

张宇, 吴静. 基于深度学习的商品评论情感分析[J]. 智能计算机与应用, 2024, 14(8): 54-58. DOI: 10.20169/j.issn.2095-2163.240809

基于深度学习的商品评论情感分析

张宇, 吴静

(浙江理工大学 计算机科学与技术学院, 杭州 310018)

摘要: 商品评论情感分析的研究通常是对商品的评论进行情感分类, 从中挖掘出用户的兴趣爱好。传统的情感分析模型在提取语义特征时不够全面, 不能准确全面的获取商品评论中蕴含的信息, 使得情感分类的准确率较低, 本文提出了一种基于深度学习的商品评论情感分析方法 (BERT-CNN-BiLSTM-Attention, BCBA)。首先, 使用 BERT 模型进行词向量表达, 获取商品评论的特征词向量; 其次, 通过 CNN 获取商品评论中的局部特征信息, 通过 BiLSTM 获取商品评论中的上下文语义特征信息, 将获取到的两种特征信息进行融合; 最后, 将融合后的特征信息输入到注意力机制中, 对重要特征信息赋予更多的权重, 完成情感分类任务。通过与其他传统情感分析模型对比的实验结果表明, 本文提出的 BCBA 模型能够有效地提高情感分类的准确率。

关键词: 情感分析; BERT; CNN; BiLSTM

中图分类号: TP391.1

文献标志码: A

文章编号: 2095-2163(2024)08-0054-05

Sentiment analysis of commodity reviews based on deep learning

ZHANG Yu, WU Jing

(School of Computer Science & Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China)

Abstract: In the research of sentiment analysis of commodity reviews, it is usually used to classify commodity reviews into sentiment to dig out users' interests and hobbies. The traditional sentiment analysis model is not comprehensive enough when extracting semantic features, and can not accurately and comprehensively obtain the information contained in commodity reviews, which makes the accuracy of sentiment classification low. This paper proposes a deep learning based product comment sentiment analysis (BERT-CNN-BiLSTM-Attention). Firstly, BERT model is used to express the word vectors to obtain the feature word vectors of commodity reviews. Secondly, the local feature information in commodity reviews is obtained through CNN and the contextual semantic feature information in commodity reviews is obtained through BiLSTM, and the two features obtained are integrated. Finally, the fused feature information is input into the attention mechanism, and more weight is given to the important feature information to complete the emotion classification task. Compared with other traditional emotion analysis models, the experimental results show that the BCBA model proposed in this paper can effectively improve the accuracy of emotion classification.

Key words: sentiment analysis; BERT; CNN; BiLSTM

0 引言

近年来, 对商品评论进行情感分析成为国内外研究的热点, 情感分析的主要任务是帮助用户在面对某个对象评价正负面问题时, 对具有感情色彩的文本进行计算和分析^[1]。在情感分析时, 需要将商品评论文本转化为词向量的形式, 传统的词向量转换做法是通过 One-hot 编码、Word2Vec 等方式, 这些方式都是基于词语本身含义进行转化, 对中文表

达中一词多义的情况无法解决^[2-3]。如某款手机的评论“手机待机时间很久”和“物流等待了很长的时间”, 其中的“久”字在两个语句中表达了不同的情感极性, 在第一个语句中表示积极, 在第二个语句中表示消极, 如果使用同样的词向量进行表示, 显然是有误差的。

目前关于情感分析的研究主要包括 3 种: 基于情感词典的情感分析、基于机器学习的情感分析和基于深度学习的情感分析^[4]。深度学习的情感分

作者简介: 吴静 (1997-), 女, 硕士研究生, 主要研究方向: 自然语言处理。

通讯作者: 张宇 (1982-), 女, 博士, 副教授, 主要研究方向: 数据挖掘, 自然语言处理。Email: yzh@zstu.edu.cn

收稿日期: 2023-04-23

析方法使用神经网络模型作为情感分类器,是目前情感分类方法中表现最好的一种^[5]。常见的神经网络模型包括:卷积神经网络模型(Convolutional Neutral Network, CNN)和循环神经网络模型(Recurrent Neutral Network, RNN)。RNN 的变体形式 BiLSTM(Bi-directional Long Short-Term Memory)有效地解决了模型的梯度爆炸等问题,并且能够获取文本的上下文信息,使得文本的信息得到充分利用^[6]。

