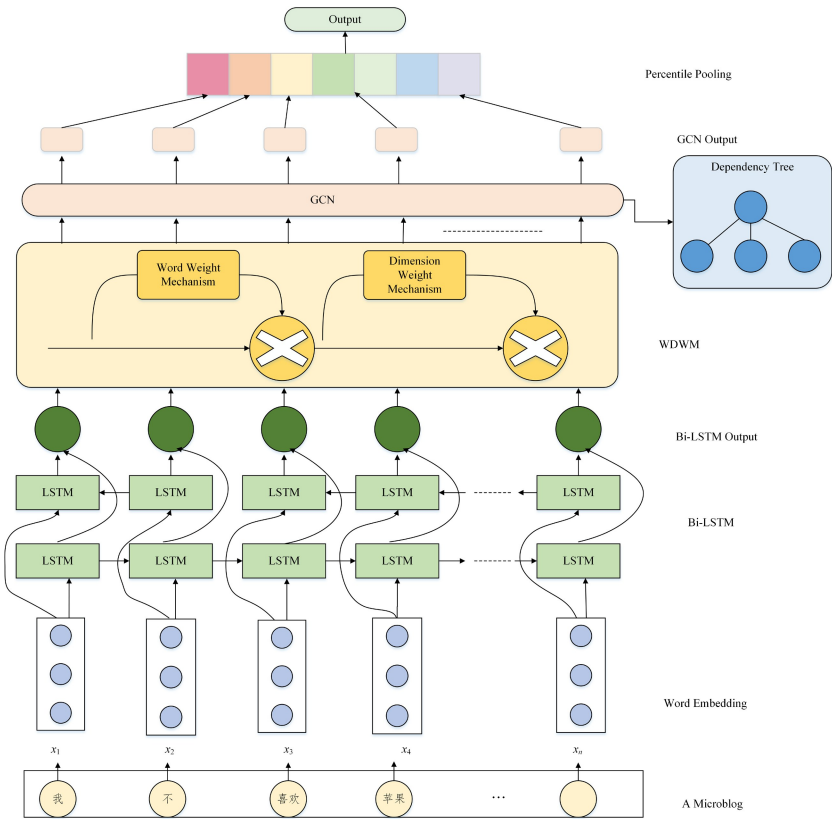


图 1 模型结构

Fig. 1 Architecture of the model

对于组成一个句子的每个单词,通过以下步骤进行情感分类:1)对输入的微博文本进行预处理,将文本词汇向量化,

生成每个单词的 Word Embedding;2)生成的词向量输入 Bi-LSTM,执行 Bi-LSTM,提取微博文本特征,生成特征矩阵;

3)执行 Word Weight Mechanism,通过 Maxpooling 和 Avgpooling 抽取特征并分别通过 MLP 整合特征信息,最终将其合并并通过 Softmax 函数生成每个单词对应的权重,从而得到权重向量,将步骤 2)中生成的特征矩阵与权重向量相乘,关注每个单词对微博文本情感分类任务的影响,得到新的特征矩阵并将其作为下一层的输入;4)执行 Dimension Weight Mechanism,通过 Maxpooling 和 Avgpooling 分别抽取词向量的重要维度特性,进行横向拼接,融合多种特征信息进行全连接,抽取深层次特征,最终通过 Softmax 函数生成词向量的每个维度对应的权重,将步骤 3)中输出的特征矩阵与权重矩阵相乘,抽取单词的重要维度特性,对单词的多种维度特性进行有效融合,从而得到新的特征矩阵并将其作为下一层的输入;5)在句子的依存关系树上执行 GCN,引入语法信息,对步骤 4)中输出的特征矩阵进行进一步的特征抽取,从而得到新的特征矩阵并将其作为下一层的输入;6)通过 Percentile Pooling 函数进行分类,最终分为 7 类:幸福、悲伤、喜欢、愤怒、讨厌、恐惧和惊奇。

由于微博评论没有固定的语法和模式,因此需要对评论文本去噪并使用 jieba 分词进行词语分割。对于词嵌入,采用 Word2vec 对词向量进行处理,并在网络训练期间对向量进行更新。在经过预处理后,将每个微博文本设置为 $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$,并将其作为 Bi-LSTM 模型的输入。其中 $\mathbf{x}_i \in \mathbb{R}^{300}$ ($i=1,2,\dots,n$), $\mathbf{x}_i$ 为第 i 个单词的向量表示, n 为评论文本的词汇数量。

