

文章编号:1671-6833(2020)01-0058-05

融合位置注意力机制和改进 BLSTM 的食品评论情感分析

李 勇,金庆雨,张青川

(北京工商大学 农产品质量安全追溯技术及应用国家工程实验室,北京 100048)

摘 要: 为了对食品评价的情感倾向进行更加精确的分类,在进行情感语义分析时,卷积神经网络(convolutional neural networks,CNN)方法在提取目标的结构特征方面具有一定的优势,可以提取到多种局部特征,循环神经网络(recurrent neural networks,RNN)具有记忆功能,在序列特征提取方面具有一定的优势,双向长短时记忆网络(bidirectional long short-term memory,BLSTM)在提取远距离依赖序列语义特征方面可以得到很好的效果。在 BLSTM 的基础上,又引入基于食品领域的语义角色标注与位置相结合的位置注意力机制,来实现距离相关的序列语义特征提取,利用 CNN 实现序列语义特征的情感语义分类,从而构造出了一种基于 BLSTM 和位置注意力机制的食品评论情感分析模型。实验结果表明,设计的模型在情感分类方面取得了很好的分类效果,与之前的情感分类模型进行比较,在准确率结果上有所提高。

关键词: 情感分析;评论;双向长短时记忆网络;卷积神经网络;位置注意力机制

中图分类号: TP391.1 **文献标志码:** A **doi:**10.13705/j.issn.1671-6833.2020.01.006

0 引言

随着互联网的快速发展,社交网站和各大电商平台迅猛发展。互联网上的消费者用户在浏览、购买商品的同时,也倾向于发表自己的意见,对所购买的商品、店家的服务等进行评论和反馈^[1]。在线购物的消费者,他们在购物的过程中产生的浏览记录、购买记录和商品评论等数据,生成了数据量庞大的互联网信息。这些非结构化的互联网信息,包含了众多消费者对各类商品的观点和态度。商家可以根据这些商品评论信息制定销售战略,消费者也可以将这些评论数据作为参考,选择合适的商品。因此,对这些互联网信息进行分析 and 利用是十分必要的。随着互联网的普及,人们越来越倾向于在网上进行购物,用户评论数据信息规模不断地增大,仅仅依靠人工的方式进行处理已经变得不再现实,越来越多的学者利用自然语言处理技术对互联网信息进行情感分析,这一研究成为自然语言处理领域的一个研究

热点^[2]。情感分析是对人们的观点评价进行情感的倾向性分析,对产品、服务或者事件进行挖掘和分析以及一系列推理和归纳的技术^[3]。在商品的评论中不是每个词都是包含情感信息成分的,或者并不能明显地表明评论者的态度,含有情感成分的词主要是形容词、动词和部分名词等,这些词是情感分析关注的重点。目前,情感分析的主要研究方法是基于机器学习的传统算法,基于机器学习的方法需要通过使用大量的人工标注数据的特征来确定给定文本的情感极性,这项工作非常费时费力^[4]。随着深度学习、人工智能等相关技术的发展,深度神经网络技术成为了自然语言处理领域的关键技术,而且在文本情感分析中也得到了很好的应用,取得了不错的效果^[5]。笔者提出了一种基于长短时记忆网络与位置注意力机制融合并结合卷积神经网络的情感分类模型。该模型充分利用 BLSTM 的特性,挖掘评论文本的语义特征,并与

收稿日期:2019-06-16;修订日期:2019-10-09
基金项目:国家重点研发计划项目(2016YFD0401205);北京市自然科学基金资助项目(4202014);北京市自然科学基金资助项目(4184084);教育部人文社会科学研究青年基金资助项目(17YJCZH007);科技创新服务能力建设项目(19005857028);北京市教委科技类一般项目(SQKM201710011008)
作者简介:李勇(1979—),男,河北邢台人,北京工商大学讲师,博士,主要从事食品安全大数据研究,E-mail:liyong@btbu.edu.cn。

位置注意力机制相结合,通过 BLSTM 的训练获取评论中详细的特征信息,使用位置注意力机制计算,使情感相关的词语对整个评论起决定性的作用。最后通过 CNN 来进行特征的分类,从而提高了对食品评论信息情感分类的精度。实验结果也表明了该模型在情感分类方面得到了非常大的提高,获得了不错的分类效果。

