

基于 CNN 与词语邻近特征的情感分类模型

吕 超, 杨 超, 李仁发

(湖南大学 信息科学与工程学院, 长沙 410082)

摘 要: 基于卷积神经网络(CNN)的方法在情感分类任务中得到广泛应用,该方法使用词向量作为网络的输入,但是,在卷积过程中每个词向量只能表征一个单词,并不蕴含上下文信息,导致了信息传递连续性的降低。为此,构建一种基于词语邻近特征的 CNN 模型,在卷积过程中让每个词向量携带邻近词语的特征,既保证信息传递的连续性,也保证词向量在局部范围内的序列性。实验结果表明,在 COAE2014(二分类)和 COAE2015(三分类)的情感分类任务上,该模型的准确率分别达到 89.43% 和 85.61%,验证了其可行性和高效性。

关键词: 情感分类;卷积神经网络;词向量;连续性;序列性;邻近特征

中文引用格式: 吕 超, 杨 超, 李仁发. 基于 CNN 与词语邻近特征的情感分类模型[J]. 计算机工程, 2018, 44(5): 182-187.

英文引用格式: LÜ Chao, YANG Chao, LI Renfa. Sentiment Classification Model Based on CNN and Word Adjacent Feature[J]. Computer Engineering, 2018, 44(5): 182-187.

Sentiment Classification Model Based on CNN and Word Adjacent Feature

LÜ Chao, YANG Chao, LI Renfa

(School of Information Science and Engineering, Hunan University, Changsha 410082, China)

[Abstract] The method based on Convolutional Neural Network (CNN) is widely used in the task of sentiment classification. This method mainly uses word vectors as the input of the network. However, in the convolution process, each word vector can only represent one word and does not contain context information, which reduces the continuity of information transmission. Aiming at this problem, a CNN model based on word adjacent feature is proposed, which allows each word vector to carry the characteristics of its adjacent words in the convolution process, which not only guarantees continuity of information transmission, but also ensures the sequentiality of word vectors in local range. The experimental results show that the accuracy rate of the proposed method is 89.43% and 85.61% respectively in the COAE2014 (2 classification) and COAE2015 (3 classification) emotional classification tasks, which verifies the feasibility and efficiency of the model.

[Key words] sentiment classification; Convolutional Neural Network (CNN); word vector; continuity; sequentiality; adjacent feature

DOI: 10.19678/j.issn.1000-3428.0046274

0 概述

随着微博、微信等社交媒体的兴起,网络上带有评论性的文本资源数量与日俱增。对于该大量非结构化的文本,迫切需要通过技术手段对其情感倾向进行分类判断,以达到多方面的目的,如获取大众对于有关政策的意见、对相关产品的满意程度和对各种社会热点事件的看法等^[1]。但是,微博文本篇幅较短、经常存在错别字和语法错误、充满大量网络词汇和口语词汇,这些问题都给情感分类任务带来了挑战,引发了国内外学者极大的关注^[2]。

目前,基于卷积神经网络(Convolutional Neural Network, CNN)的情感分类方法得到了广泛关注。CNN 通过多个并行的卷积操作获取文本的局部特征^[3],经过池化层将筛选出的局部特征拼接后传递给全连接层进行分类。然而大多数的 CNN 模型在卷积操作和特征拼接环节中,很大程度上忽略了文本的有序性,即词语之间的序列性,因此,其不能保证信息传递的连续性。例如,“我/喜欢/自然/语言/处理/但是/它的/学习/门槛/很高”,假设卷积核的尺寸是 5,则被卷积的分片依次是“我/喜欢/自然/语言/处理”“喜欢/自然/语言/处理/但是”“自然/