3.2 Bi-LSTM 模型

在输入表示微博文本的矩阵 $\mathbf{X}$ 后,利用 Bi-LSTM 模型提取文本特征。LSTM 模型主要包括输入门 i_t 、遗忘门 f_t 、输出门 o_t 和记忆单元 C_t 等,其计算过程如下:

$$f_t = \sigma(\mathbf{W}_f \cdot [\mathbf{m}_{t-1}, \mathbf{x}_t] + \mathbf{b}_f) \quad (1)$$

$$i_t = \sigma(\mathbf{W}_i \cdot [\mathbf{m}_{t-1}, \mathbf{x}_t] + \mathbf{b}_i) \quad (2)$$

$$C_t = f_t \times C_{t-1} + i_t \times \tanh(\mathbf{W}_c \cdot [\mathbf{m}_{t-1}, \mathbf{x}_t] + \mathbf{b}_c) \quad (3)$$

$$o_t = \sigma(\mathbf{W}_o \cdot [\mathbf{m}_{t-1}, \mathbf{x}_t] + \mathbf{b}_o) \quad (4)$$

$$\mathbf{m}_t = o_t \cdot \tanh(C_t) \quad (5)$$

其中, $\mathbf{W}_i, \mathbf{W}_f, \mathbf{W}_o$ 分别为输入门、遗忘门、输出门的权重矩阵; $\mathbf{b}_i, \mathbf{b}_f, \mathbf{b}_o$ 分别为输入门、遗忘门、输出门的偏置矩阵; $\sigma, \tanh$ 为激活函数, $\mathbf{m}_t, \mathbf{m}_{t-1}$ 表示 t 时刻和 $(t-1)$ 时刻隐藏层的状态。

Bi-LSTM 模型每一时刻的状态由正向和反向两个方向的 LSTM 共同决定。因此, Bi-LSTM 模型会同时获得正向序列 $\mathbf{M}_1 = \{\mathbf{m}_{11}, \mathbf{m}_{12}, \dots, \mathbf{m}_{1n}\}$ 和反向序列 $\mathbf{M}_2 = \{\mathbf{m}_{21}, \mathbf{m}_{22}, \dots, \mathbf{m}_{2n}\}$ 。将这两个序列进行拼接,得到 $\mathbf{M} = \begin{bmatrix} \mathbf{M}_1 \\ \mathbf{M}_2 \end{bmatrix} = \{\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_n\}$, 其中 $\mathbf{m}_i = \begin{bmatrix} \mathbf{m}_{1i} \\ \mathbf{m}_{2i} \end{bmatrix}$, $i=1,2,\dots,n$ 表示每个单词的特征向量。将 $\mathbf{m}_i$ 作为 WDWM 的输入,执行 3.3 节中的 WDWM。

3.3 WDWM

对于 3.2 节中生成的特征向量 $\mathbf{M}$, WDWM 从横向(Word Weight Mechanism)和纵向(Dimension Weight Mechanism)两个方向的维度计算权重,生成权重向量 $\mathbf{P}_w$ 和 $\mathbf{P}_d$ 。然后将 $\mathbf{P}_w$ 和 $\mathbf{P}_d$ 分别与输入的特征矩阵相乘进行特征的自适应学习。

$$\mathbf{M}' = \mathbf{P}_w(\mathbf{M}) \times \mathbf{M} \quad (6)$$

$$\mathbf{M}'' = \mathbf{P}_d(\mathbf{M}') \times \mathbf{M}' \quad (7)$$

其中, $\mathbf{M} \in \mathbb{R}^{360}$, $\mathbf{P}_w \in \mathbb{R}^{360}$ 和 $\mathbf{P}_d \in \mathbb{R}^{140}$ 分别是 Word Weight Mechanism 和 Dimension Weight Mechanism 生成的权重向量。接下来,本文将分别对 Word Weight Mechanism 和 Dimension Weight Mechanism 两种不同的权重机制进行描述。

(1) Word Weight Mechanism

在生成单词的特征向量时,并非所有的单词对文本的语义表示作出同等贡献。因此,本文引入 Word Weight Mechanism 对每个单词分配不同的权重,抽取核心单词的特征信息。它的结构如图 2 所示。

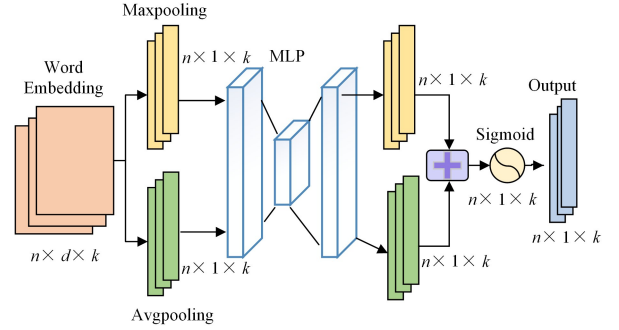


图 2 单词权重机制

Fig. 2 Word weight mechanism

Word Weight Mechanism 首先通过 Bi-LSTM 模型输出的特征矩阵 $\mathbf{M}$ 分别进行一维最大池化和一维平均池化,从而得到两个不同的特征描述,将 $\mathbf{M}_{\max}$ 和 $\mathbf{M}_{\text{avg}}$ 输入 MLP,分别进行非线性变化,得到 $\mathbf{y}_{\max}$ 和 $\mathbf{y}_{\text{avg}}$,再将 $\mathbf{y}_{\max}$ 和 $\mathbf{y}_{\text{avg}}$ 两个一维矩阵进行矩阵相加,经过 Sigmoid 函数得到 $\mathbf{P}_w$ 。具体计算过程如下:

