

文章编号: 2095-2163(2022)11-0186-06

中图分类号: TP183

文献标志码: A

基于 Bert-BiLSTM 的商品评论情感分析研究

徐鹏, 罗梓汛, 黄昕凯

(广东东软学院, 广东 佛山 528225)

摘要: 商品评论为商家的选品和用户的购买提供了重要的决策帮助。为了获得商品评论的情感特征,并在评论中捕捉更多的情感信息,本研究提出一种 Bert-BiLSTM 的情感分类模型,运用斯坦福情感分析数据集、亚马逊商品评论数据集做一个情感分类模型,该模型利用 Bert 嵌入层对句子进行分割并将其转换为词向量,然后将其传递到 BiLSTM 模型中,以获取评论文本中的属性和情感词;训练后的模型使用混淆矩阵作为评价指标,相比其他深度学习模型在最终结果上表现出明显优势。将训练好的模型对结果进行分类和预测,从而分析结果的情感,这就为用户与商户在购买商品或选品时提供了建议方向和引导情绪。

关键词: 文本分析; Bert; BiLSTM; 深度学习; NLP; 情感分析; 机器学习

Commodity review sentiment analysis based on Bert-BiLSTM model

XU Peng, LUO Zixun, HUANG Xinkai

(Neusoft Institute Guangdong, Foshan Guangdong 528225, China)

[Abstract] Commodity reviews provide important decision-making assistance for merchants' selection and users' purchase. In order to obtain the emotional characteristics of commodity reviews and capture more emotional information in sentences, this study proposes an emotion classification model of commodity reviews. By using Stanford sentiment analysis data set and Amazon products review dat set, the research uses the Bert embedding layer to segment sentences and convert them into word vectors, and then passes them into the BiLSTM model to obtain the attributes and emotional words in the review text. The trained model uses confusion matrix as the evaluation index, and has a better performance compared with other deep learning models. The trained model is used to classify and predict the results, so as to analyze the emotion of the results. The results could provide users and merchants with suggested directions and guided emtions during the process of purchasing goods.

[Key words] text analysis; Bert; BiLSTM; deep learning; NLP; sentiment analysis; machine learning

0 引言

在互联网迅速发展、大数据来临,以及中国电商行业竞争加剧的背景环境下,用户于电商平台上购买商品的行为规模呈指数级上升,评论数据也随之增多。通常情况下,商品的评论数据中,蕴含着用户观点态度、情感倾向以及个人的见解。因此挖掘出用户的评论数据中的深层意向至关重要。在此前提下,商品评价分析应运而生。

商品评价分析是指通过对用户评论的处理,分析用户对商品的关注程度和情感态度,为商户选品和用户购买提供一定的决策辅助^[1]。在本文 Bert 与 BiLSTM 的商品评论情感分析研究中,商品评论情感分析占据主要地位,重点旨在发现用户是否有购买该商品的意愿并起到推荐作用,引导更多用户进行购买及评论。传统的关键词提取算法和预训练模型不能很好地联系上下文的语境和词向量稀疏等问题,不能精准有效地标出关键词。针对此问题,本

文采用 Bert 模型对评论数据进行词向量处理,避免分词造成的歧义,同时 BiLSTM 能结合上下文的语境,使得模型更为精确,实现评论文本情感分析。

1 模型构建与研究

1.1 循环神经网络

循环神经网络(Recurrent Neural Network, RNN)能够挖掘出特征中语义信息,使深度学习模型在处理语言数据时运用更加广泛。循环神经网络在原有的神经网络基础上,增加了对隐藏层的循环结构,使隐藏层既可不受输入影响,还能接收上一时刻隐含层的影响^[2]。循环神经网络设计展开如图 1 所示。

1.2 LSTM

相比于 RNN, LSTM 引入了输入门 i 、遗忘门 f 、输出门 o 以及内部记忆单元 c , 通过门控状态来控制传输状态,记住需要记忆的信息,忘记不重要的信息;而不是和 RNN 一样,只能够做记忆上的叠加。LSTM 能有效解决传统 RNN 在处理时间序列长期

作者简介: 徐鹏(2001-),男,本科生,主要研究方向:数据科学与大数据技术;罗梓汛(2001-),男,本科生,主要研究方向:信息系统设计与开发;黄昕凯(2001-),男,本科生,主要研究方向:信息系统设计与开发。