2018 年,Google 的 Devlin 等^[7]构建了一个新的预训练模型(Bidirectional Encoder Representation from Transformers, BERT)。BERT 模型是一个基于微调的将文本转化为特征向量的词表示模型,由输入层、编码层和输出层组成。BERT 模型将词向量、文本向量和位置向量的加和作为模型的输入向量^[8-9]。BERT 模型采用双向 Transformer 编码器中的编码部分,编码部分主要包含前馈神经网络和多头自注意力。BERT 模型对输入的信息进行多头计算、残差计算和归一化处理,得到的输出即为词向量。

本文采用 BERT 模型将文本词语转化为动态词向量的形式,融合评论文本的词语信息和位置信息,使得词向量的表达更加精准。使用 CNN 和 BiLSTM 分别提取评论文本的局部特征和上下文语义关系,充分挖掘文本中隐含的深层次特征信息;再将并行组合之后的模型融入到注意力机制中,使得模型高度关注重要词语的特征,完成最终的情感分类任务。

1 相关工作

1.1 词向量

计算机中的文本数据是以自然语言的形式保存的,由于计算机不能直接识别人类的语言,因此可以将这些原始的人类语言转化为计算机可以识别的数据类型,常用的方法是将文本转化为向量或者矩阵进行存储^[10]。One-hot 编码为每个词语确定唯一的索引,若在语料库中找到了文本中的词语,则将该词语对应的位置表示为 1,剩余的位置表示为 0。Word2vec 是由 Google 研究团队提出的词向量工具,包括两种结构:词袋模型结构(Continues Bag Of Words, CBOW)和跳字模型结构(Skip-gram)^[11-12]。

1.2 情感分析技术

随着网络的不断发展,越来越多人开始在网络中分享自己的个人感受和看法,国内外很多学者也逐步将情感分析作为自己研究的方向。Turney

等^[13]通过对基本的褒贬词汇进行扩充,使用情感极性语义算法分析情感倾向,情感分类的准确率达到了 74%;Pang 等^[14]分别使用朴素贝叶斯、支持向量机和最大熵算法对电影的评论数据进行分析,实验结果表明,支持向量机方法在电影评论数据中的分类结果最好;Cao 等^[15]通过 CNN 算法对训练样本和测试样本进行实验,结果表明与其他传统分类方法相比,该方法的分类效果更好;Xu 等^[16]使用 BiLSTM 进行文本情感分析,解决了 RNN 中容易出现梯度爆炸等问题。深度学习的情感分析方法具有自适应、自组织和实时学习的特点,利用神经网络进行情感分析,解决了传统的情感分析方法在特征提取时的问题,如花费大量代价进行计算,特征提取不全面从而导致的准确度较低。

2 情感分析模型

本文提出情感分析模型(BERT-CNN-BiLSTM-Attention, BCBA),使用 BERT 模型来获取评论文本的词向量,再使用 CNN 和 BiLSTM 的组合模型进行特征提取,将 CNN 和 BiLSTM 的输出向量进行拼接,得到含有评论文本局部特征和上下文特征的向量,最后通过注意力机制赋予重要特征更多的权重,输出情感分析的结果。具体流程如图 1 所示。

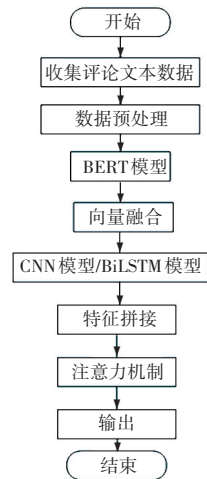


图 1 BCBA 情感分析流程图

Fig. 1 BCBA emotion analysis flow chart

2.1 输入层

获取用户对商品的评论文本,用户评论是用户购买商品后的反馈意见,用户发表的评论为自由评论,一般并不是标准化数据,因此需要去除数据集中的异常评论和没有价值的评论,主要包括:系统默认好评和用户的重复评论内容;利用停用词典去除助词、拟声词和虚词等,使用 jieba 分词进行中文分