$$\mathbf{M}_{\max} = \text{MaxPool}\{\mathbf{M}\} \quad (8)$$

$$\mathbf{M}_{\text{avg}} = \text{AvgPool}\{\mathbf{M}\} \quad (9)$$

$$\mathbf{y}_{\max} = \mathbf{W}_1(\text{relu}(\mathbf{W}_0 \mathbf{M}_{\max} + \mathbf{b}_0)) + \mathbf{b}_1 \quad (10)$$

$$\mathbf{y}_{\text{avg}} = \mathbf{W}_1(\text{relu}(\mathbf{W}_0 \mathbf{M}_{\text{avg}} + \mathbf{b}_0)) + \mathbf{b}_1 \quad (11)$$

$$\mathbf{P}_w = \text{sigmoid}(\mathbf{y}_{\max} + \mathbf{y}_{\text{avg}}) \quad (12)$$

(2) Dimension Weight Mechanism

Dimension Weight Mechanism 抽取单词向量的重要维度,并通过 Sigmoid 函数生成词向量的每个维度对应的权重,对单词的多维度特性进行有效融合。维度权重机制的模型结构如图 3 所示。对输入的特征向量矩阵 $\mathbf{M}$ 的转置分别进行一维最大池化和一维平均池化,抽取特征,得到 $\mathbf{M}_{\max}$ 和 $\mathbf{M}_{\text{avg}}$ 。 $\mathbf{M}_{\max}$ 和 $\mathbf{M}_{\text{avg}}$ 通过横向拼接 Concatenate 函数,将抽取到的两种特征合并并进行全连接,通过激活函数 σ 得到输出 $\mathbf{P}_d$, σ 指 Sigmoid 函数,具体公式如下:

$$\mathbf{M}_{\max} = \text{MaxPool}\{\mathbf{M}\} \quad (13)$$

$$\mathbf{M}_{\text{avg}} = \text{AvgPool}\{\mathbf{M}\} \quad (14)$$

$$\mathbf{P}_d = \sigma(\text{relu}(\mathbf{W}_{(2)}(\text{Concatenate}(\mathbf{M}_{\max}, \mathbf{M}_{\text{avg}}) + \mathbf{b}_{(2)}))) \quad (15)$$

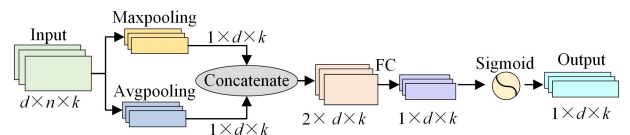


图 3 维度权重机制

Fig. 3 Dimension weight mechanism

3.4 基于解析依赖树的 GCN 模型

利用文本的解析依赖树来构建 GCN,使得 GCN 能够关注句子的语法信息。其中,解析依赖树能识别主语-动词、副词和其他多种语法结构,语言技术平台(Language Technology Platform,LTP)的依存句法分析关系如表 1 所列。

表 1 LTP 的依存句法分析关系

Table 1 Dependency syntactic analysis of LTP

关系类型	Tag
状中结构	ADV
主谓关系	SBV
定中关系	ATT
动补关系	CMP
动宾关系	VOB
并列关系	COO
介宾关系	POB
核心关系	HED

实际上,语法结构对于模型识别文本的情感非常重要,特别是对于包含相同单词但表达不同情感的微博。例如:“我不是很喜欢。”和“我很不喜欢!”,前者表示的是一种消极的情感,但后者负面的感情比前者更加强烈。通过分析句子中不同元素之间的关系,来提高细粒度情感分类模型的性能。

对于每个微博,基于解析依赖树的 GCN 模型都会建立一个图, $G=(V,E)$,其中, V 是顶点集,由构成微博的所有单词组成, E 是边集,表示单词与单词之间的依赖关系。其中,单词与单词之间的依赖关系有:不相关、自环关系、对头依赖关系和头对依赖关系,分别用数字“0”“1”“1”“1”来表示。例如:我不喜欢苹果,图 4 标注了不同的语法结构,根据不同单词之间的语法依赖关系可构造对应的邻接矩阵(见图 5)。又因为每个微博都不能超过 140 个字,所以将每个邻接矩阵的大小设置为 $[140 \times 140]$ 。将上一层 Dimension Weight Mechanism 层的输出 H 作为 GCN 层的输入,根据 Kipf&Welling^[34],GCN 网络可以计算为:

$$S=\text{relu}(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H\phi)$$
 (16)

