

Research on Chinese Microblog Sentiment Classification Based on TextCNN-BiLSTM Model

Haiqin Tang and Ruirui Zhang*

Abstract

Currently, most sentiment classification models on microblogging platforms analyze sentence parts of speech and emoticons without comprehending users' emotional inclinations and grasping moral nuances. This study proposes a hybrid sentiment analysis model. Given the distinct nature of microblog comments, the model employs a combined stop-word list and word2vec for word vectorization. To mitigate local information loss, the TextCNN model, devoid of pooling layers, is employed for local feature extraction, while BiLSTM is utilized for contextual feature extraction in deep learning. Subsequently, microblog comment sentiments are categorized using a classification layer. Given the binary classification task at the output layer and the numerous hidden layers within BiLSTM, the Tanh activation function is adopted in this model. Experimental findings demonstrate that the enhanced TextCNN-BiLSTM model attains a precision of 94.75%. This represents a 1.21%, 1.25%, and 1.25% enhancement in precision, recall, and F1 values, respectively, in comparison to the individual deep learning models TextCNN. Furthermore, it outperforms BiLSTM by 0.78%, 0.9%, and 0.9% in precision, recall, and F1 values.

Keywords

Chinese Microblog Review, Deep Learning, Sentiment Classification, TextCNN-BiLSTM

1. Introduction

In the era of networked media, a growing number of individuals have turned to online platforms to gather information, voice their opinions, and express their emotions [1]. As a widely-used social medium disseminating copious real-time information, analyzing the emotional content of microblog comments holds significant practical importance for platform management and public opinion regulation [1]. Sentiment analysis primarily involves scrutinizing the content of generated text to discern its emotional polarity: positive or negative [2]. There are three principal approaches to sentiment analysis: those based on emotion dictionaries [3], machine learning [4], and deep learning [5] for emotional analysis.

The emotion dictionary-based approach necessitates the construction of an emotion lexicon. HowNet stands as the most prevalent Chinese emotion dictionary [6]. This lexicon constitutes a crucial emotional resource for sentiment analysis, applicable to tasks of varying granularity such as words, phrases, and attribute sentences [7]. Leveraging HowNet and SentiWordNet dictionaries, Zhou et al. [8] deconstructed Chinese words, calculated their emotional polarity, established the SLHS Chinese dictionary, and utilized

※ This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Manuscript received April 28, 2022; first revision October 17, 2022; accepted December 11, 2022.

* Corresponding Author: Ruirui Zhang (zhangruiruisw@163.com)

School of Business and Tourism, Sichuan Agricultural University, Chengdu, China (1696562280@qq.com, zhangruiruisw@163.com)

the SVW classifier to analyze emotions in microblog texts, achieving an accuracy of 83.84%. Addressing issues of scale and colloquial word scarcity, Zhao et al. [9] constructed a 100,000-scale emotion lexicon based on extensive microblog data, resulting in a 1.13% improvement in microblog emotion classification performance. Jiang et al. [10] addressed unknown emotional words in text content by selecting HowNet as the seed and constructing a domain-specific emotion dictionary using pointwise mutual information (PMI) and Word2vec algorithms, yielding a substantial enhancement in dictionary accuracy compared to others.

Machine learning models have progressively found application in sentiment classification. Pang et al. [11] pioneered the incorporation of machine learning into emotional analysis. Through experimental comparisons on film reviews, support vector machines emerged with the highest efficacy, achieving an accuracy of 82.92%. Li et al. [12] introduced a multi-label maximum entropy-based machine learning model, applying it to datasets comprising Twitter, microblog, and other comments, achieving an accuracy of 86.06%. Kaur et al. [13] leveraged N-grams for feature extraction and, in conjunction with the k-nearest neighbor classification algorithm, raised accuracy to 82%. Zhang et al. [14] integrated the emotional dictionary with a machine learning model, augmenting emotional characteristics with negative words in the health domain, and subsequently applied the model to classify negative topics on microblogs, yielding an accuracy of 74.1%.