词,得到数据集 D 。

将数据集 D 输入到 BERT 模型,文本数据通过词向量、分段向量和位置向量后,将数据集 D 依次传入前馈神经网络和多头自注意力,再通过线性变换、归一化等操作获得词语的向量表示,将 BERT 语义特征向量 b_i 定义为式(1):

$$b_i = w_i(\omega + \sigma + \rho) \quad (1)$$

其中, ω 、 σ 、 ρ 分别表示词向量、分段向量和位置向量。

2.2 情感分析层

在局部特征提取中,CNN 主要由输入层、卷积层、池化层和全连接层 4 个部分组成。CNN 的核心组成部分为卷积层,CNN 由多个卷积层组成,并且每个卷积层中可以有多不同的卷积核。输入层后面连接的是大小不同的滤波器,每一个滤波器都对应一个滤波映射后得出的新特征图,表示提取到的不同角度的特征信息。将经过 BERT 处理后的词向量输入到 CNN 中,滤波器经过一个高度为 h 的词序列窗口时的卷积操作结果如式(2):

$$c_i = f(W \otimes H_{i:i+h-1} + b) \quad (2)$$

其中, c_i 表示从 i 个词向量到第 $i+h-1$ 个词向量卷积操作后的特征图; f 为激活函数; W 为模型参数矩阵; \otimes 表示卷积运算; $H_{i:i+h-1}$ 表示第 i 个到第 $i+h-1$ 个词向量卷积运算; b 为偏置项。

采用 ReLU 函数作为激活函数:

$$f(x) = \max(0, x) \quad (3)$$

经过卷积运算后,通过池化层可以减少模型的参数数量:

$$z_i = \max(c_i) \quad (4)$$

最后,将池化操作后的输出进行拼接,再通过全连接层进行组合,得到输出向量 H_s 。

在上下文语义提取模型中,本文采用 BiLSTM 结构。BiLSTM 可以向两个时序相反的 LSTM 网络中同时输入,分别学习当前词语的上文表示和下文表示,并将上文表示和下文表示进行拼接。双向 LSTM 共同决定了 BiLSTM 的最终结果,计算过程如下:

$$h_l = \overrightarrow{\text{LSTM}}(b_i) \quad (5)$$

$$h_g = \overleftarrow{\text{LSTM}}(b_i) \quad (6)$$

$$H_r = W_l h_l + W_g h_g + b_r \quad (7)$$

其中, h_l 表示前向 LSTM 的输出结果; h_g 表示后向 LSTM 的输出结果; H_r 表示 BiLSTM 的输出结果。

2.3 注意力机制层

注意力机制层的输入为通过 CNN 的输出向量 H_s 和 BiLSTM 的输出向量 H_r 拼接之后的结果,将拼接后的文本特征向量表示为 H , 具体计算如下:

(1) 计算注意力权重 u_i :

$$u_i = v_i \tan(W_u H + b_u) \quad (8)$$

其中, W_u 表示权重矩阵, b_u 表示偏置向量。

(2) 使用 Softmax 函数进行归一化处理:

$$a_i = \text{Softmax}(u_i) \quad (9)$$

(3) 将特征 H 和权重系数 a_i 进行加权求和,获取文本中每个词的重要性信息的文本向量 g^i :

$$g^i = \sum_{i=1}^n a_i H \quad (10)$$

2.4 输出层

输出层使用的激活函数为 *sigmoid* 函数,输出的内容为情感极性的预测分类结果, y_i 为第 i 条评论的情感极性预测:

$$y_i = \text{sigmoid}(w_i \times g^i) + b_i \quad (11)$$

其中, w_i 表示训练权重矩阵; b_i 表示偏置项; y_i 为第 i 条评论的情感极性预测。

模型训练中的 *loss* 损失函数:

$$\text{loss} = -\frac{1}{n} \sum_{i=1}^n y_i \times \log \hat{y}_i + (1 - y_i) \times \log (1 - \hat{y}_i) \quad (12)$$

其中, n 表示输入数据的总数; y_i 为 0 或 1,表示数据的真实预测分类结果是否正确; \hat{y}_i 表示通过模型得到的预测结果。

3 实验与分析

3.1 实验数据集

本文从京东购物网站爬取商品牙膏的评论数据作为研究的初始数据集。爬取到的数据共有 103 850 条商品评论,包括商品 ID、用户名、评论文本和评分。本文的部分实验数据集见表 1。