其中, A 为邻接矩阵, $\tilde{D}$ 为度矩阵, $\tilde{D}_{ii}=\sum_j A_{ij}$, H 为上一层 Dimension Weight Mechanism 层的输出, ϕ 是由网络训练得到的权重向量。

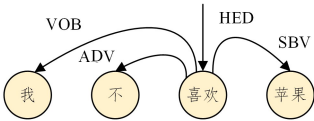


图 4 一个解析依赖树的例子

Fig. 4 Example of dependency parse tree

	我	不	喜欢	苹果
我	1	0	1	0
不	0	1	1	0
喜欢	1	1	1	1
苹果	0	0	1	0

图 5 一个句子对应的邻接矩阵

Fig. 5 Adjacency matrix of a sentence

3.5 百分比池化

池化层的目的是改善神经网络模型的不变性和效率,它

包括平均池化、最大池化等方式。但是,模型中可能存在一些不可控制的意外噪声,最大池化不够稳定,平均池化受异常值的影响较大。为了解决该问题,文献[35]采用了一种百分比池化,它将元素进行升序排序后,取最低 $p\%$ 值,第 p 个百分位数比平均值更健壮,因为它对异常值不敏感。百分比池化将向量 Z 的 p 个百分位数表示为函数 $f_p(Z)$,其中 p 的范围为 $0\sim 100$ 。例如, $f_{100}(Z)$ 是最大值, $f_{50}(Z)$ 是 Z 的中位数。将 $f_p(Z)$ 用作百分比池化的池化函数,通过实验可知,当 $p=50$ 时,取得最佳结果。

3.6 损失函数

本文在损失函数中添加了正则项,其是由权重矩阵 W_i 的转置与它本身的乘积和单位矩阵 I 的差值构成,以控制梯度消失和爆炸的问题,计算式如下:

$$Loss=loss(y,f_w(x))+\lambda\sum_i\|W_i^TW_i-I\|^2$$
 (17)

其中, $loss$ 是原始损失函数, y 是标签, f_w 是预测类, λ 是罚系数, W_i 是权重矩阵,而 I 是单位矩阵。

4 实验

本文将模型应用于微博细粒度情感分类。在实验中,所有代码均在 Linux CUDA 平台上运行的 PyTorch1.2.0 上实现,在 1080Ti GPU 计算机上训练 WDW-M-GCN 模型大约需要 6 h。

4.1 数据集

本文选择了 NLP&CC2013 数据集 4 和搜狗新闻数据集,为提高模型的泛化能力,本文采用文献[35]中的 15664 条带标签的新浪微博,并将微博文本情感分别分为 7 类和 2 类,7 类为幸福、悲伤、喜欢、愤怒、讨厌、恐惧和惊奇,2 类为积极和消极。搜狗新闻数据集包括近 19 万条搜狗新闻文本,将搜狗新闻文本最终分为 10 类,分别为金融、现实、股票、教育、科学、社会、政治、体育、游戏、娱乐。

4.2 性能指标

本文中的测试数据集源自中文微博文本 task5 中的 NLP&CC 2013 情绪分析,为了方便比较,本文使用与 2020 年发表的微博文本细粒度情感分类相同的指标。评估指标是宏观精度 $macro_precision$ 、微观精度 $micro_precision$ 、宏观召回率 $macro_recall$ 、微观召回率 $micro_recall$ 以及 F 度量的宏观平均值 $macro_f$ 和微观平均值 $micro_f$,其定义如下:

$$macro_precision=\frac{1}{7}\sum_i\frac{\#s_correct(emo=i)}{\#s_proposed(emo=i)}$$
 (18)

$$macro_recall=\frac{1}{7}\sum_i\frac{\#s_correct(emo=i)}{\#gold(emo=i)}$$
 (19)

$$macro_f=\frac{2\times macro_precision\times macro_recall}{macro_precision+macro_recall}$$
 (20)

$$micro_precision=\frac{\sum_i\#s_correct(emo=i)}{\sum_i\#s_proposed(emo=i)}$$
 (21)

$$micro_recall=\frac{\sum_i\#s_correct(emo=i)}{\sum_i\#gold(emo=i)}$$
 (22)

$$micro_f=\frac{2\times micro_precision\times micro_recall}{micro_precision+micro_recall}$$
 (23)