基金项目: 国家自然科学基金(61173036); 国家高技术研究发展计划项目(2012AA01A301-01); 湖南省科技计划项目(2015GK3015)。

作者简介: 吕 超(1987—),男,硕士研究生,主研方向为深度学习、自然语言处理; 杨 超,副教授、博士; 李仁发,教授、博士生导师。

收稿日期: 2017-03-08 **修回日期:** 2017-04-25 **E-mail:** chao9038@hnu.edu.cn

语言/处理/但是/它的”“语言/处理/但是/它的/学习”“处理/但是/它的/学习/门槛”及“但是/它的/学习/门槛/很高”,则得到的局部信息可能是“喜欢”“自然”“语言”“但是”“门槛”“很高”,也可能是“自然”“语言”“它的”“处理”“学习”“很高”,前者明显表达了负面情感,但后者很可能被误判为带有正面情感,因为其中的“很高”带有积极情绪。在现实生活中,人们也常使用一些脱离语境的极性词对情感进行评估,比如“大”“小”“高”“低”等,然而这些词在不同的语境中有着不同的含义,因此,无法仅凭某些极性词就分辨出整个文本的情感倾向,此时必须要考虑语境。

针对上述问题,本文建立一种基于词语邻近特征的CNN模型,在卷积过程中考虑每个词语前后邻近词的特征,以此保证信息传递的连续性。

1 相关工作

近年来,深度学习模型在计算机视觉^[4]和语音识别^[5]邻域取得了一定成果。深度学习模型通常是具有多层结构的神经网络,其强调从大规模的数据集中学习各种现实事物的特征表示,且这些特征能够被计算机直接应用于各种计算模型中。在自然语言处理领域,大多数深度学习方法通过神经网络模型学习单词的词向量表达^[6],在此基础上再进一步进行处理,如进行语义分析^[7]、搜索词检索^[8]、句子分类^[9]等。目前,主流的神经网络模型是递归神经网络(如RNN、LSTM)和卷积神经网络(如DCNN)。递归神经网络将文本看作有序的词语序列,并且考虑词语间的关联性,其充分利用上下文信息从而较好地进行了语言建模。其中RNTN模型^[10]利用情感树库在二元化的句法树结构上进行语义合成,文献^[11]在此模型基础上,通过对依存句法树进行变换,完成句中指定评价对象的情感极性分析。文献^[12]通过添加语义合成函数的pooling层,使得RNTN可以自适应地选择最适合当前节点的语义合成函数,从而大幅提高情感分类效果。但是该类RNN模型仅考虑文本的有序性而忽略词语在语义上的结构性,且该类模型在训练过程中存在梯度爆炸和梯度消失问题^[13],导致训练过程难以收敛。由于微博文本的不规范性,导致各种句法分析结果通常不准确,因而CNN模型成为微博情感分类的首选模型。文献^[9]提出一种DCNN模型,通过学习句子结构来提高情感语义合成性能。文献^[14]提出CharSCNN模型,为英语单词增加字符表示层的特征,以提高对于英文Twitter中不规则单词的表示能力。文献^[15]首次尝试将CNN应用到句子分类任务中,通过词向量训练工具word2vector生成的词向量,训练一个

简单的3层CNN,其在二分类和多分类数据集上均取得了很好的分类性能,说明预先训练的词向量是一类通用的特征向量,可以在不同的分类任务中直接使用。文献^[15]采用的MCNN是CNN的一种,能够对多个窗口大小内的词向量执行语义合成操作,其在多个不同的情感数据集上也取得很好的分类性能。文献^[16]基于MCNN提出EMCNN模型,其将基于表情符号的情感空间映射与深度学习模型MCNN进行结合,增强了MCNN捕捉情感语义的能力。文献^[17]提出一种LSTM-CNN模型,尝试从子句的角度来分析文本的情感极性,利用LSTM(长短时记忆网络)归纳某个子句的上下文信息后将其拼接在子句的词向量表达形式的两侧,再使用CNN对子句进行特征提取、分类,该模型在影评数据集上能取得比传统分类方法更好的性能。

上述模型都基于传统的CNN模型,在做卷积操作时依然无法保证词向量之间的序列性以及信息传递的连续性。本文在参考了文献^[15]的工作后,分别建立3种基于词语邻近特征的CNN模型LAF-CNN、RAF-CNN、LRAF-CNN,在卷积过程中让每个词向量携带邻近词语的特征,这样既保证信息传递的连续性,且在局部范围内确保词语的序列性不会被打乱,也在一定程度上使其具备RNN模型的归纳能力。