表 1 商品评论数据集

Table 1 Product review data sheet

商品 ID	用户名	评论文本	评分
100026078644	菜 * * * 7	口感不错,很清新	4
100012545121	J * * * b	清洁效果很好	5
100032188115	乐 * * * i	泡沫细腻,口气清新	4
100023070757	U * * * h	一直用这款,防蛀牙	3
100026782820	料 * * * 丹	味道有点奇怪	2

在数据集中,用户的评分为 1~5 分。通过对评

论进行分析,发现评分为 1~2 分的数据表现为消极;评分为 3 分的数据表现为中性;评分为 4~5 分的数据表现为积极。为了方便情感分析的划分,本文将 5 个类别的数据集归纳为 3 个类别的数据集,将 1~2 分的数据归为 0 分,表示消极;将 3 分的数据归为 1 分,表示中立;将 4~5 分的数据归为 2 分,表示积极。

在实验中,模型采用 6:2:2 的比例对数据集进行划分,即将 94 436 条评论数据随机地分为 57 462 条作为训练集,验证集和测试集的评论数据为 18 487 条。

3.2 评估指标

本文采用准确率和 $F1 - Score$ 作为模型的评价指标

(1) 准确率 (*Accuracy*): 预测正确的样本数目占总样本的比例:

$$Accuracy = \frac{C}{N} \quad (13)$$

其中, C 表示预测正确的样本数目, N 表示总样本数目。

(2) $F1 - Score$: 是精确度 (P) 和召回率 (R) 的调和均值:

$$P = \frac{TP}{TP + FP} \quad (14)$$

$$R = \frac{TN}{TN + FN} \quad (15)$$

$$F1 - Score = \frac{2PR}{P + R} \quad (16)$$

其中, TP 表示将正向文本样本预测为正向样本; FN 表示将正向文本样本预测为负向样本; TN 表示将负向文本样本预测为负向样本; FP 表示将负向文本样本预测为正向样本。

3.3 对比实验模型

(1) CNN^[16]: 使用 Word2vec 工具获取文本词向量,通过 CNN 获取特征,完成情感分类任务;

(2) BiLSTM^[17]: 采用双向 LSTM 模型获取上下文语义特征,完成情感分类任务;

(3) CNN-BiLSTM^[18]: 使用 Word2vec 工具获取文本词向量,采用 CNN 和 BiLSTM 分别获取文本局部特征和上下文语义关系特征,将两种特征向量进行融合,最后通过 Softmax 层输出情感分类的结果;

(4) CNN - BiLSTM - Attention^[19]: 在 CNN 和 BiLSTM 的基础上引入了 Attention 机制,对于情感分类任务中的不同词语分配不同的权重,最后完成

情感分类。

3.4 实验结果与分析

通过对 94 436 条评论数据进行分析,该数据在各个模型中的分类结果见表 2。

表 2 各个模型的实验结果

Table 2 Experimental results of each model		
模型	准确率	$F1 - Score$
CNN	86.83	86.80
BiLSTM	87.51	87.82
CNN-BiLSTM	88.57	88.84
CNN-BiLSTM-Attention	91.12	91.34
BERT- CNN-BiLSTM-Attention	93.06	93.11

从表 2 可以看出,融合局部文本特征和上下文语义的 CNN-BiLSTM 模型与单独使用模型相比,提取文本特征有明显的改进,情感分类的效果也较好;通过在 CNN-BiLSTM 模型中加入 Attention 机制,与 CNN-BiLSTM 模型相比,引入 Attention 模型的准确率和 $F1 - Score$ 都有提升,说明在将重要词赋予更大的权重,情感分析模型的效果也有所提高。模型 CNN-BiLSTM-Attention 中采用 Word2vec 工具,模型 BERT- CNN-BiLSTM-Attention 中采用 BERT 模型,其分类效果比采用 Word2vec 工具好,说明 BERT 模型的动态词向量分类效果较好。本文提出的 BCBA 模型相比于 CNN-BiLSTM 模型准确率提升了大约 4.5%,相比于 CNN-BiLSTM-Attention 模型准确率提升了大约 2%,并且 BCBA 模型的 $F1 - Score$ 相比于其他模型也均有所提升。