收稿日期: 2022-05-12

哈尔滨工业大学主办 ◆ 专题设计与应用

依赖中的梯度消失和梯度爆炸的问题,且对很多需要“长期记忆的”任务来说,效果显著,误差较小^[3]。LSTM 单元结构如图 2 所示。

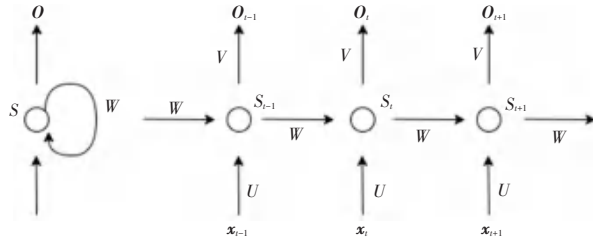


图 1 RNN 结构图

Fig. 1 RNN structure

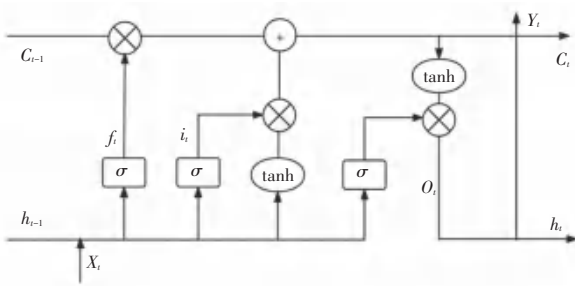


图 2 LSTM 单元结构

Fig. 2 LSTM unit structure

由图 2 可知, f_t 称为遗忘门, 表示 C_{t-1} 需要用作于计算 C_t 的特征; “ \otimes ”表示 LSTM 中最重要的门机制; W_f 是遗忘门的权重; b_f 是遗忘门的偏置。此处

需用到的数学公式可写为:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

对于输入门 i 来说, 可用于控制输入 x 和当前计算的状态更新到记忆单元的程度大小。相应的数学表示形式为:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (3)$$

i_t 能够控制 \tilde{C}_t 的部分特征, 用来更新 C_t , 与 f_t 相同, 可由如下公式进行计算:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (4)$$

对输出层来说, 其设计原理的数学表达式为:

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = O_t * \tanh(C_t) \quad (6)$$

上述所有表达式中, σ 通常是指 Sigmoid 函数, 主要起到门控作用, 其输出为 0~1, 当输出接近 0 或 1 时, 才能符合物理意义上的开或关。

1.3 BiLSTM

BiLSTM 是对 LSTM 的改进方案。该方案有效解决了 LSTM 无法学习反向特征的问题, 因此本文采用双向 BiLSTM 模型用于情感分类, 一个用于保存上一词语的前后文, 另一个用于存储下一词语的前后文^[4]。BiLSTM 结构如图 3 所示。