2 基于词语邻近特征的CNN模型

2.1 词语邻近特征

词语的邻近特征指与该词相邻词语的特征。假设原始句子为:

$$X = (w_1, w_2, \dots, w_{i-1}, w_i, \dots, w_n)^T \quad (1)$$

对应的词向量矩阵为:

$$M = (v_1, v_2, \dots, v_{i-1}, v_i, \dots, v_n)^T \quad (2)$$

其中, w_i 表示组成句子的各个词语, v_i 表示 w_i 对应的词向量特征。在卷积过程中,对于某个词语 w_i 采取3种不同的邻近特征附加策略:附加左邻近特征(Left Adjacent Feature, LAF),附加右邻近特征(Right Adjacent Feature, RAF),附加左右邻近特征(Left Right Adjacent Feature, LRAF),其作用都是为了增强词语在特定上下文中所表达的语义。若使用LAF策略,则卷积过程中使用词向量 M_{LAF} ,即从第一个词向量开始,每一个词向量都会与其左侧的词向量特征进行合并操作,该做法的好处是上一个词向量的信息可以传递给下一个词向量,因此,下一个词向量在参加卷积操作时可以提供相对充足的情感信息供卷积层提取有效特征。 M_{RAF} 中每一个词向量与其右侧的词向量特征合并, M_{LRAF} 中每一个词向量与其左、右侧的词向量特征合并。3种策略的直观描述如图1所示,其中 v_0 表示零向量。

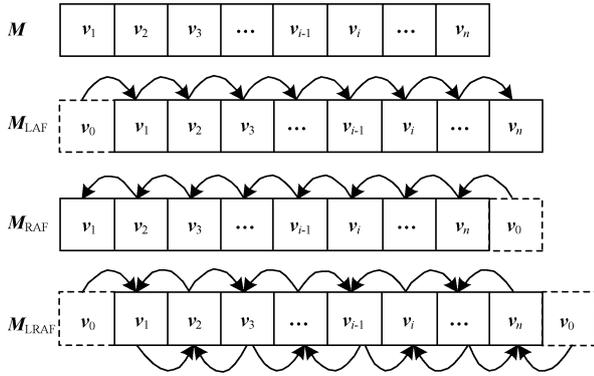


图1 3种词语邻近特征附加策略示意图

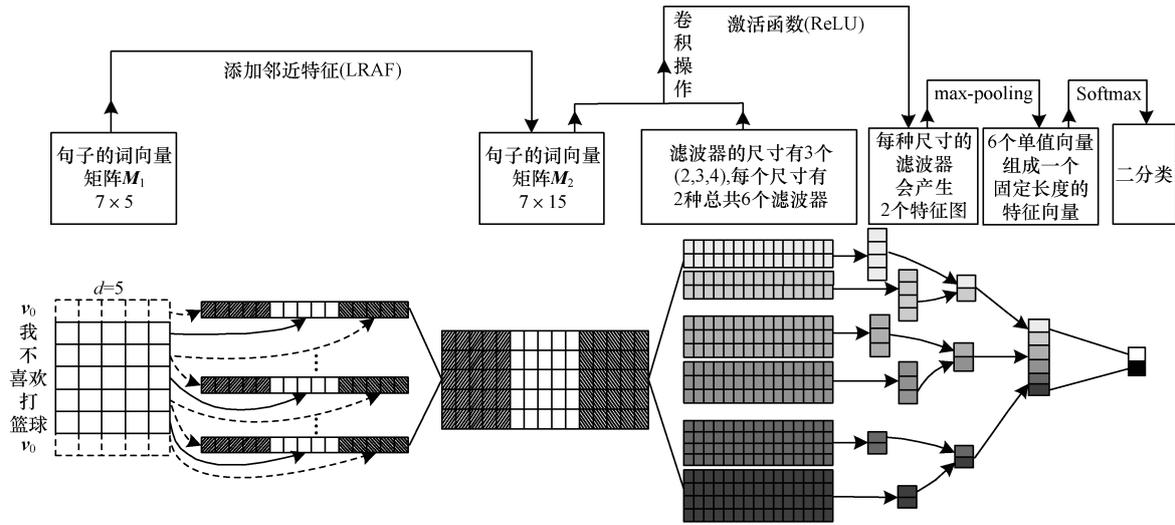


图2 基于词语邻近特征的 CNN 模型结构

假设输入层句子 S 的长度为 n , $v_i \in \mathbb{R}^d$ 表示 d 维的词向量, 其对应句子 S 中的第 i 个词语, 则句子 S 可以表示成:

$$\mathbf{v}_{1:n} = \mathbf{v}_1 \oplus \mathbf{v}_2 \oplus \cdots \oplus \mathbf{v}_i \oplus \mathbf{v}_{i+1} \oplus \cdots \oplus \mathbf{v}_n \quad (3)$$

其中, \oplus 表示串联操作, 也可以说, $\mathbf{v}_{i:i+j}$ 表示词向量 $\mathbf{v}_i, \mathbf{v}_{i+1}, \dots, \mathbf{v}_{i+j}$ 的串联。

如果使用邻近特征附加策略, 则句子 S 会有另外 3 种表示形式 (图 2 中使用 $\mathbf{v}_{1:n}^{\text{LRAF}}$)。

$$\mathbf{v}_{1:n}^{\text{LAF}} = (\mathbf{v}_0 \uplus \mathbf{v}_1) \oplus (\mathbf{v}_1 \uplus \mathbf{v}_2) \oplus \cdots \oplus (\mathbf{v}_{i-1} \uplus \mathbf{v}_i) \oplus \cdots \oplus (\mathbf{v}_{n-1} \uplus \mathbf{v}_n) \quad (4)$$

$$\mathbf{v}_{1:n}^{\text{RAF}} = (\mathbf{v}_1 \uplus \mathbf{v}_2) \oplus (\mathbf{v}_2 \uplus \mathbf{v}_3) \oplus \cdots \oplus (\mathbf{v}_i \uplus \mathbf{v}_{i+1}) \oplus \cdots \oplus (\mathbf{v}_{n-1} \uplus \mathbf{v}_n) \quad (5)$$

$$\mathbf{v}_{1:n}^{\text{LRAF}} = (\mathbf{v}_0 \uplus \mathbf{v}_1 \uplus \mathbf{v}_2) \oplus (\mathbf{v}_1 \uplus \mathbf{v}_2 \uplus \mathbf{v}_3) \oplus \cdots \oplus (\mathbf{v}_{i-1} \uplus \mathbf{v}_i \uplus \mathbf{v}_{i+1}) \oplus \cdots \oplus (\mathbf{v}_{n-1} \uplus \mathbf{v}_n \uplus \mathbf{v}_0) \quad (6)$$

其中, 运算符 \uplus 为 2 个向量的合并符号, 如向量 $\mathbf{a} = [1, 2, 3]$, 向量 $\mathbf{b} = [4, 5, 6]$, 则有 $\mathbf{a} \uplus \mathbf{b} = [1, 2, 3, 4, 5, 6]$ 。

卷积层利用大小为 $h \times k$ 的滤波器对输入特征矩阵进行卷积操作, 即:

$$c_i = f(w \cdot \mathbf{v}_{i:i+h-1}^{\text{LRAF}} + b) \quad (7)$$

2.2 CNN 模型

基于词语邻近特征的 CNN 模型如图 2 所示 (使用 LRAF 附加策略), 图中以“我不喜欢打篮球”为例。其中使用的卷积核有 3 个尺寸: 2, 3, 4, 每个尺寸的卷积核有 2 种, 每个卷积核对句子矩阵进行卷积操作生成长度可变的特征图后, 对其进行最大池化 (max-pooling) 操作, 即取出每个特征图中的最大值, 由这些单值生成一个固定大小的特征向量后将其以全连接的形式传给 Softmax 层, Softmax 层使用该特征向量作为输入同时计算出各个类别上的概率分布 (这里假设其是二分类, 因此, 图 2 中只有 2 个分类状态)。