Advancements in technology have facilitated the integration of deep learning into natural language processing (NLP)-based emotion analysis. In cases of intricate content, hybrid models prove superior to singular models for certain sentiment classification tasks. Sun et al. [15] harnessed the GloVe model for word vector training, employing the bidirectional gated recurrent unit (BiGRU) for contextual feature extraction, and incorporated an attention mechanism to achieve emotion classification. This model exhibited an accuracy of 91.21% on the IMDB dataset. Zhao et al. [16] introduced the bidirectional long short-term memory (BiLSTM) and a convolutional neural network (CNN) serial hybrid network to comprehend lexeme information in texts, achieving an accuracy of 91.31% on a dataset encompassing review texts from six fields. Miao et al. [17] argued that the CNN-BiGRU model effectively captured static and sequential features through CNN and bidirectional GRU, thereby simplifying the feature extraction process, reducing dimensionality, and enhancing accuracy and training efficiency. Yang et al. [18] leveraged GloVe for preprocessing, incorporating an attention mechanism into BiGRU to extract critical information from texts, and further extracted features through the forward attention mechanism, resulting in an augmented effectiveness in sentiment classification. In response to the challenge of embedding models struggling to incorporate word sentiment information characteristics, Yan et al. [19] proposed the parallel CNN and BiLSTM-attention model for evaluating JD e-commerce review datasets, substantially enhancing overall effectiveness. Fan and Li [20] generated character and word vectors using FastText, employing the GRU neural network for sentiment classification, leading to a 92% improvement in sentiment classification accuracy for networked short texts. Yang et al. [21] addressed the issue of neglecting the potency of sentiment feature words and introduced an enhanced BiLSTM-CNN+Attention model based on a comprehensive emotional dictionary, proficiently extracting semantic features and elevating accuracy. Shen et al. [22] introduced a model integrating knowledge-enhanced semantic representation with a dual attention mechanism, amalgamating contextual and sentiment features to enhance sentiment analysis accuracy. Ali Al-Abyadh et al. [23] conducted a comparative study between single models and various hybrid deep learning models, revealing that the hybrid deep learning-support

vector machine model elevated sentiment analysis accuracy by 91.3%. Hu et al. [24] integrated a multi-layer attention mechanism, combined with BiGRU and multi-granularity convolutional neural network, achieving a sentiment analysis accuracy of 92.75% in the hotel review dataset. Khan et al. [25] applied a combination of machine learning models and deep learning hybrid modes to scrutinize the sentiment of Urdu language reviews. The experiments demonstrated that models utilizing Bert pre-trained word embedding exhibited exceptional effectiveness. Wu et al. [26] engineered a sentiment classifier based on word embedding and lexical polarity analysis, implementing a two-level long short-term memory network, effectively resolving the challenge of applying the model to high-quality training sets characterized by high label accuracy.

The method founded on emotion dictionaries only considers the semantics of individual words, neglecting contextual semantic information. Additionally, it demands substantial time investment for constructing an emotion lexicon, imposing inherent limitations. Machine learning methods hinge on human annotation and struggle to discern deeper semantics. Although deep learning models excel in sentiment classification, there remains a need for refinement in efficiently eliminating special symbols and accurately comprehending contextual semantic information, particularly in the face of complex microblog comments, regardless of whether a singular or hybrid model is employed. In this study, a TextCNN-BiLSTM hybrid model is employed to comprehensively extract both local and contextual features, thereby enhancing the efficiency of sentiment classification for microblog comments.

2. Theoretical Basis

2.1 Word2vec Vectorization

Initially, sentiment analysis necessitates the vectorization of the target comment texts. Common vectorization methods encompass One-hot, Word2vec, Glove, and BERT. Word2vec constructs word embeddings through skip-grams or continuous bag-of-words (CBOW). CBOW determines the central word based on several consecutive words preceding and following it, enabling it to predict the target word in context. In contrast, Skip-gram predicts its context from the central word. This article employs the CBOW model for word embedding.

2.2 TextCNN Network Structure

CNN, a feedforward neural network with a convolution structure, comprises convolution and pooling layers, proficient at extracting local features for text classification [27]. TextCNN represents a variant of CNN that can simultaneously employ filters of varying sizes to extract features of different dimensions from the text, thereby obtaining representative features. The TextCNN model consists of convolution, pooling, and classification layers. The schematic representation of the TextCNN structure is illustrated in Fig. 1.

At the heart of TextCNN lies the convolutional layer, responsible for acquiring diverse text features through multiple convolutional kernels. The calculation formula is as follows:

$$h_i = f(w_{i(x,y)} * c_{(x,y)} + b_i), \quad (1)$$

where f denotes the activation functions, encompassing Tanh, ReLU, sigmoid, and others. $w_{i(x,y)}$ signifies the weight of the filter input node corresponding to the i node in the output matrix; (x, y) denotes the value of the node $c_{(x,y)}$ in the filter; b_i represents the bias corresponding to the i node. Local feature extraction is achieved by employing three filters with convolution kernel sizes of 1, 3, and 5, culminating in the final result of the convolution layer h_i .