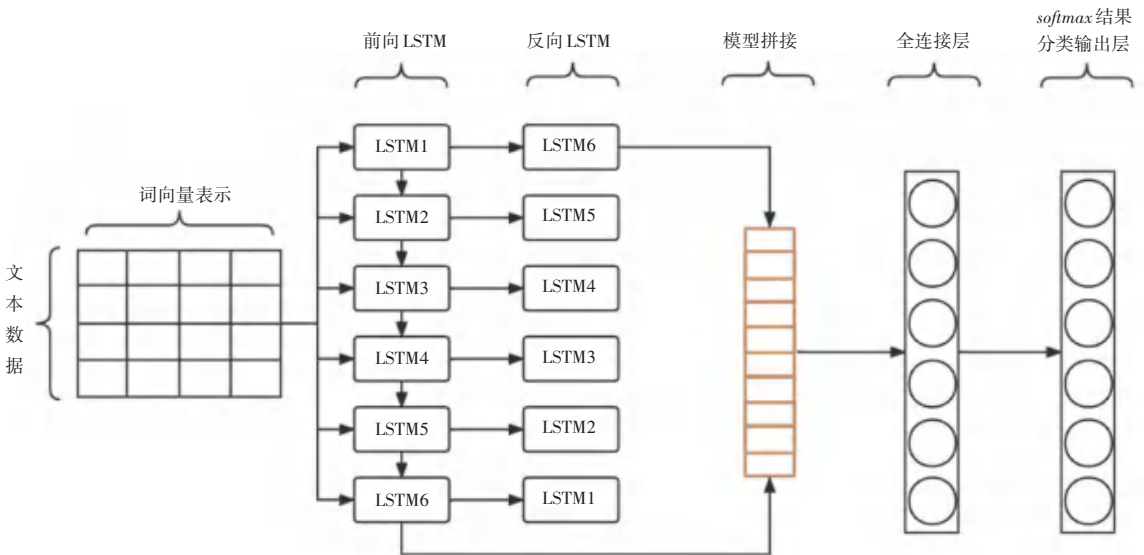


图 3 BiLSTM 结构图

Fig. 3 BiLSTM structure

1.4 Bert

Bert 是一个预训练的语言模型, 重点强调了不再采用传统的单向语言模型, 或者把 2 个单向语言模型通过浅层拼接的方法来进行预训练, 而是通过使用遮挡语言模型 (MLM) 生成深度的双向语言表征^[5]。

Bert 的提出为自然处理领域带来明显提升。前

期的模型是单向训练, 并将 2 个单向训练相结合, 而 Bert 则为了将句子转化为词向量而使用了多层 Transformer^[6], 对语境的分析相比单向模型来说会更加透彻^[2]。Bert-base 的 Encoder 是由 12 层结构相同的 Transformer Encoder 结构堆叠而成。虽然结构上是相同的, 但相互间的权重并不共享^[7]。Bert

Transformer Encoder 结构如图 4 所示。

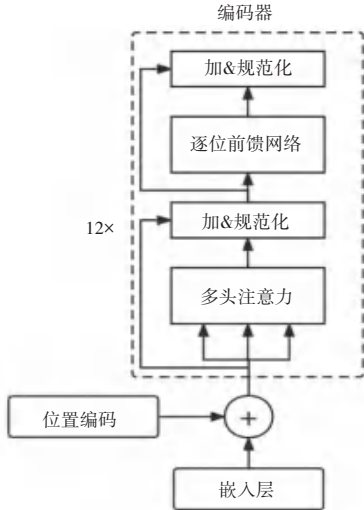


图 4 Bert Transformer Encoder 结构

Fig. 4 Bert Transformer Encoder structure

Bert 的模型架构上采用了 Transformer 的

encoder 部分,输入由字嵌入 (token embedding)、段嵌入 (segment embedding) 和位置嵌入 (position embedding) 三部分相加构成,每个输入的起始 token 会固定设置为 'CLS',用于下游的分类任务,2 个不同的 sentence 间会加入 'SEP' 作为分隔,输入的尾部同样会加入一个 'SEP'^[8]。嵌入层架构如图 5 所示。

研究中,由 MLM 和下一句预测 (NSP) 来进行 Bert 的模型训练。MLM 对输入中的部分词语进行随机选取替换,致力于通过训练能够正确预测出原始输入的替换词,进而联合双向上下文达到双向编码的效果^[9]。NSP 能够预测 2 个句子是否连在一起,用于挖掘句子关系。在预处理时会以 50% 的概率从其他文档中随机选取首个 text 的下一个 text,在预训练中预测其后的 text 是否为前一个 text 的真实上下文,即可整理得到句子的逻辑关系^[10]。