其中, c_i 为特征图中第 i 个特征值, $f(\cdot)$ 为卷积核函数, $w \in \mathbb{R}^{h \times k}$ 为滤波器, h 为滑动窗口大小, b 为偏置值, $\mathbf{v}_{i:i+h-1}^{\text{LRAF}}$ 为由第 i 行到第 $i+h-1$ 行组成的局部特征矩阵。因此, 特征图 C 为:

$$C = [c_1, c_2, \dots, c_{n+h-1}] \quad (8)$$

在卷积层后, 加入了 ReLU 层, 将 ReLU 作为激活函数, 可以加快随机梯度下降的收敛速度^[16]。

下采样层使用文献 [18] 提出的 max-pooling 方法进行特征采样, 得到的特征值 \hat{c} 为:

$$\hat{c} = \max \{ C \} \quad (9)$$

卷积层和下采样层共同构成了 LRAF-CNN 的特征提取层。LRAF-CNN 由多个不同尺寸的特征提取层 (有多个 h 值) 并列组成, 每种尺寸的特征提取层各有 m 个, 因此, 全连接层的特征向量 \mathbf{z} 为:

$$\mathbf{z} = [\hat{c}_{1,h_1}, \hat{c}_{2,h_1}, \dots, \hat{c}_{i,h_1}, \dots, \hat{c}_{m,h_1}, \dots, \hat{c}_{1,h_j}, \hat{c}_{2,h_j}, \dots, \hat{c}_{i,h_j}, \dots, \hat{c}_{m,h_j}, \dots] \quad (10)$$

其中, \hat{c}_{i,h_j} 表示第 j 种类型的滤波器产生的第 i 个特征值。

相比普通的CNN模型,基于词语邻近特征的神经网络结构可以有效提取出与正负面情感标签相关的词向量特征,并用于最终的情感分类。

下采样层输出的多个单值特征作为全连接层的输入,全连接层将其进行拼接,形成特征向量,然后利用Softmax多分类函数输出各个分类的概率大小,从而确定分类结果,同时根据训练数据的实际分类标签,采用反向传播算法对模型参数进行梯度更新,即:

$$P(y=l|V;\theta) = \frac{e^{z^T\theta_l}}{\sum_{k=1}^K e^{z^T\theta_k}} \quad (11)$$

其中, y 表示句子的实际情感标签, θ 表示模型的参数集合,即神经元之间的权重矩阵,训练过程中 θ 会不断变化,直至收敛。

通过式(11)可以计算出给定句子的情感标签和模型参数集合 θ 的条件概率分布图,进而得出分类结果。

2.3 正则化

在网络的倒数第2层,引入Dropout策略并使用L2-constraint规则权重约束条件对网络参数进行正则化处理^[19]。Dropout通过在前向传播过程中随机丢弃一定比例的神经元(本文Dropout比例为0.5)以阻止神经元之间的相互适应。对于倒数第2层的输入 z ,不再使用式(12)计算输出神经元 y 。

$$y = w \cdot z + b \quad (12)$$

取而代之的是式(13)所示的Dropout策略。

$$y = w \cdot (z \circ r) + b \quad (13)$$

其中, \circ 为向量内积符号, $r \in \mathbb{R}^m$ 为遵从Bernoulli随机分布的“掩饰向量”,向量中元素值为1的概率是 p ,为0的概率是 $1-p$ 。梯度只在未进行掩饰处理的神经元中反向传播。

3 实验结果与分析

3.1 数据集

本文使用COAE2014、COAE2015任务4的微博数据集(<http://115.24.12.5/web/index.html>),其中观点倾向性分析评测(Chinese Opinion Analysis Evaluation,COAE)始于2008年,是国内情感分析方面的权威评测会议。COAE2014任务4(二分类)的微博数据集共包含40000条数据,其中官方公布了7000条微博的极性(负面有3224条,正面有3776条),另33000条是混淆语料,涵盖手机、翡翠、保险3个领域。COAE2015任务4(三分类)的微博数据集共包含133201条数据,其中官方公布了20154条微博的极性(负面有6209条,中性有1639条,正面有12306条),