The pooling layer serves to reduce the dimensionality of features post-convolution and guards against overfitting. Given that BiLSTM must be integrated after TextCNN, retaining the spatial information of the text becomes imperative. Since pooling would lead to the loss of this information, it is omitted in this study.

The classification layer amalgamates the features extracted from the pooling layer into a composite vector. Subsequently, this vector undergoes classification using the softmax classifier to accomplish emotion classification, with dropout implemented as a preventive measure against overfitting.

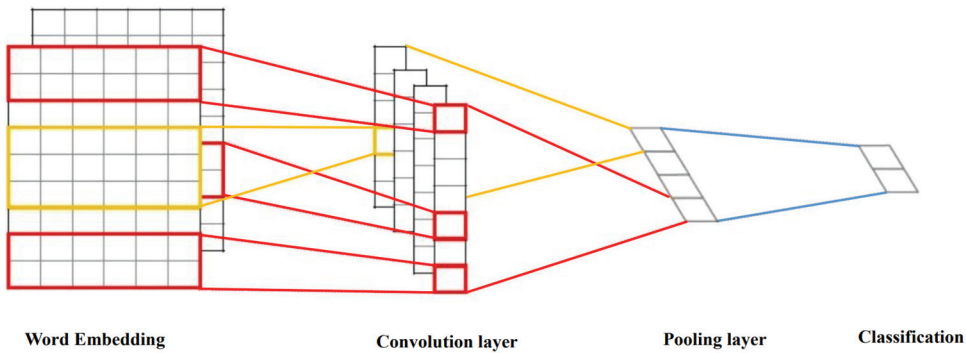


Fig. 1. Text convolutional neural network.

2.3 BiLSTM Network Structure

LSTM network, a type of recurrent neural network [28], establishes a crucial link between texts and sentences due to the diversification and complexity inherent in Chinese texts. Microblog comments encapsulate users’ sentiments and perspectives expressed in varying tenses. Hence, to gain a comprehensive contextual understanding, employing bidirectional LSTM becomes imperative.

BiLSTM comprises two LSTM networks operating in reverse, jointly determining the output of the entire network. It remedies the limitation of LSTM’s inability to encode information in a bidirectional manner. The internal structure of BiLSTM is depicted in Fig. 2.

f_t represents the forgetting gate. This calculation allows us to choose which information from the upper layer should be forgotten. The formula for the forgetting gate is as follows: σ denotes the activation function, W_f signifies the weight of the forgetting gate, and b_f represents the offset of the forgetting gate.

$$f_t = \sigma(W_f * [h_{t-1}, x_t] + b_f). \tag{2}$$

i_t represents the input gate. The gated state’s input includes the previous moment’s hidden layer h_{t-1} and the current input word x_t . At this stage, there exists a temporary cell state \tilde{c}_t . The calculation formulas for i_t and c_t are as follows: W_i represents the weight of the input gate, b_i denotes the offset of the input gate, W_c signifies the weight of cell information, and b_c represents the offset of cell information.

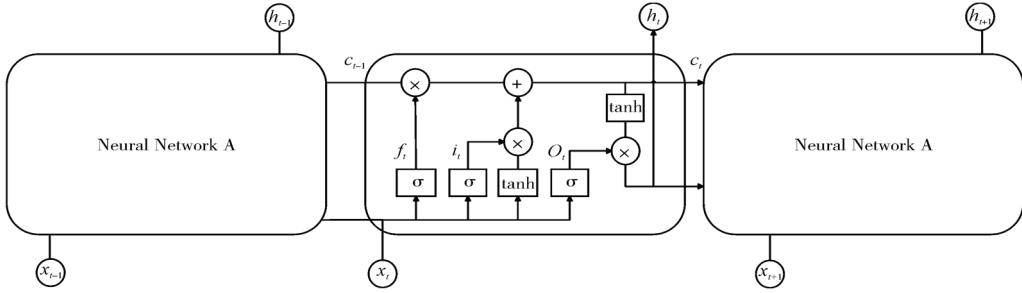


Fig. 2. The internal structure of bidirectional long short-term memory model.

$$i_t = \sigma(W_i * [h_{t-1}, x_t] + b_i), \quad (3)$$

$$\tilde{c}_t = \tanh(W_c * [h_{t-1}, x_t] + b_c). \quad (4)$$

Following the application of the forgetting gate and the input gate, the cell state from the last moment, c_{t-1} , undergoes an update to yield the new cell state, c_t . This can be calculated using the following formula:

$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t. \quad (5)$$

o_t stands for the output gate, governing the output information of the network structure. The final output is represented by h_t . The formulas for h_t and o_t are as follows: W_o denotes the weight of the output gate, and b_o signifies the offset of the output gate.