1 相关工作

现有的情感分析方法主要由基于规则的情感分类方法、基于机器学习的情感分类方法向基于深度学习的情感分析方法的方向发展,分析结果的准确率也在不断地提高^[6]。

在进行情感分析时,采用基于规则方法的过程中需要有很多人参与的工作,其中包括情感词典的构建和一些语言结构的归纳总结等。对这些内容进行分析时需要通过构建相关的情感词典和那些文本数据中包含的情感词进行对比来计算文本的情感倾向性。Kim 等^[7]在对话题评价对象进行情感分析时,使用了概率的方法对每个词赋予一定的情感强度,并根据这些情感词的强度来进行评估得分,最后通过把这些情感词的分数相加得出每个话题对象的情感倾向,取得了非常好的效果。王志涛等^[8]通过基于词典和规则集的中文微博情感分析方法,并根据其在微博中的特性,定义了不同语言层面的规则,将情感词典应用在从词到句子的不同文本中,进行了多粒度的情感计算,并在实验数据集上证明了该方法是可行的和有效的。

基于机器学习的情感分析方法通过输入大量的标注语料以及这些标注语料的情感标签,训练和这些语料相关的评论数据的情感分类器,然后通过这些训练好的分类器来预测新的文本数据的情感^[9]。情感分析的主要研究方法还是一些基于机器学习的传统算法,例如,信息熵、支持向量机、条件随机场等^[10]。这些机器学习的方法大体可以分为 3 类:有监督的机器学习、无监督的机器学习和半监督的机器学习^[11]。王新宇^[12]通过对旅游网络点评的情感倾向性进行分析,使用向量空间模型来表示评论,用情感词典对特征空间进行降维,通过 SVM 机器学习模型进行分类,最终得到了有效的分类结果。通过使用机器学习的方法来进行情感分析,最终的分类结果往往是由对特征对象的选取来决定的,特征对象的选取直接影响分类的效果,由于个体之间存在着很大的差

异性,通过人工选择特征有着很大的局限性和不确定性,难以真正发现和挖掘文本深层次的特征。

深度学习的过程实际上是在模拟人的神经元之间进行信息传递的过程^[13],深度学习是为了使得最终的模型可以像人一样进行数据的学习和分析,进而解释数据。目前主要应用在图像处理、声音识别和文本分析等领域。李阳辉等^[14]通过采用降噪自编码对文本数据进行无标记的特征学习来进行情感分类,并通过实验获得了比较好的结果。李章晓等^[15]将深度学习的技术应用到了金融领域,通过建立模型对外汇预测和投资组合优化进行实验,表明实验方法具有有效性。随着深度学习理论研究的逐渐深入,深度学习也应用到了情感分析方面,李杰等^[16]通过采用卷积神经网络对短文本评论信息进行了情感分类,得到了高准确率的分

2 食品评论情感分析模型

笔者设计的情感分析模型如图 1 所示,其中 X 代表模型第 1 部分的输出。主要由两部分组成:①文本情感收集器;②情感信息分类器。

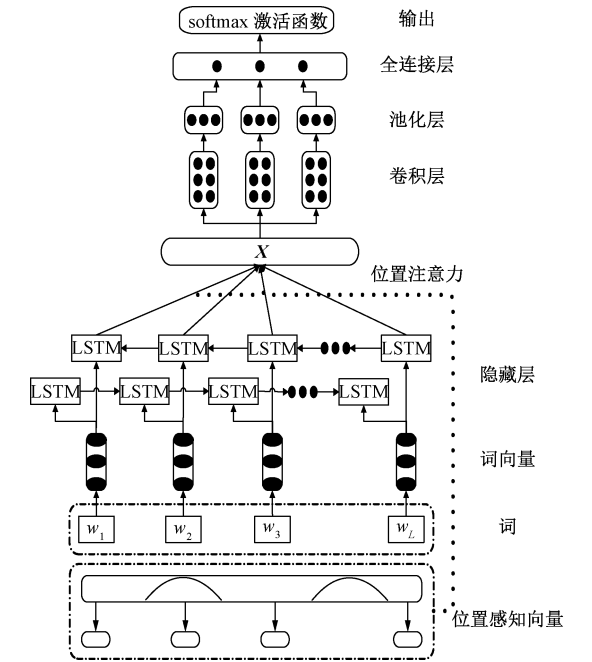


图 1 情感分析模型结构

Figure 1 Sentiment analysis model structure

文本情感收集器是基于改进的 BLSTM 并融合位置注意力向量来对评论中的情感信息进行抽取。情感信息分类器是将文本情感收集器的输出作为输入,通过卷积神经网络来进一步进行情感