另113047条是混淆语料,涵盖汽车、电子、手机、美食、娱乐、宾馆等多个领域。2个数据集的概要信息如表1所示。

表1 COAE2014、COAE2015数据集概要信息

数据集	c	l	N	V	V_{pre}	CV
COAE2014	二	100	7000	13993	13451	10
COAE2015	三	102	20154	19389	18384	10

其中, c 表示目标数据集是几分类, l 表示该数据集中最长的句子包含多少个词语, N 表示数据集中有多少条标注数据, V 表示标注数据的词汇量大小, V_{pre} 表示标注数据中含有多少已存在于预训练词向量中的词汇, $CV=10$,表示采用十折交叉验证法划分实验数据。

3.2 数据预处理与词向量预训练

在数据预处理方面,本文直接去除文本中所有标点符号和表情符号,只留下含义较多的中、英文单词,然后利用ICTCLAS(<http://ictclas.nlpir.org/>)分词工具对实验数据集进行分词。在词向量预训练方面,本文使用新浪微博提供的数据抽取接口,抽取1000万条微博数据,过去燥和预处理之后,再使用Google开源的词向量训练工具word2vec(<https://code.google.com/archive/p/word2vec/>)训练词向量。经过训练后,得到的词向量文件包含80552个词语,在COAE2014、COAE2015任务4数据集上的词汇覆盖率分别是96.13%、95.33%。

3.3 模型对比

各模型的训练过程如下:

1)CNN:是基准模型,使用预训练词向量对所有词语进行初始化,不在预训练词向量中的词语随机初始化(随机初始化范围是 $[-0.25, 0.25]$),且在训练过程中所有词向量保持不变。

2)LAF-CNN:词向量的初始化方案与CNN相同。训练过程中为每个词语添加左邻近特征,如果左邻近词语不在预训练词向量中,则随机初始化左邻近特征,且在训练过程中所有词向量保持不变。

3)RAF-CNN:词向量的初始化方案与CNN相同。训练过程中,为每个词语添加右邻近特征,如果右邻近词语不在预训练词向量中,则随机初始化右邻近特征,且在训练过程中所有词向量保持不变。

4)LRAF-CNN:词向量的初始化方案与CNN相同。在训练过程中,为每个词语添加左、右邻近特征,如果左、右邻近词语不在预训练词向量中,则随机初始化左、右邻近特征,且在训练过程中所有词向量保持不变。

上述4个模型除了邻近特征的附加策略不同,其余超参数设置均相同,超参数值如表2所示。

表2 模型的可调超参数值

可调参数	参数值
卷积核函数	ReLU
词向量维度	300
批处理大小	64
滤波器尺寸	3,4,5
每个尺寸的滤波器个数	128
参数更新比例	0.5
权重惩罚系数	1.0
迭代次数	200

模型在训练过程中,采用文献[20]提出的 Adadelta Update Rule 规则进行随机梯度下降更新参数。

表3 各模型在 COAE2014 数据集上的分类结果对比

模型	负面			正面			准确率
	查准率	召回率	F 值	查准率	召回率	F 值	
CNN	87.30	82.72	84.94	85.75	89.63	87.65	86.43
LAF-CNN	87.38	85.49	86.43	87.73	89.36	88.54	87.57
RAF-CNN	88.20	83.02	85.53	86.08	90.43	88.20	87.00
LRAF-CNN	90.97	85.98	88.40	88.21	92.47	90.29	89.43

表4 各模型在 COAE2015 数据集上的分类结果对比

模型	负面			中性			正面			准确率
	查准率	召回率	F 值	查准率	召回率	F 值	查准率	召回率	F 值	
CNN	83.02	80.13	81.55	56.00	23.46	33.07	85.64	94.32	89.77	83.78
LAF-CNN	80.19	88.67	84.22	55.91	30.95	39.85	90.83	91.29	91.06	85.41
RAF-CNN	76.29	88.70	82.03	64.00	9.82	17.02	89.71	92.75	91.20	84.87
LRAF-CNN	84.52	83.72	84.12	50.00	21.23	29.81	87.80	94.23	90.90	85.61