$$o_t = \sigma(W_o * [h_{t-1}, x_t] + b_o), \quad (6)$$

$$h_t = o_t * \tanh(c_t). \quad (7)$$

3. Network Structure Design

3.1 Research Process

The research process in this article unfolds as follows: initially, comment data is inputted. We employ the Jieba Word segmentation tool to process the inputted comments, forming word vectors through Word2vec. Subsequently, the results of word vectorization are fed into the TextCNN model for feature extraction, while BiLSTM is employed for context feature extraction. Finally, sentiment polarity classification is accomplished, categorizing the polarity of each sentence comment into two classes. The sentiment analysis research process is illustrated in Fig. 3.

3.2 TextCNN-BiLSTM Network Structure

Firstly, the preprocessed microblog comments are trained into a word vector using the CBOW in the Word2vec model, which serves as the input for the hybrid model. Next, TextCNN, without a pooling layer, is utilized to extract local features, while BiLSTM is employed to capture global features reflecting contextual information. Ultimately, a softmax layer is employed for the sentiment classification of microblog comments. The structure of the hybrid model is depicted in Fig. 4.

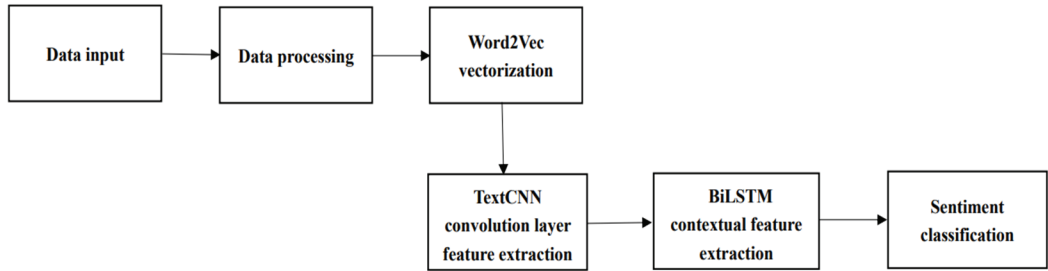


Fig. 3. The research process.

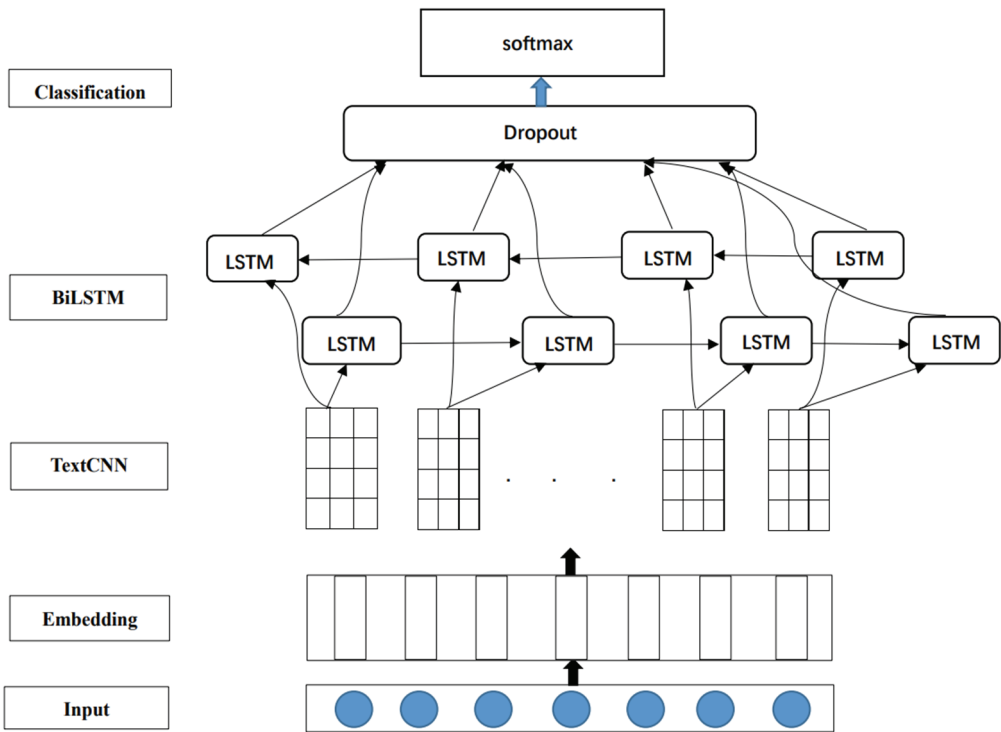


Fig. 4. Model structure diagram.