其中, $\#gold$ 是测试集手动注释的标签数量, $\#s_proposed$ 是在测试集上分类的标签数量, $\#s_correct$ 是正确分类的

微博数量。 i 是幸福、悲伤、喜欢、愤怒、讨厌、恐惧和惊奇之间的一种情感类型。

4.3 参数设置

WDWM-GCN 模型的参数设置如表 2 所列。

表 2 参数设置

Table 2 Parameter settings

Parameters		Values / Descriptions
Word embedding	Embedding size	300
Bi-LSTM	Max time step	140
	Hidden neurons	180
	Hidden layers	2
	Batch normalization	True
	Dropout rate	0.5
GCN	Network size (size of Θ)	[360,7]
	Layers number	1
Pooling	Pooling method	50th percentile pooling
Regularization	Loss function	Cross entropy
	L2 regularization	1×10^{-8}
	Orthogonal regularization	1×10^{-8}
Training	Optimizer	Adam
	Mini-batch size	32
	Learning rate	0.001
	Weight decay rate	1×10^{-8}

4.4 模型对比

本文将情绪的极性分为 7 类,即幸福、悲伤、喜欢、愤怒、讨厌、恐惧和惊奇,同时选择了几种情感分类算法作为基准,包括传统的机器学习方法和最新的神经网络体系结构。表 3 列出了针对 NLP&CC2013 数据集、基准模型以及最新的神经网络模型的实验结果。

表 3 不同模型在 NLP&CC2013 数据集上的表现

Table 3 Performance of different models on NLP&CC2013 dataset

(单位: %)

Method	<i>macro_f</i>	<i>micro_f</i>
SVM vote ^[36]	23.62	31.11
MCNN ^[37]	34.62	43.65
EMCNN ^[38]	35.17	44.22
CNN ^[35]	67.42	74.23
LSTM ^[35]	68.32	73.42
LSTM+CNN ^[35]	69.89	76.42
LSTM+GCN	74.81	77.99
(without syntax information) ^[35]		
Syntax-based GCN ^[35]	79.93	82.32
WDWM-GCN(Our Approach)	81.88	84.02

- (1)CNN:它采用卷积和池化操作来抽取文本的情感特征。
- (2)LSTM:从文本的两个方向来抓取词间关系,从而抽取更加丰富的文本特征。
- (3)CNN+LSTM:将 CNN 与 LSTM 相结合,抽取文本的关键语义和关键特征,但并未将文本转化成为图的形式。
- (4)LSTM+GCN(无语法信息):将 LSTM 与 GCN 相结合,但并未涉及语法信息。
- (5)Syntax-based GCN:将 GCN 与 Dependency Tree 相结合,引入句子的语法信息,使得分类结果更加准确。但并未考虑单词的多维度特性,WDWM-GCN 模型引入了一种双权重机制,不仅能关注句子的语法信息,还能够对单词的多维度特性进行有效融合。

4.5 情感七元分类模型的性能对比

表 3 列出了不同模型在 NLP&CC2013 数据集上的实验结果,并对不同模型的实验结果的文献来源都进行了标注。

传统的机器学习模型分类的准确率并不理想,相比传统的机器学习,深度学习模型 CNN 和 LSTM 的分类性能有很大的提高,其中 GCN 与 LSTM 的结合相对于 CNN 与 LSTM 的结合,*macro_f* 提高了 4.92%,这充分说明 GCN 能够捕获长距离的信息依赖,2020 年提出的 Syntax-based GCN 融入了句子的语法信息,其 *macro_f* 比 LSTM 与 GCN 的结合提高了 5.12%,而 WDWM-GCN 的 *micro_f* 比 Syntax-based GCN 提高了 1.7%。

为了进一步分析实验结果,本文针对宏观精度 *macro_precision*、微观精度 *micro_precision*、宏观召回率 *macro_recall*、微观召回率 *micro_recall* 以及 F 度量的宏观平均值 *macro_f* 和微观平均值 *micro_f* 这 6 个性能指标都进行了对比。由图 6—图 11 可知,WDWM-GCN 模型在 6 个性能指标上的表现均优于 Syntax-based GCN,充分说明了 WDWM-GCN 模型不仅能关注句子的语法信息,更重要的是它能选择对情感分类任务有效的关键词并抽取单词的重要维度特性,获取更加丰富的特征信息,从而提高文本分类的准确率。

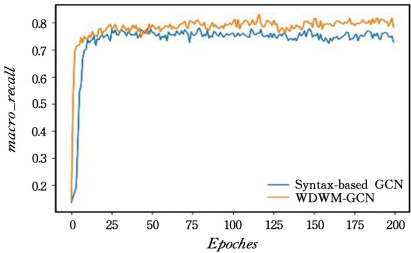


图 6 WDWM-GCN 和 Syntax-based GCN 的宏观召回率比较

Fig. 6 Comparison of *macro_recall* performance between WDWM-GCN and Syntax-based GCN

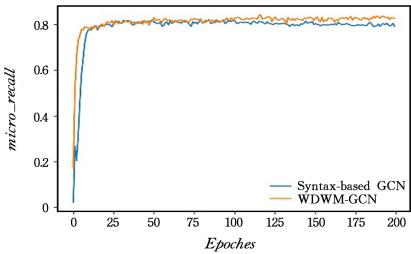


图 7 WDWM-GCN 和 Syntax-based GCN 的微观召回率比较

Fig. 7 Comparison of *micro_recall* performance between WDWM-GCN and Syntax-based GCN