语义特征的分类。

2.1 文本情感收集器

此部分主要是对评论中的情感信息进行抽取,每条食品的评论信息由 $w = \{w_1, w_2, \dots, w_L\}$ 来表示,其相应的词向量 $v = \{v_1, v_2, \dots, v_L\}$ 。模型的输入是由食品评论句子的各个词组成,每个词转成对应的词向量。每一组词向量通过输入左右两个 LSTM 模型中进行训练,将左右两侧模型训练的结果进行合并,得到更为合理的信息矩阵,同时引入位置注意力机制,每个词都有自己对应的注意力参数,最终得到更为优化的情感信息。

2.2 LSTM 模型

由于 RNN 在处理长序列输入时还存在缺陷,如梯度消失的问题。为了解决这个问题,在 RNN 中加入更多的记忆单元来控制信息在不同时刻的流动,从而解决了梯度消失问题,模型的网络通过记忆单元更新各个节点信息,从而可以学习文本序列中那些需要进行远距离依赖的特性,提高模型的准确率。

将评论语料中的每个句子转为向量表示,将词向量作为模型的一个输入序列 $v = \{v_1, v_2, \dots, v_L\}$,LSTM 计算隐藏向量序列 $h = [h_1, h_2, \dots, h_L]$ 和输出矩阵序列 $X = [x_1, x_2, \dots, x_L]$ 。LSTM 模型引入记忆单元进行信息之间的传递,LSTM 中的这些记忆单元帮助它解决避免梯度消失的问题,适合学习长期依赖的上下文语义。与传统的 RNN 相比,LSTM 增加了输入门、遗忘门和输出门。笔者采用 LSTM 模型,将上一个细胞状态同时引入到输入门、遗忘门以及新信息的计算当中^[17],以下为 LSTM 的计算过程:

$$i_t = \sigma(W_i[h_{t-1}; w_t] + b_i); \tag{1}$$

$$f_t = \sigma(W_f[h_{t-1}; w_t] + b_f); \tag{2}$$

$$o_t = \sigma(W_o[h_{t-1}; w_t] + b_o); \tag{3}$$

$$g_t = \tanh(W_c[h_{t-1}; w_t] + b_c); \tag{4}$$

$$c_t = i_t \odot g_t + f_t \odot c_{t-1}; \tag{5}$$

$$h_t = o_t \odot \tanh(c_t), \tag{6}$$

式中: i 为输入门; f 为遗忘门; o 为输出门; b 为偏执; W 为对应权重。

通常一个标准的 LSTM 只从一个方向对序列进行编码。然而两个 LSTM 也可以堆叠起来作为双向使用编码器,称为双向 LSTM,笔者采用此种方式进行模型的设计。通过两层相反方向流处理数据兼顾了历史信息和未来信息,一层从左到右的顺序,另一层从右到左的顺序,最终将两层输出作为一个整体,作为 BLSTM 隐藏层的输出。为了

突出情感词在句子中的作用,笔者加入了位置注意力机制,将 BLSTM 的输出进行了微调。

2.3 位置感知注意力机制

随着深度学习的不断发展,基于注意力机制的神经网络模型越来越多地应用到自然语言处理领域、图像识别领域、语音识别领域等不同的研究内容当中。注意力机制最早是在视觉图像领域提出来的,Bahdanau 等^[18]将这种注意力机制应用在机器翻译任务上,并将翻译任务和对齐工作同时进行,这些应用逐渐扩展到各种自然语言处理(natural language processing, NLP)任务中。

笔者提出了一个位置感知的注意力机制,首先利用语义角色标注的自然语言处理技术对评论进行处理,将句子分割成不同的句子成分,将每个句子成分与情感词库进行相似度的比对,当句子成分中的词与情感词库的词相似度大于 0.85 时,确定该句子成分中的词为情感分析的核心词。在一个句子成分中,起着关键作用的核心词对周边词的影响程度会随着距离的变化而变化。因核心词对其周边词的影响程度是不同的,笔者通过使用高斯核函数来模拟基于位置感知的影响传播:

$$Kernel(u) = \exp\left(\frac{-u^2}{2\sigma^2}\right), \tag{7}$$

式中: u 代表核心词与句子成分中当前词的距离; σ 是一个约束传播范围的参数; $Kernel(u)$ 表示基于内核距离为 u 所得到的相应的影响。

$$K(i, u): N(Kernel(u), \sigma'), \tag{8}$$

式中: $K(i, u)$ 表示第 i 个维度上的核心词在对距离为 u 的词的影响; N 是具有 $Kernel(u)$ 值的期望值和标准差 σ' 的正态密度。

$$x_i = k_i h_i. \tag{9}$$

2.4 情感信息分类器

在以上工作完成后,将商品的评论信息嵌入到矩阵 X 中,由于每条评论信息中总会有或多或少的噪音,为了得到精确的数据信息,笔者设计了情感信息分类器,此部分是由 3 个并列的过滤器组成,每一个过滤器独自抽取矩阵 X 中的情感信息,最后将每部分进行结合得到输出结果,经过进一步的特征提取获得最终的分类类别。