4. Experimental Results and Analysis

4.1 Experimental Data

The dataset utilized in this article is derived from Weibo_senti_100K, comprising over 100,000 microblog comments annotated with emotions. Weibo_senti_100K encompasses 50,000 positive and 50,000 negative comments, designated as 1 and 0, respectively. For the experiment, 10,000 data samples were selected from both positive and negative sentiment datasets, resulting in a total of 20,000 microblog comment data pieces for the model experiment. The proportion of training, testing, and validation datasets was set at 8:1:1. Sample comments are illustrated in Figs. 5 and 6.

label	review
0	1 早上好汪星人//@奥迪张强:基本上[挖鼻屎]可以判断是四驱的//@湖北中基奥迪:[哈哈],各...
1	1 这锅到你家把作用最大化啦[鼓掌][鼓掌][鼓掌]
2	1 [呵呵] //@厨缘觅友:[good][good][good][赞][赞][赞][鼓掌][鼓掌]...
3	1 找妹哦[爱你][哈哈][嘻嘻][呵呵] //@贺宁_:哈哈,得瑟啊,不过这个职位是不错,推荐!
4	1 父亲节送了老爸一块菜地,这个主意不错吧[兔子],现实版开心农场,豆角,?,?,玉米。。不少...
5	1 [哈哈]全国土地利用数据的制图综合。随着数据库的兴起,电子地图的普及,涉及到地图学、美学、地...
6	1 惭愧。我除了中小学上过写字课临过帖,以后就没有临帖。我的钢笔字楷书照印刷楷体练,行草练的是黄...
7	1 [不要]//@知心山水:不应过于苛责李、梦!孩子涉罪,此刻父母最揪心——可怜天下父母心——...
8	1 霸气//@马逸鉴:[哈哈]有新意。//@云南昆明毕胜
9	1 我靠!中间那个男人也能当世界小姐~[吃惊] //@石破天001:中间那位表情也不错 //@蜜...
10	1 活动启程,开启鸡血模式,准备好了哦!@云朗带你游台湾 @中华航空公司 @去哪儿网酒店控 [嘻...]
11	1 同问[哈哈] //@Lio李勇:余总,起床了吗?[呵呵]
12	1 【“春打六九头”:2月4日6时3分,立春】农历正月初五6时3分将迎来马年的第一个节气——“立...
13	1 有没有觉得我最可爱啊?[挤眼][害羞][偷笑][太开心]
14	1 笑一笑十年少~[哈哈]

Fig. 5. Examples of positive comments.

label	review
11100	0 这孩子太悲剧了,是大家救了你啊//@刘炎迅:这孩子太年轻了 //@作业本:那么多爱马仕只有一...
11101	0 直接枪毙//@美女时钟:畜生,不得不爆粗[怒]
11102	0 居然还有几位女性。先性别筛选一下再人肉是不是好些?这软件太雷人了。 //@中青报曹林:[抓狂]...
11103	0 十年,我在支付宝花了很多钱...说多都是泪[泪][泪][泪][泪] #十年记忆 支付时光#
11104	0 不断加我的小朋友们,我真的是郭涛,但我真的不是石头的爸爸.....[晕]
11105	0 帝都的空气质量真心不好!这才几点啊,天已经灰蒙蒙的,需要开灯了.....[悲伤]
11106	0 扎了2针,抽了5管血化验。然后一个月后还要被从喉咙插根管子进行检查。同学们,爱惜自己的胃吧。[泪]
11107	0 [吃惊][吃惊]//@赵波-赵校:回复@reagandady:孔雀是金巧巧。卧槽我怎能八卦全...
11108	0 烘焙展不好玩,还是有的吃才好!下午茶时间在夏朵花园补午午餐,5个人点了披萨、海鲜饭、龙...
11109	0 又到了吃货们懊恼“假期多少天,长胖多少斤”的时候鸟[抓狂][抓狂],谁叫你们错把他两贴在一起...
11110	0 代小儿给老父上柱香、烧香纸。好多话没来得及说,好多心情来不及表达。就是个慢热的人,亲人离开一...
11111	0 //@望月者:书的封面出来了么?可我在伊朗啥也看不见[抓狂][抓狂]
11112	0 //@三石出版痴语:听罢,泪如雨下!必须听必须听! //@朱学东:[蜡烛] //@吴晓波:...
11113	0 要不要?麽?? [泪]
11114	0 我也干过这事,一早兴冲冲的打开音乐一路狂奔,忽然...一脸苦逼相了! //@GGT:我也干...