从2个数据集上的实验结果可以总结出:

1) 当加入邻近特征后,模型在综合性能方面得到了一定提升,这说明在卷积过程中,给每个词语添加邻近特征,一方面可以强化词语在特定上下文中所表达的语义,另一方面相邻词语之间的相互关系提高了信息传递的连续性。这2点在一定程度上可确保在卷积过程中,即使漏掉了个别情感词语,也不会对最终的分类结果产生很大影响。

2) LAF-CNN 的分类准确率均略高于 RAF-CNN,其原因可能跟词语的序列性有关,中文文本语义的表达是从左至右、先有上文后有下文,而 LAF-CNN 正具有这种承上启下的作用。然而具体是否因为这个原因造成该现象,还有待进一步验证。

3) LAF-CNN、RAF-CNN 的分类准确率均略低于 LRAF-CNN,其原因跟词语在特定上下文中所蕴含的语义有关,由于 LAF-CNN 只携带左邻近特征,RAF-CNN 只携带右邻近特征,而 LRAF-CNN 兼顾

3.4 测试结果分析

在 COAE2014 数据集(二分类)和 COAE2015 数据集(三分类)上的模型对比实验结果如表3和表4所示。实验效果评价指标为查准率、召回率、F 值及准确率,其中基准模型是未考虑任何邻近特征的 CNN。从表3可以看出,相比基准模型 CNN, LAF-CNN、RAF-CNN、LRAF-CNN 在查准率、召回率、F 值及准确率方面均有所提升,其中准确率分别提升了 1.14%、0.57%、3%。从表4可以看出,相比基准模型 CNN, LAF-CNN、RAF-CNN、LRAF-CNN 在查准率、召回率、F 值方面浮动比较大,这是由各个分类的训练集数量不均衡所导致,但模型在准确率方面依然提升了 1.63%、1.09%、1.83%。

左、右邻近特征,其可以提取到更精确的语义信息,因此,该模型的分类效果更好。

4 结束语

为在卷积过程中增强词语在当前上下文中的语义、保证词语在局部范围内的序列性,本文建立一种基于词语邻近特征的 CNN 模型,并为其设计了3种词语邻近特征附加策略。对比实验结果表明,相对传统 CNN 模型,该模型具有可行性、高效性。但是在附加邻近特征时,本文只考虑当前词语前、后邻近词的特征,下一步将在计算能力允许的情况下,扩大邻近词的数量从而获取更精确的上下文信息;此外,在语料预处理时,本文直接去掉所有标点符号和表情符号,然而某些标点符号和表情符号的去除会影响上下文的情感极性转移,对模型最后的分类效果产生一定的负面影响,因此,解决由标点符号和表情符号引起的情感极性转移问题也是今后的研究方向。