综上所述,本文提出的 BCBA 情感分析模型能够有效增强情感分类的能力,提高分类的准确性。

4 结束语

本文采用 BERT 模型来获取词向量矩阵,解决了传统词向量无法应对商品评论的一词多义问题;通过 CNN 获取商品评论的局部信息, BiLSTM 获取商品评论的上下文语义信息,充分挖掘商品评论中的蕴含的深层含义,并使用注意力机制使得模型对于重要词语赋予高度关注,完成最终的情感分类任务。通过对比实验的结果可以看出,本文提出的 BCBA 情感分析模型的准确率和 $F1 - Score$ 与其他模型相比均有一定的提升,验证了本文提出的情感分析模型的可行性和有效性。

参考文献

[1] HIRSCHBERG J, MANNING C D. Advances in natural language

- processing[J]. *Science*, 2015, 349(6245): 261-266.
- [2] DAHOUDA M K, JOE I. A deep-learned embedding technique for categorical features encoding [J]. *IEEE Access*, 2021, 9: 114381-114391.
- [3] TANG H, ZHU H, WEI H, et al. Representation of semantic word embeddings based on SLDA and word2vec model [J]. *Chinese Journal of Electronics*, 2022, 31(5): 1-8.
- [4] WANKHADE M, RAO A C S, KULKARNI. A survey on sentiment analysis methods, applications, and challenges [J]. *Artificial Intelligence Review*, 2022, 55(7): 5731-5780.
- [5] SEO S, KIM C, KIM H, et al. Comparative study of deep learning-based sentiment classification [J]. *IEEE Access*, 2020, 8: 6861-6875.
- [6] 魏苏波, 张顺香, 朱广丽, 等. 基于正交投影的 BiLSTM-CNN 情感特征抽取方法 [J]. *南京师大学报(自然科学版)*, 2023, 46(1): 139-148.
- [7] DEVLIN J, CHANG M W, LEE K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding [J]. *arXiv preprint arXiv:1810.04805*, 2018.
- [8] ACHEAMPONG F A, NUNOO - MENSAH H, CHEN W. Transformer models for text-based emotion detection: A review of bert-based approaches [J]. *Artificial Intelligence Review*, 2021, 54(8): 5789-5829.
- [9] NOZZA D, BIANCHI F, HOVY D. What the [mask]? making sense of language-specific BERT models [J]. *arXiv preprint arXiv:2003.02912*, 2020.
- [10] 汤英杰, 刘媛华. 基于预训练模型融合深层特征词向量的中文文本分类 [J]. *上海理工大学学报*, 2023, 45(2): 189-197, 204.
- [11] LEE Y H, KIM D W, LIM M T. A two-level recurrent neural network language model based on the continuous Bag-of-Words model for sentence classification [J]. *International Journal on Artificial Intelligence Tools*, 2019, 28(1): 1950002.
- [12] ZHANG C, LIU X, BIS D. An analysis on the learning rules of the skip-gram model [C]//*Proceedings of 2019 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2019: 1-8.
- [13] TURNER P D, LITTMAN M L. Unsupervised learning of semantic orientation from a hundred-billion-word corpus [J]. *Artificial Intelligence*, 2002. DOI:10.48550/arXiv.cs/0212012.
- [14] PANG B, LEE L, VAITHYANATHAN S. Thumbs up? Sentiment classification using machine learning techniques [J]. *arXiv preprint cs/0205070*, 2002.
- [15] CAO G, MA Y, MENG X, et al. Emotion recognition based on CNN [C]//*Proceedings of 2019 Chinese Control Conference (CCC)*. IEEE, 2019: 8627-8630.
- [16] XU G, MENG Y, QIU X, et al. Sentiment analysis of comment texts based on BiLSTM [J]. *IEEE Access*, 2019, 7: 51522-51532.
- [17] 周泳东. 基于卷积神经网络的商品评论情感分析的研究 [D]. 南京: 南京邮电大学, 2019.
- [18] 汪林, 蒙祖强, 杨丽娜. 基于多级多尺度特征提取的 CNN-BiLSTM 模型的中文情感分析 [J]. *计算机科学*, 2023, 50(5): 248-254.
- [19] 杨秀璋, 郭明镇, 候红涛, 等. 融合情感词典的改进 BiLSTM-CNN+Attention 情感分类算法 [J]. *科学技术与工程*, 2022, 22(20): 8761-8770.