Fig. 6. Examples of negative comments.

4.2 Data Processing

- 1) Text segmentation: Chinese text lacks explicit segmentation, necessitating methods for segmentation to train the word vector via the pre-training model. Common Chinese word segmentation tools include Jieba, SnowNLP, NLPIR, and THULAC. This article employs Jieba as the word segmentation tool.
- 2) Using stop-word list: Chinese texts often contain unnecessary high-frequency words like conjunctions, prepositions, and modal words, which can impact the efficiency and accuracy of classification models. Thus, these words should be removed during data processing. Microblog data often contains interference symbols and URLs. After regular expression processing, the stop-word list is applied to process the data. Due to the specific nature of microblog comments, this study combines Harbin Institute of Technology (HIT) and Baidu stop-word lists, supplementing them with additional specialized stop words to form an enhanced mixed stop-word list, thereby

improving data processing effectiveness. The process is as follows: first, the texts are segmented using the combined Baidu and HIT stop-word list. Then, based on the word segmentation results, high-frequency symbols and English words are added to the combined list, resulting in a new, more comprehensive mixed stop-word list. The added stop words are detailed in Fig. 7.

In the experimental data (Section 4.1), 30 positive and negative data samples were processed using the Jieba segmentation tool along with the newly integrated mixed stop-word list. The results of word segmentation for these examples are illustrated in Figs. 8 and 9.

≈ -- V [55555 555555 ~ / // "" [] {}] 【 】 ps Ps。 {} { }

Fig. 7. Examples of added stop words.

- 3) Vectorization processing: After data preprocessing, the processed data undergoes vectorization, transforming the result of word segmentation into the model’s input vector. In this paper, the CBOW model in Word2vec is adopted for word vectorization. Partial results after word vectorization are depicted in Fig. 10.

1	早上好 汪星 挖鼻屎 判断 四驱 励志 襄阳 东富 奥迪 史上 尊贵 进取 动感 汪星 请问 型号,1
2	这锅 你家 作用 最大化 鼓掌 鼓掌 鼓掌,1
3	赞 赞 赞 鼓掌 鼓掌 鼓掌 厨二代 客气 兄弟 时间 沙龙 做客 阿誉 菜品 设计 总监 一家 师兄 握手 腌渍涛 大厨 国际 厨皇会 花心,1
4	找妹 爱 嗨嗨 贺宁 嗨瑟 职位 不错 推荐,1
5	父亲节 送 老爸 一块 菜地 主意 不错 免子 现实 版 开心 农场 豆角 玉米 样 每周 老爸 老妈 采摘 开心,1
6	全国 土地利用 数据 制图 综合 数院库 兴起 电子地图 普及 涉及 地图学 美学 地理学 测绘 专业 学科 地图 综合 技术 研究 越来越,1
7	惭愧 中小学 写字课 临帖 钢笔字 楷书 印刷 练 行草 黄若舟 汉字 写法 十万 旌旗 骨里 笔力 沉着 端庄 刚劲 清卓 疏朗 确实 下过 苦功 鼓掌,1
8	不应 苛责 李 梦 孩子 涉罪 此刻 父母 揪心 可怜天下父母心 将心比心 行者 老丁 同复 凤 嗨嗨 凤凰 坑爹 儿子,1
9	霸气 新意,1
10	男人 世界 小姐 吃惊 那位 表情 不错 蜜 土豆 选 七见 影像 厅 天 赞成 这条 微博 会 火 全靠 右边 那位 表情,1
11	活动 启程 开启 鸡血 模式 嗨嗨 嗨嗨 嗨嗨 可爱 可爱 可爱,1
12	同问 余 总 起床,1
13	春打 六九头 月 日 时分 立春 农历 正月初五 分将 迎来 马年 第一个 节气 立春 民间 咬春 习俗 吃 春饼 春卷 萝卜 春到 人间 一卷 太 开心,1
14	有没有 可爱 挤眼 害羞 偷笑 太开心,1
15	笑一笑 十年 少 哈哈,1

Fig. 8. The word segmentation results of positive comments in Fig. 5.