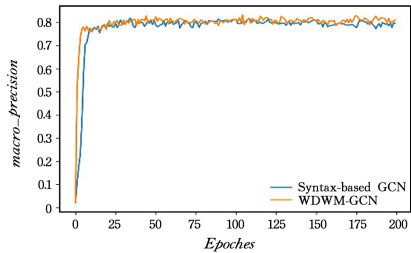


图 8 WDWM-GCN 和 Syntax-based GCN 的宏观精确率比较

Fig. 8 Comparison of *macro_precision* performance between WDWM-GCN and Syntax-based GCN

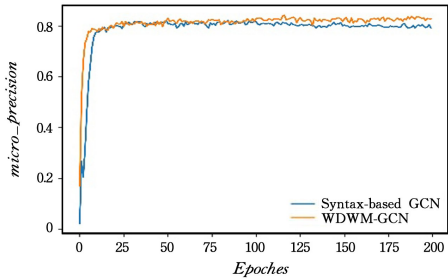


图 9 WDW-GCN 和 Syntax-based GCN 的微观精确率比较

Fig. 9 Comparison of *micro_precision* performance between WDW-GCN and Syntax-based GCN

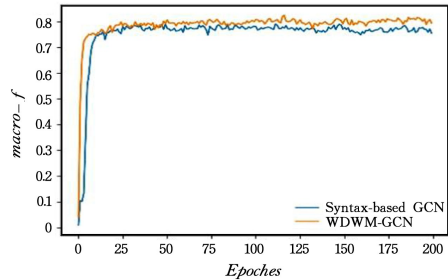


图 10 WDW-GCN 和 Syntax-based GCN 的 F 度量的宏观平均值比较

Fig. 10 Comparison of *macro_f* performance between WDW-GCN and Syntax-based GCN

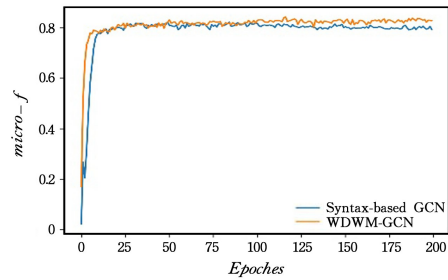


图 11 WDW-GCN 和 Syntax-based GCN 的 F 度量的微观平均值比较

Fig. 11 Comparison of *micro_f* performance between WDW-GCN and Syntax-based GCN

4.6 情感二元分类模型的性能对比

本文将情绪的极性分为两类:积极和消极。在这种情况下,我们删除了惊讶以及含糊不清的情感类型,将幸福的情感类型划分为积极,将悲伤、恐惧、讨厌、愤怒的情感类型划分为消极。为了适应二元分类问题,本文将图卷积的输出维度设置为 2,其余参数保持不变。同样,我们对不同模型的实验结果的文献来源都进行了标注。实验结果如表 4 所列。

表 4 情感二元分类模型的比较

Table 4 Comparison of emotional binary classification models (单位: %)

Method	Accuracy	Precision	Recall
LSTM ^[35]	90.01	90.07	90.15
CNN ^[35]	91.65	91.62	92.65
Syntax-based GCN ^[35]	92.04	92.19	92.04
WDWM-GCN	92.77	92.76	92.70

2020 年提出的 Syntax-based GCN 将 LSTM 与基于解析依赖树的 GCN 模型相结合,其准确率比单独使用 CNN 进行情感分类提高了 0.39%,比单独使用 LSTM 进行情感分类提高了 2.03%,但在微博情感分类中,每个单词对情感分类任务的重要性是不同的,并且单词的每一维度对单词的特性(词义、词性、语义)的影响也是有区别的,WDWM-GCN 提出了一种双重权重机制 WDW,并且将其与基于解析依赖树的 GCN 模型相融合,它的准确率比 Syntax-based GCN 提高了 0.73%,因为它能选择每条微博中的关键单词并且有效地抽取单词的多种重要维度特性,最大程度地保留文本的关键信息,从而提高了情感分类的准确率。

4.7 搜狗新闻数据集上文本分类模型的性能对比

通过实验可知,双重权重机制 WDW 能显著提高文本分类模型的性能。为了验证这一点,本文在搜狗新闻数据集上对不同的文本分类模型进行了实验,并且对不同模型的文献来源进行了标注。实验结果如表 5 所列。

表 5 文本分类模型加入 WDW 前后的结果比较

Table 5 Comparison of the different text classification model before and after the intergration of WDW (单位: %)