2.5 卷积神经网络

卷积神经网络广泛应用在图像识别和文本分类领域。它是文本分类使用最多的深度学习网络结构,其底层由卷积层、池化层交替组成,顶端使用全连接层来完成具体的任务。

本文中卷积层是由 3 种窗口大小不同的卷积

核组成,用来进行提取数据内部的语义特征。通过最大池化层提取其中的主要特征,在最后一层通过全连接层完成情感极性正向情感或者负向情感的映射。卷积核 $w \in \mathbf{R}^{n \times k}$,卷积的窗口大小为 n ,输出特征为:

$$s_i = f(w_i \cdot x_i + b_i)。$$

(10)

在本实验中考考虑到收敛速度的问题,采用relu函数作为激活函数进行非线性操作,得到的 s_i 代表通过卷积获得的局部特征,最终得到特征向量集合 S 。通过以上卷积过程所有的输出特征都是独立计算的,对于每一个过滤器来说,采用max-pooling方式来降低特征向量的大小:

$$m_i = \max S。$$

(11)

模型的最后一部分是输出层。笔者设计了3个过滤器在全连接层的输出:

$$y = \beta_1 m^1 + \beta_2 m^2 + \beta_3 m^3。$$

(12)

最后,将全连接层的输出 y 输入到softmax函数中,将输出转换为概率进行分类:

$$p_i = \frac{e^{y_i}}{\sum_{k=1}^n e^{y_k}}。$$

(13)

3 对比实验与结果分析

3.1 数据集

在京东商城食品类别的用户评论数据集上评估笔者设计的方法。将整个数据集划分为训练集、测试集。其中的训练数据有80 000条,测试数据有20 000条。数据如表1所示。

表 1 实验数据

Table 1 The experimental data	
数据类别	数据量/条
精选肉类	50 000
新鲜水果	20 000
休闲食品	30 000

3.2 实验参数设置

实验参数的调整对实验结果的影响很大,通过比较,最终选取了以下参数:在BLSTM中,词向量的维度为200,层数为2,学习率为0.001,dropout的值为0.5,设置epoch为50。在CNN中,词向量的维度为200,窗口分别为2、3、4,激活函数为relu函数。

3.3 实验评价指标

采用准确率 P 和召回率 R 来评价实验分类结果的好坏,准确率表示预测类的样本中有多少是预测正确的;召回率表示真实标签为测试类的

有多少是被预测正确的,具体公式如下:

$$P = \frac{C}{O};$$

(14)

$$R = \frac{C}{L},$$

(15)

式中: C 表示模型返回测试类预测正确的数量; O 表示模型返回的总数量; L 表示测试类的总数量。使用 F_1 测度来评价准确率和召回率:

$$F_1 = \frac{2 \cdot P \cdot R}{P + R}。$$

(16)

3.4 实验结果与分析

笔者对比了以下几组实验,如表2所示。

表 2 对比实验结果

Table 2 The results of the contrast experiments			
模型	准确率	召回率	F_1 测度
CNN	0.784 7	0.761 5	0.772 9
Att-CNN	0.805 1	0.793 8	0.799 4
BLSTM	0.846 2	0.835 4	0.840 8
Att-BLSTM	0.852 7	0.843 5	0.848 1
本文模型	0.869 4	0.854 2	0.861 7

从实验结果可以发现,笔者提出的模型方法的准确率要比其他几种基准方法都要高。

在精选肉类中有这样一条评论:“什么精品好肉,一塌糊涂,腥的一踏糊涂,直接处理掉了”。这条评论在其他的几种方法都归类为了积极的评价,只有本文的模型归类为消极评价。这充分体现了本文模型的一个优点。对于情感词库中没有的词,通过相似度计算将“一塌糊涂”及包含一个错字的“一踏糊涂”归为了消极情感的词,通过位置注意力机制对其向量进行调整从而准确地将此评论归为消极评论,提高了情感分类的准确率。

4 结论

笔者在传统BLSTM模型的基础上,通过将食品领域相关情感词的位置感知引入注意力机制,突出了情感信息在评论中的情感语义极性,融合CNN来实现情感语义特征分类,从而提出了一种面向情感信息抽取和情感语义分类的食品评论情感分析方法。对比实验的结果表明,笔者提出的方法是可行的和有效的,通过对评论情感信息抽取部分进行组合提取可以获取句子中更多的语义特征,从而提高了情感分类的精度。

参考文献:

[1] 任远远,王卫平.中文网络评论的产品特征提取及情感倾向判定[J].计算机系统应用,2014,23(11):

22-27.