1	孩子 太 悲剧 救 孩子 太 年轻 爱马仕 真的 别墅 名字 豪车 名字 美美 太 可怜 泪,0
2	枪毙 畜生 爆粗 怒,0
3	几位 女性 先 性别 筛选 再人肉 好些 软件 太雷人 抓狂 李开复 这是 人脸 搜索 眼镜 搜索 偷笑,0
4	十年 支付宝 花钱 说 泪 泪 泪 十年 记忆 支付 时光,0
5	不断 加 小朋友 真的 郭涛 真的 石头 爸爸 晕,0
6	帝都 空气质量 真心 不好 这才 几点 天 灰蒙蒙 开灯 悲伤,0
7	扎 针 抽 管 血 化验 一个月 喉咙 插根 管子 检查 同学 爱惜 胃 泪,0
8	吃惊 吃惊 回复 孔雀 金巧巧 卧槽 怎能 八卦 全 怒 贵圈 真乱 哑巴 换个 假胸 手感 哑巴 差 不知 孔雀 哑巴,0
9	烘焙 展 不好玩 吃 下午茶 时间 夏朵 花园 补 午饭 餐 点 披萨 海鲜 饭 龙虾 面 几样 前菜 份 甜点 抓狂 抓狂 抓狂,0
10	吃货 懊恼 假期 天 长 胖 斤 鸟 抓狂 抓狂 错 两贴 是因为 电影 热映 只顾 般配,0
11	代 小儿 老父 上柱 香 烧香纸 好多 话说 好多 心情 表达 个慢 热 亲人 离开 一年 伤感 慢慢 渗透 泪 敲 一记 钟 逝者 已去 日 追 ,0
12	望月 书 封面 伊朗 看不见 抓狂 抓狂,0
13	听罢 泪如雨下 听听 蜡烛 北京 厨子 听 上海 话 唱 听 大颖 沙鼠 宁财神 北京 厨子 大颖 沙鼠 五行 泪 蜡烛,0
14	要 不要 麽 泪,0
15	干过 这事 一早 冲冲 打开 音乐 一路 狂奔 一脸 苦逼 相 干过 这事 并线 衰脸,0

Fig. 9. The word segmentation results of negative comments in Fig. 6.

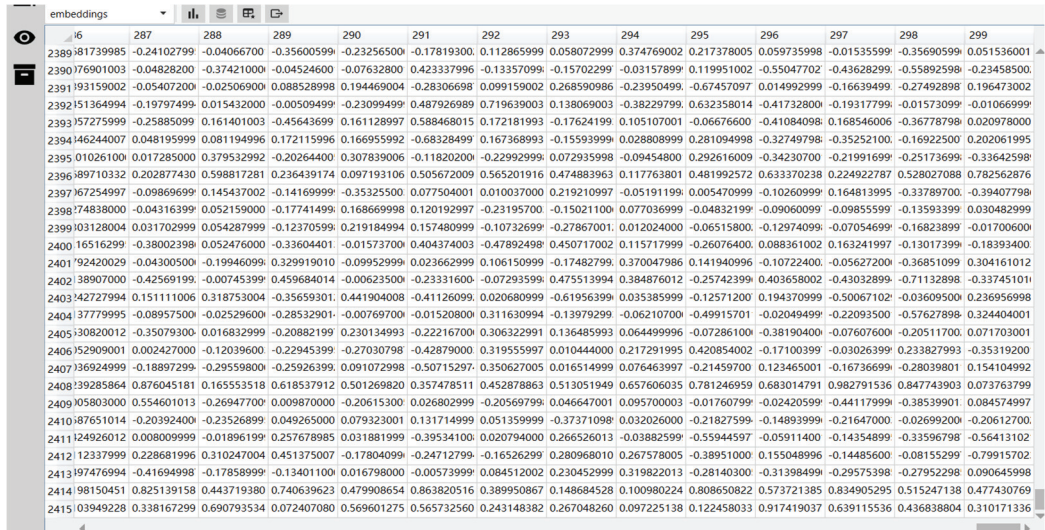


Fig. 10. The result of vectorization by continuous bag-of-word model.

4.3 Experimental Environment and Evaluation Indicators

4.3.1 Experimental environment

The environmental parameters are shown in Table 1.

Table 1. The experimental environment and environmental parameter

Experimental environment	Environmental parameter
Operating system	Windows 10
CPU	Intel Core i7-6700HQ CPU
GPU	NVIDIA GeForce GTX 965M (8 G)
Internal storage	16 GB
Programming language	Python 3.9
Deep learning framework	PyTorch
Exploitation environment	PyCharm

4.3.2 Evaluation indicators

The standard evaluation indicators for sentiment classification include accuracy rate (Acc), precision rate (P), recall rate (R), and comprehensive evaluation value (F1). The calculation formulas are as follows:

$$Acc = \frac{T}{N}, \tag{8}$$

$$P = \frac{TP}{TP+FP}, \tag{9}$$

$$R = \frac{TP}{TP+FN}, \tag{10}$$

$$F1 = \frac{2 * P * R}{P + R}. \tag{11}$$

The quantity of correctly predicted results by the classification model is represented by T, while N denotes the total number of samples. TP signifies the number of positive class data correctly forecasted as positive, FP stands for the number of negative class data incorrectly forecasted as positive, and FN represents the number of positive class data incorrectly forecasted as negative. This paper adopts the average values of the positive and negative indicators, specifically AP, AR, and AF1, as the evaluation criteria.