Method	Accuracy	Accuracy(WDW-based)
TextCNN ^[39]	90.79	91.61
TextRNN ^[40]	90.81	91.10
FastTest ^[41]	91.79	92.36
DPCNN ^[42]	89.99	90.62
Transformer ^[31]	89.50	90.09
Bert ^[43]	94.24	94.72

根据表 5 所列的实验结果可知,WDWM 能够提高文本分类模型的性能,分类效果平均提高了 0.5%。其中,TextCNN 模型的分类结果提高了近 1%,TextRNN 模型的分类结果没有明显改进,提高了 0.29%。但整体上,所有模型的分类结果都有提高,这说明 WDW 在文本分类任务中能提高分类的准确性和精度。

结束语 针对细粒度文本情感分类模型所面临的挑战,本文提出了一种双重权重机制 WDW,通过将基于解析依赖树的 GCN 模型与 WDW 相结合,有效地进行细粒度情感分类。实验结果表明,WDWM-GCN 模型能捕获关键单词的信息并有效融合单词的重要维度特性,最大程度地保留文本的关键信息,其性能优于最新的算法。同时,此机制可以有效地与分类模型结合,提高模型的性能。通过相关实验,进一步验证了此模型的通用性。

本文提出的 WDW 采用的是一维池化操作,在一定程度上会丢失部分信息,后期我们会对 WDW 作出进一步的改进,使其在未来能够应用于更多的领域。

参 考 文 献

[1] HU M, LIU B. Mining and Summarizing Customer Reviews [C] // Tenth Acm Sigkdd International Conference on Knowledge Discovery & Data Mining ACM. 2004:168-177.

- [2] LI R, LIN Z, LIN H L, et al. A Review of Text Sentiment Analysis[J]. Journal of Computer Research and Development, 2018, 55(1):30-52.
- [3] ZHAO J Q, GUI X L, ZHANG X J. Deep Convolution Neural Networks for Twitter Sentiment Analysis[J]. IEEE Access, 2018, 6(1):23253-23260.
- [4] WEI X, TAO L. Aspect Based Sentiment Analysis with Gated Convolutional Networks[J]. arXiv:1805.07043, 2018.
- [5] ALIAKSEI S, MOSCHITTI A. Twitter Sentiment Analysis with Deep Convolutional Neural Networks[C]// International ACM Sigir Conference. ACM, 2015:959-962.
- [6] PWINT OO S H, THEERAMUNKONG T, HUNG N D. Sentence Sentiment Classification Using Convolutional Neural Network in Myanmar Texts[C]// 2020 2nd International Conference on Image, Video and Signal Processing (IVSP '20). 2020.
- [7] SOCHER R, PERELYGIN A, WU Y J, et al. Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank[C]// Conference on Empirical Methods in Natural Language Processing (EMNLP 2013). 2013.
- [8] ABDUL-MAGEED M, UNGAR L. Emonet: Fine-grained Emotion Detection with Gated Recurrent Neural Networks[C]// Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL'17). 2017:718-728.
- [9] SACHIN S, TRIPATHI A, MAHAJAN N, et al. Sentiment Analysis Using Gated Recurrent Neural Networks[J]. SN Computer Science, 2020, 1(2):1-13.
- [10] BAZIOTIS C, PELEKIS N, DOULKERIDIS C. Datastories at semeval-2017 Task 4: Deep LSTM with Attention for Message-level and Topic-based Sentiment Analysis [C]// Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017). 2017.
- [11] LU Y, ZHOU J, DAI H N, et al. Sentiment Analysis of Chinese Microblog Based on Stacked Bidirectional LSTM[C]// 2018 15th International Symposium on Pervasive Systems, Algorithms and Networks (I-SPAN). 2018.
- [12] QIAN Q, HUANG M, LEI J, et al. Linguistically Regularized LSTM for Sentiment Classification[C]// Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. 2017.
- [13] BAHDANAU D, CHO K, BENGIO Y. Neural Machine Translation by Jointly Learning to Align and Translate[J]. arXiv:1409.0473, 2014.
- [14] TRAN N K, NIEDERREÉ C. Multihop Attention Networks for Question Answer Matching[C]// The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR'18). 2018:325-334.
- [15] LIU P, LIU T W, SHI J, et al. Aspect level sentiment classification with unbiased attention and target enhanced representations [C]// Proceedings of the 35th Annual ACM Symposium on Applied Computing (SAC '20). 2020:843-850.
- [16] LEI Z, YANG Y, YANG M, et al. A Multi-sentiment-resource Enhanced Attention Network for Sentiment Classification[C]// Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Short Papers). 2018.
- [17] SHAO X L. Fine-Grained Sentiment Analysis Based on Hierarchical Attention Networks[J]. Computer Ence and Application, 2019, 9(11):2143-2153.
- [18] WANG Y, HUANG M, ZHU X, et al. Attention-based LSTM for Aspect-level Sentiment Classification[C]// Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. 2016.
- [19] WU X H, CHEN L, WEI T T, et al. Sentiment analysis of Chinese short text based on Self-Attention and Bi-LSTM[J]. Journal of Chinese Information Processing, 2019, 33(6):100-107.
- [20] WANG Z T, YU Z W, GUO B, et al. Sentiment analysis of Chinese microblog based on lexicon and rule set[J]. Computer Engineering & Applications, 2015, 51(8):218-225.
- [21] VILARES D, ALONSO M A, GÓMEZ-RODRÍGUEZ C. Supervised Sentiment Analysis in Multilingual Environments[J]. Information Processing & Management, 2017, 53(3):595-607.
- [22] GAMON M. Sentiment classification on customer feedback data [C]// Proceedings of the 20th International Conference on Computational Linguistics. 2004:841-847.
- [23] LIU P F, MENG H L. Conditional random fields labeling and maximum entropy classification for aspect based sentiment analysis[C]// International Workshop on Semantic Evaluation. 2014.
- [24] BO P, LILLIAN L, SHIVAKUMAR V. Thumbs up? Sentiment Classification using Machine Learning Techniques[C]// EMNLP. 2002:79-86.
- [25] CHEN K, LIANG B, KE W D, et al. Sentiment Analysis of Chinese Weibo Based on Multi-channel Convolutional Neural Network[J]. Computer Research and Development, 2018, 55(5):945-957.
- [26] ZHENG S M, YANG M. A New Method of Improving BERT for Text Classification [C]// International Conference on Intelligent Science and Big Data Engineering. 2019:442-452.
- [27] ZHENG Z, LI L, JING C. Deeply Hierarchical Bi-directional LSTM for Sentiment Classification[J]. Computer Science, 2018, 45(8):213-217.
- [28] WANG J L, YANG Y X, WANG X M. ResLCNN model for short text classification[J]. Journal of Software, 2017, 28(s2):61-69.
- [29] YANG Z C, YANG D Y, DYER C, et al. Hierarchical attention networks for document classification[C]// ACL. 2016:1480-1489.
- [30] LUONG M, PHAM H, MANNING C. Effective approaches to attention-based neural machine translation[C]// EMNLP. 2015.
- [31] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]// Proceedings of the 31st International Conference on Neural Information Processing Systems. 2017:6000-6010.