[2] MANKE S N,SHIVALE N.A review on:opinion mining and sentiment analysis based on natural language processing[J].International journal of computer applications, 2015,109(4):29-32.

[3] 李晨,朱世伟,魏墨济,等.基于词典与规则的新闻文本情感倾向性分析[J].山东科学,2017,30(1):115-121.

[4] 杜昌顺,黄磊.分段卷积神经网络在文本情感分析中的应用[J].计算机工程与科学,2017,39(1):173-179.

[5] 陈珂,梁斌,柯文德,等.基于多通道卷积神经网络的中文微博情感分析[J].计算机研究与发展,2018,55(5):945-957.

[6] 郭利敏.基于卷积神经网络的文献自动分类研究[J].图书与情报,2017(6):96-103.

[7] KIM S M, HOVY E. Extracting opinions, opinion holders, and topics expressed in online news media text [C]//Proceedings of the Workshop on Sentiment and Subjectivity in Text - SST '06, July 22, 2006. Morristown: Association for Computational Linguistics, 2006: 1-8.

[8] 王志涛,於志文,郭斌,等.基于词典和规则集的中文微博情感分析[J].计算机工程与应用,2015,51(8):218-225.

[9] 邢长征,李珊.文本情感分析的深度学习方法[J].计算机应用与软件,2018,35(8):102-106.

[10] 刘智鹏,何中市,何伟东,等.基于深度学习的商品评价情感分析与研究[J].计算机与数字工程,2018,46(5):921-927.

[11] 陶宏曜,梁栋屹.基于卷积神经网络的职位描述文本分类方法[J].软件,2017,38(6):30-34.

[12] 王新宇.基于情感词典与机器学习的旅游网络评价情感分析研究[J].计算机与数字工程,2016,44(4):578-582,766.

[13] 焦李成,杨淑媛,刘芳,等.神经网络七十年:回顾与展望[J].计算机学报,2016,39(8):1697-1716.

[14] 李阳辉,谢明,易阳.基于深度学习的社交网络平台细粒度情感分析[J].计算机应用研究,2017,34(3):743-747.

[15] 李章晓,宋微,田野.基于深度学习和进化计算的外汇预测与投资组合优化[J].郑州大学学报(工学版),2019,40(1):92-96.

[16] 李杰,李欢.基于深度学习的短文本评论产品特征提取及情感分类研究[J].情报理论与实践,2018,41(2):143-148.

[17] 梁军,柴玉梅,原慧斌,等.基于极性转移和 LSTM 递归网络的情感分析[J].中文信息学报,2015,29(5):152-159.

[18] BAHDANAU D, CHO K, BENGIO Y. Neural machine translation by jointly learning to align and translate [J]. Computer and language, 2014(9):1-15.

Improved BLSTM Food Review Sentiment Analysis
with Positional Attention Mechanisms

LI Yong, JIN Qingyu, ZHANG Qingchuan

(National Engineering Laboratory for Agri-product Quality Traceability, Beijing Technology and Business University, Beijing 100048 China)

Abstract: In order to more accurately classify the emotional tendency of food reviews. In the case of emotional semantic analysis, CNN had certain advantages in extracting the structural features of the target, and could extract a variety of local features. With memory function, RNN had certain advantages in sequence feature extraction. Bidirectional Long Short-Term Memory (BLSTM) could achieve good results in extracting remote-dependent sequence semantic features. Based on BLSTM, this paper introduced a positional attention mechanism based on the combination of semantic role labeling and location in the food field. The distance-related sequence semantic feature extraction was realized, and the sentiment semantic classification of sequence semantic features was realized by CNN, and a food comment sentiment analysis model based on BLSTM and positional attention mechanism was constructed. Experimental results showed that the model designed could achieve good results in emotional classification, and an improved accuracy rate compared with the previous sentiment classification model.

Key words: sentiment analysis; review; blstm; CNN; positional attention mechanisms