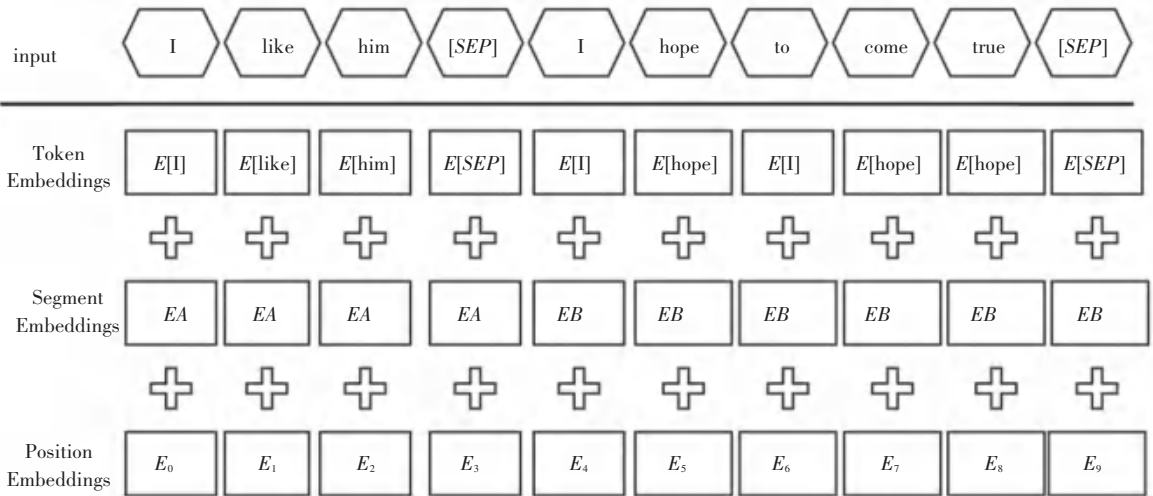


图 5 Bert 嵌入层架构

Fig. 5 Bert embedded layer structure

注意力机制如图 6 所示。运算结构为一层 Decoder 与一层 Encoder 对应。在 Encoder 中,输入经过 Embedding 后,需进行位置嵌入 (Positional encoding),再经过 Multi-Head Attention,最后是全连接层^[11]。

1.5 Bert-BiLSTM 模型构建

本文采用 Bert 的嵌入层将词序列抽取输出 embedding 序列、再转换成词向量形式,同时将 Bert 输出的词向量经过 BiLSTM 进一步再做特征提取,将其传送到输出层。输出层由全连接层和 softmax 层构成,在全连接层调整特征向量的维数,并使用