[32] DU C, HUANG L. Text Classification Research with Attention-based Recurrent Neural Networks[J]. International Journal of Computers, Communications & Control (IJCCC), 2018, 13(1): 50.

[33] ZHANG J, DUAN L G, LI A P, et al. Fine-grained Sentiment Analysis Based on Combination of Attention and Gated Mechanism[J]. Computer Science, 2021, 48(8): 226-233.

[34] KIP F T, WELLING M. Semi-supervised classification with graph convolutional networks[C]// ICLR. 2016.

[35] LAI Y, ZHANG L, HAN D, et al. Fine-Grained Emotion Classification of Chinese Microblogs Based on Graph Convolution Networks[J]. World Wide Web, 2020, 23(4): 1-17.

[36] WEN S, WAN X. Emotion Classification in Microblog Texts Using Class Sequential Rules[C]// AAAI. 2014: 187-193.

[37] JIANG F, LIU Y, LUAN H, et al. Microblog Sentiment Analysis with Emoticon Space Model[C]// Chinese National Conference on Social Media Processing. Berlin: Springer, 2014.

[38] YE P, KUMAR J, KANG L, et al. Unsupervised feature learning framework for no-reference image quality assessment[C]// 2012 IEEE Conference on Computer Vision and Pattern Recognition. 2012: 1098-1105.

[39] YOON K. Convolutional Neural Networks for Sentence Classification[C]// Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. 2014.

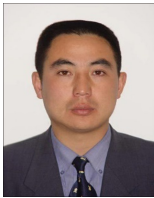
[40] LIU P F, QIU X P, HUANG X J, et al. Recurrent Neural Network for Text Classification with Multi-Task Learning[C]//

Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence. 2016: 2873-2879.

[41] JOULIN A, GRAVE E, BOJANOWSKI P, et al. Bag of Tricks for Efficient Text Classification[J]. arXiv:1607.01759, 2016.

[42] JOHNSON R, ZHANG T. Deep Pyramid Convolutional Neural Networks for Text Categorization[C]// Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Long Papers). 2017.

[43] DEVLIN J, CHANG M W, LEE K, et al. Pre-training of Deep Bidirectional Transformers for Language Understanding [J]. arXiv:1810.04805, 2018.



LI Hao, born in 1970, professor, Ph.D. His main research interests include distributed computing, grid and cloud computing research.



KANG Yan, born in 1972, Ph.D, associate professor. Her main research interests include software engineering, system optimization, big data processing and mining.

(责任编辑:柯颖)