参考文献

- [1] PANG B, LEE L. Seeing stars: exploiting class relationships for sentiment categorization with respect to rating scales [C]//Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics. New York, USA: ACM Press, 2005: 115-124.
- [2] 韦航,王永恒. 基于主题的中文微博情感分析[J]. 计算机工程, 2015, 41(9): 238-244.
- [3] CUN L, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [4] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet classification with deep convolutional neural networks [C]//Proceedings of International Conference on Neural Information Processing Systems. [S. l.]: Curran Associates, Inc., 2012: 1097-1105.
- [5] GRAVES A, MOHAMED A, HINTON G. Speech recognition with deep recurrent neural networks [C]//Proceedings of 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. Washington D. C., USA: IEEE Press, 2013: 6645-6649.
- [6] MIKOLOV T, SUTSKEVER I, CHEN K, et al. Distributed representations of words and phrases and their compositionality [C]//Proceedings of the 26th International Conference on Neural Information Processing Systems. [S. l.]: Curran Associates, Inc., 2013: 3111-3119.
- [7] YIH W T, HE X, MEEK C. Semantic parsing for single-relation question answering [EB/OL]. [2017-02-25]. <https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/SingleRelationQA-YihHeMeek-ACL14.pdf>.
- [8] SHEN Y, HE X, GAO J, et al. Learning semantic representations using convolutional neural networks for Web search [C]//Proceedings of the 23rd International Conference on World Wide Web. New York, USA: ACM Press, 2014: 373-374.
- [9] KALCHBRENNER N, GREFFENSTETTE E, BLUNSOM P. A convolutional neural network for modelling sentences [EB/OL]. [2017-02-26]. <http://aclweb.org/anthology/P14-1062>.
- [10] SOCHER R, PERELYGIN A, WU J Y, et al. Recursive deep models for semantic compositionality over a sentiment treebank [EB/OL]. [2017-02-15]. https://nlp.stanford.edu/~socherr/EMNLP2013_RNTN.pdf.
- [11] DONG L, WEI F, TAN C, et al. Adaptive recursive neural network for target-dependent twitter sentiment Classification [EB/OL]. [2017-02-20]. <http://acl2014.org/acl2014/P14-2/pdf/P14-2009.pdf>.
- [12] DONG L, WEI F, ZHOU M, et al. Adaptive multi-compositionality for recursive neural models with applications to sentiment analysis [C]//Proceedings of the 28th AAAI Conference on Artificial Intelligence. Menlo Park, USA: AAAI Press, 2014: 1537-1543.
- [13] KOLEN J, KREMER S. Gradient flow in recurrent nets; the difficulty of learning long term dependencies [M]. Washington D. C., USA: Wiley-IEEE Press, 2001: 237-243.
- [14] SANTOS C N D, GATTI M. Deep convolutional neural networks for sentiment analysis of short texts [EB/OL]. [2017-02-10]. <http://www.aclweb.org/anthology/C14-1008>.
- [15] KIM Y. Convolutional neural networks for sentence classification [EB/OL]. [2017-02-15]. <http://aclweb.org/anthology/D14-1181>.
- [16] 何炎祥,孙松涛,牛菲菲,等. 用于微博情感分析的一种情感语义增强的深度学习模型[J]. 计算机学报, 2016, 39(4): 1-19.
- [17] CHEN Peng, XU Bing, YANG Muyun, et al. Clause sentiment identification based on convolutional neural network with context embedding [C]//Proceedings of the 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery. Washington D. C., USA: IEEE Press, 2016: 1532-1538.
- [18] COLLOBERT R, WESTON J, BOTTOU L, et al. Natural language processing (almost) from scratch [J]. Journal of Machine Learning Research, 2011, 12(1): 2493-2537.
- [19] HINTON G E, SRIVASTAVA N, KRIZHEVSKY A, et al. Improving neural networks by preventing co-adaptation of feature detectors [J]. Computer Science, 2012, 3(4): 212-223.
- [20] ZEILER M D. Adadelta: an adaptive learning rate method [J]. Computer Science, 2012, 3(4): 325-330.
- 编辑 吴云芳
- (上接第181页)
- [11] 任晓龙,吕琳媛. 网络重要节点排序方法综述[J]. 科学通报, 2014, 59(13): 1175-1197.
- [12] FREEMAN L C. A set of measures of centrality based on betweenness [J]. Sociometry, 1977, 40(1): 35-41.
- [13] BONACICH P. Factoring and weighting approaches to status scores and clique identification [J]. Journal of Mathematical Sociology, 1972, 2(1): 113-120.
- [14] PARSONS T D. The search number of a connected graph [C]//Proceedings of the 9th Conference on Combinatorics, Graph Theory, and Computing. Winnipeg, Canada: [s. n.], 1978: 549-554.
- [15] NEWMAN M E J. The structure of scientific collaboration networks [J]. Proceedings of the National Academy of Sciences, 2001, 98(2): 404-409.
- [16] BORGATTI S P, MEHRA A, BRASS D J, et al. Network analysis in the social sciences [J]. Science, 2009, 323(5916): 892-895.
- [17] 罗由平,周召敏,李丽娟,等. 基于幂率分布的社交网络规律分析[J]. 计算机工程, 2015, 41(7): 299-304.
- 编辑 索书志