softmax 分类器对评论文本进行分类,实现情感分析。模型设计架构如图 7 所示。

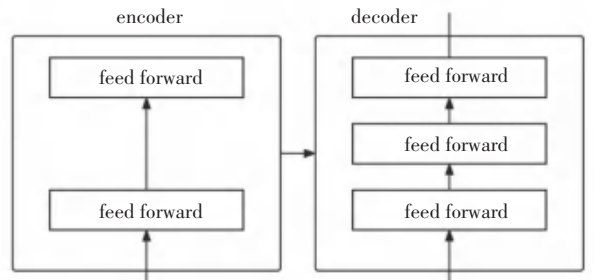


图 6 Attention 运算

Fig. 6 Attention operation

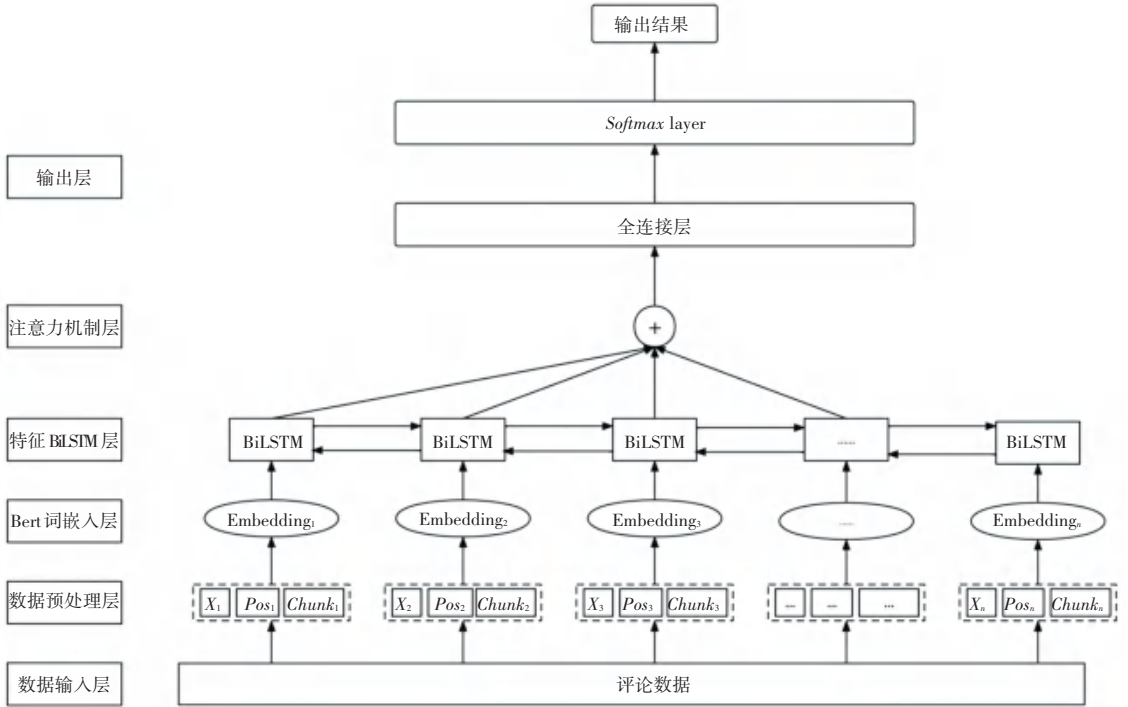


图 7 Bert-BiLSTM 模型架构

Fig. 7 Bert-BiLSTM model structure

在 BiLSTM 的输出层引入 Multi-Head Attention 注意力机制, 词向量在 BiLSTM 层产生的输出向量 h_i 进入 Multi-Head Attention 层 $q \in R^{d_q}, k \in R^{d_k}, v \in R^{d_v}$ 输出, 提高文本情感信息的利用反馈。计算公式见如下:

$$h_i = f(W_i^{(q)} q, W_i^{(k)} k, W_i^{(v)} v) \in R^{p_o} \quad (7)$$

其中, 参数包括 $W_i^{(q)} \in R^{p_q \times d_q}, W_i^{(k)} \in R^{p_k \times d_k}$ 和 $W_i^{(v)} \in R^{p_v \times d_v}$, 以及池化注意力 f 表示可加性注意力。Multi-Head Attention 输出需经一个线性转换, 对应的是拼接后 h 个头的结果。对此可表示为:

$$W_o [h_1 \cdots h_h] \in R^{p_o}$$

2 实验与分析

2.1 实验数据准备

实验使用的商品评论数据集为 Julian McAuley, UCSD 整理的 Amazon product data-Clothing, Shoes and Jewelry 数据集, 总共 278 677 条数据, 数据未带情感标注, 实验抽取 10 000 条评论作为实验数据, 并对文本内容以自然语言处理 (SNOWNLP) 来做分类辅助的人工标注, 标注结果可分为 *positive*、*neutral*、*negative*, 选取数据集 70% 作为训练集, 20% 作为测试集, 10% 的商品评论作为校验集。实验部分数据示例见表 1。

表 1 实验数据示例

Tab. 1 Examples of experimental data

文本示例	倾向性
Perfect red tutu for the price. I bought it as part of my daughters Halloween costume and it looked great on her.	<i>positive</i>
We bought several tutus at once, and they are got high reviews. Sturdy and seemingly well-made. The girls have been wearing them regularly, including out to play, and the tutus have stood up well. Fits the 3-yr old & the 5-yr old well. Clearly plenty of room to grow. Only con is that when the kids pull off the tutus, the waste band gets twisted, and an adult has to un-tangle. But this is not difficult.	<i>negative</i>
This was a really cute tutu the only problem is that it was super short on my 5 yr old daughter. Other than that it was really adorable	<i>positive</i>
Very cute, shorter than I was expecting, but then again, I didn't bother to calculate the measurements before ordering it. It has three layers of tulle, so it's not as full as the picture looks if stretched fully, but still a full skirt. I bought this skirt for my six year-old niece, so I highly doubt the skirt will ever be stretched to its full capacity on her. In fact, the skirt stretches pretty far but is also retractable so it can fit a variety of sizes	<i>neutral</i>
This is a great tutu and at a really great price. It doesn't look cheap at all. I'm so glad I looked on Amazon and found such an affordable tutu that isn't made poorly. A++	<i>positive</i>

为确保数据有效性,需对文本数据进行预处理。文本预处理的过程为:去除多余无用符号,对文本单词进行纠错处理,通过自定义停用词表将文本数据中的无意义单词进行剔除,使用正则表达式把文本中特殊符号删除,再使用 Sentence BERT 提取到句子主干,索引长度实现标准化,从而避免了句子过长无法训练问题。评论数据预处理结果示意图表 2。

表 2 评论数据预处理示意图

Tab. 2 Comment data preprocessing schematic

评论	预处理结果
Very cute and perfect for my 6 yr old daughter. She wore it to a princess birthday party. Lots of fun.	cute and perfect daughter wore it to party lots of fun
My 5 year old daughter get this today for Christmas. She and we are very pleased with it. Thank you so much!	daughter get this pleased Christmas Thank much

2.2 实验参数

本文模型 Bert-BiLSTM 参数设置如下。输入层采用预训练模型 BERT-Base-uncased,该模型采用 12 层 Transformer,隐层维度为 768, Multi-Head-Attention 的参数为 12,模型总参数大小为 110 MB。特征提取层主要由 BiLSTM 构成。

模型训练方面,设置批次大小为 64,最大序列长度为 512,隐藏层个数为 13, epoch 为 4, batch size 设定为 256, 优化器选用 LAMB, 防止过拟合的 dropout 率为 0.5。

2.3 实验评价指标

本文的研究内容是预测商品评论数据是否为消极或积极,是自然语言处理中常见的分类任务。若预测结果为积极,标记为 1,否则标记为 0。预测结果的混淆矩阵见表 3。

表 3 混淆矩阵

Tab. 3 Confusion matrix

预测结果	实际结果	
	1	0
1	TP	FP
0	FN	TN

表 3 中, TP 表示预测结果为积极,实际评论也为积极; FP 表示预测结果为积极,而实际为消极; FN 表示预测结果为消极,而实际为积极; TN 表示预测结果为消极,实际也为消极。接下来,对于研究中选用的评价指标,拟展开探讨分述如下。

(1) 召回率。表示模型实际为 1 的样本,预测仍为 1 的样本概率,其计算公式为:

$$R = \frac{TP}{TP + FN} \quad (8)$$

(2) 精准率。是指在所有预测为 1 的样本中,实际有多少个样本为 1,其计算公式为:

$$P = \frac{TP}{TP + FP} \quad (9)$$

(3) F_1 值。是对召回率和精准率的综合评价指标,是对其进行加权平均的结果,其计算公式为:

$$F_1 = \frac{2PR}{P + R} \quad (10)$$

2.4 实验结果分析

本文从准确率 (Precision score)、召回率 (Recall score)、 F_1 (F_1 -score) 三个方面作为评价指标。其中, Precision 是判别模型对负样本的区分能力; Recall 是模型对正样本的识别能力; F_1 是将 Precision 与 Recall 相结合的综合值,这 2 个评价指标的结合可以更加全面地反映分类性能。特别地,利用 F_1 值来评估分类器性能时,分类器的性能越好, F_1 值越接近于 1,因此本文选其作为衡量实验效果的主要评价指标。

在对比实验上,本文选取了 Bert、BiLSTM、W2V-SVM 模型进行训练与结果比较。具体描述如下。

(1) Bert: 使用预处理模型 BERT-base-uncased,参数与本文 Bert 预设参数保持一致,利用预训练文本特征后输入 Bert 情感分类。

(2) BiLSTM: 定义参数大小与 2 层双向 LSTM 的模型结构一样,并使用全连接层,再经过 SoftMax 分类器输出情感分类结果。

(3) W2V-SVM: 使用 SNOWNLP 进行分词,并做数据转化,接着将 Word2Vec 模型词表初始化,再将各个词向量用平均的方式生成整句对应的向量。用矩阵进行建模与转化,拟合 SVM 模型,并使用本文准备的文本评论数据进行训练。

对比模型的实验结果见表 4。由表 4 的整体结果分析可知, Bert-BiLSTM 模型精度与 Bert、BiLSTM 和 W2V-SVM 精度相比分别提高了 3.32%、9.1% 和 4.35%。Bert-BiLSTM 的双向神经网络相对于传统文本分类模型在语境上有较大提升,进行情感分析可以获得更好的结果。表 5 为示例分析结果展示。

表 4 对比试验模型比较

Tab. 4 Comparative test model comparison

Model	Precision	Recall	F_1	%
Bert	82.37	77.65	81.42	
BiLSTM	76.56	75.73	76.29	
W2V-SVM	81.34	68.83	74.62	
Bert-BiLSTM	85.69	84.21	84.67	

表 5 示例分析结果

Tab. 5 Samples analysis results

原文本	倾向性	文本处理	结果
My 3-year old granddaughter loves it! I also got her a ballerina costume for her Bitty Baby. Definitely a big hit!	<i>positive</i>	granddaughter loves it I got a ballerina for Bitty Baby Definitely hit	<i>positive</i>
put it in the washer bc like a normal kid my daughter got it dirty. When i took it out it was all ripped up at the ends. she was so uppset when i got rid of it for the price i paid I thought it would be a little better	<i>negative</i>	daughter got it dirty I took it out it was all ripped up she was upset I got rid of it I paid it would be a little better	<i>negative</i>
Said they would fit a certain size but when we got them they are a lot smaller then as described	<i>neutral</i>	they fit a certain size we got them they are a lot smaller	<i>neutral</i>

3 结束语

本文对基于 Bert 和 BiLSTM 的情感分析进行了研究。实验结果表明, Bert 将句子转化为词向量, 输入到 BiLSTM 模型中, 由于 BiLSTM 能兼顾上下文的语境, 提高句意的情感丰富度, 有效提升了文本分类的准确度, 得到较为优化的分类器。经过对比实验后, 所得结果均要优于传统的机器学习模型 SVM 以及 W2V-SVM、Bert、BiLSTM 等深度学习模型。只是本文研究仍有一定不足, 如某些评论数据存在歧义, 过分积极化实则隐藏为消极, 在后续研究中需要对此类数据进行特殊处理, 并且加大评论数据集, 进而提升模型整体效果, 增强泛化能力。

参考文献

- [1] 康月, 薛惠珍, 华斌. 面向深度学习网络的细粒度商品评价分析[J]. 计算机工程与应用, 2021, 57(11): 140-147.
 [2] 李博, 李洪莲, 关青, 等. 基于 CNN-BiLSTM-HAN 混合神经网络

的高校图书馆社交网络平台细粒度情感分析[J]. 农业图书情报学报, 2022, 34(04): 63-73.

- [3] 刘继, 顾凤云. 基于 BERT 与 BiLSTM 混合方法的网络舆情非平衡文本情感分析[J]. 情报杂志, 2022, 41(04): 104-110.
 [4] 李洋, 董红斌. 基于 CNN 和 BiLSTM 网络特征融合的文本情感分析[J]. 计算机应用, 2018, 38(11): 3075-3080.
 [5] 赵宏, 傅兆阳, 赵凡. 基于 BERT 和层次化 Attention 的微博情感分析研究[J]. 计算机工程与应用, 2022, 58(05): 156-162.
 [6] DEVLIN J, CHANG Mingwei, LEE K, et al. BERT: Pre-training of deep bidirectional transformers for language understanding[J]. arXiv preprint arXiv: 1810.04805, 2018.
 [7] 史振杰, 董兆伟, 庞超逸, 等. 基于 BERT-CNN 的电商评论情感分析[J]. 智能计算机与应用, 2020, 10(02): 7-11.
 [8] 胡荣磊, 芮璐, 齐筱, 等. 基于循环神经网络和注意力模型的文本情感分析[J]. 计算机应用研究, 2019, 36(11): 3282-3285.
 [9] 赵响, 张智雄, 刘欢等. 基于 BERT 模型的中文医学文献分类研究[J]. 数据分析与知识发现, 2020, 4(08): 41-49.
 [10] 杨杰, 杨文军. 基于 BERT 模型的文本评论情感分析[J]. 天津理工大学学报, 2021, 37(02): 12-16.
 [11] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. CoRR abs/1706.03762, 2017.
 [12] 冯兴杰, 张志伟, 史金钊. 基于卷积神经网络和注意力模型的文本情感分析[J]. 计算机应用研究, 2018, 35(05): 1434-1436.

(上接第 185 页)

- [8] CUI Laizhong, XU Chong, LI Genghui, et al. A high accurate localization algorithm with DV-Hop and differential evolution for wireless sensor network[J]. Applied Soft Computing, 2018, 68: 39-52.
 [9] RAMEZANI M, BAHMANYAR D, RAZMJOOY N. A new improved model of marine predator algorithm for optimization problems[J]. Arabian Journal for Science and Engineering, 2021.
 [10] FARAMARZI A, HEIDARINEJAD M, MIRJALILI S, et al. Marine predators algorithm: A nature-inspired metaheuristic[J]. Expert Systems with Applications, 2020, 152: 113377.
 [11] 赵欣. 不同一维混沌映射的优化性能比较研究[J]. 计算机应用

研究, 2012, 29(03): 913-915.

- [12] 于建芳, 刘升, 韩斐斐, 等. 基于柯西变异的蚁狮优化算法[J]. 微电子学与计算机, 2019, 36(06): 45-49.
 [13] 毛清华, 张强. 融合柯西变异和反向学习的改进麻雀算法[J]. 计算机科学与探索, 2021, 15(06): 1155-1164.
 [14] YANG Xiaojing, JIAO Qingju, LIU Xinke. Center Particle Swarm Optimization algorithm[J]. Proceedings of 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC 2019). Chengdu, China: IEEE, 2019: 2084-2087.