4.4 Analysis of the Experimental Results

The TextCNN-BiLSTM hybrid sentiment classification model is built on the PyTorch deep learning framework. Extensive parameter tuning tests were conducted to achieve optimal effectiveness (Table 2).

Table 2. Model hyperparameters

Parameter name	Value
Embedding dimension	300
The number of convolutional cores	200
Kernel size	[1,3,5]
Number of training samples	32
Epoch	20
Learning rate	1e-3
Dropout rate	0.2

1) Comparison experiment for stop-word list

As a critical step in sentiment analysis, data processing directly impacts the accuracy and efficiency of deep learning model training. In Jieba word segmentation, the stop-word list is applied to eliminate high-frequency interfering vocabulary, ensuring data reliability.

Based on the same parameters of the TextCNN-BiLSTM model, experiments were conducted using different stop-word lists: Baidu stop-word list (stopwords1), HIT stop-word list (stopwords2), Baidu and HIT mixed stop-word list (stopwords3), and the new mixed stop-word list (stopwords4). The experimental results are shown in Fig. 11.

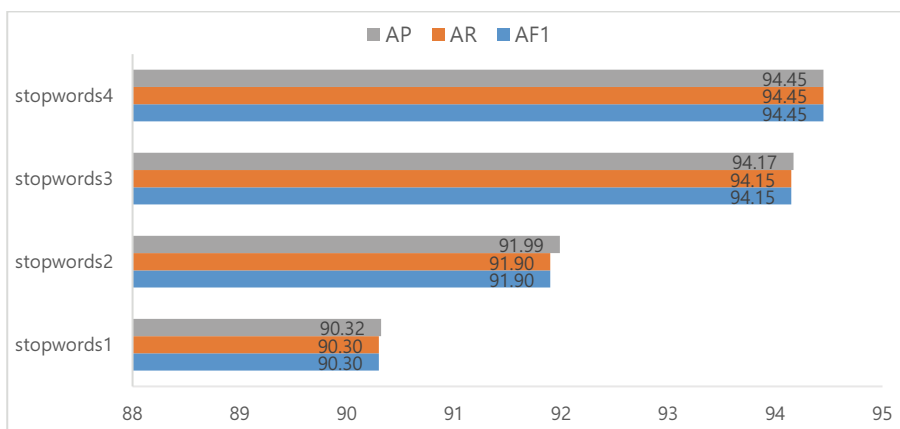


Fig. 11. The experimental results of TextCNN-BiLSTM model with different stop-word lists.

From Fig. 11, it's evident that the single Baidu and HIT stop-word lists are inadequate for data processing. The Baidu stop-word list covers only a limited number of specific words, while the HIT stop-word list lacks English stop words. Data processing based on the combined Baidu and HIT mixed stop-word list has a significantly improved effect. The new mixed stop-word list, built upon the combined Baidu and HIT list, demonstrates even better data processing effectiveness, particularly for special emoticons and interfering English words in microblog comments. The accuracy rate of the new mixed stop-word list is 4.13%, 2.56%, and 0.28% higher than that of the Baidu stop-word list, HIT stop-word list, and Baidu and HIT mixed stop-word list, respectively.

2) Comparison experiment for activation function

Common activation functions include Tanh, ReLU, and sigmoid, with ReLU being the frequently used function in the model. In this paper, while keeping other parameters constant, the activation functions of the TextCNN convolution layers and the softmax layers of the hybrid model were set to Tanh, ReLU, and sigmoid. The experimental results are shown in Fig. 12.

From Fig. 12, it's evident that the model achieves the highest precision when Tanh is used as the activation function. It outperforms ReLU by 0.3% and Sigmoid by 0.41%. Given the numerous hidden layers in BiLSTM, using Sigmoid can lead to gradient disappearance and relatively poor classification results. While ReLU helps mitigate the vanishing gradient problem, it comes with the issue of some neurons remaining inactive, preventing parameter updates. Tanh function proves effective in handling the large number of hidden layers in the hybrid model, making it more suitable for binary classification tasks, resulting in overall superior performance.

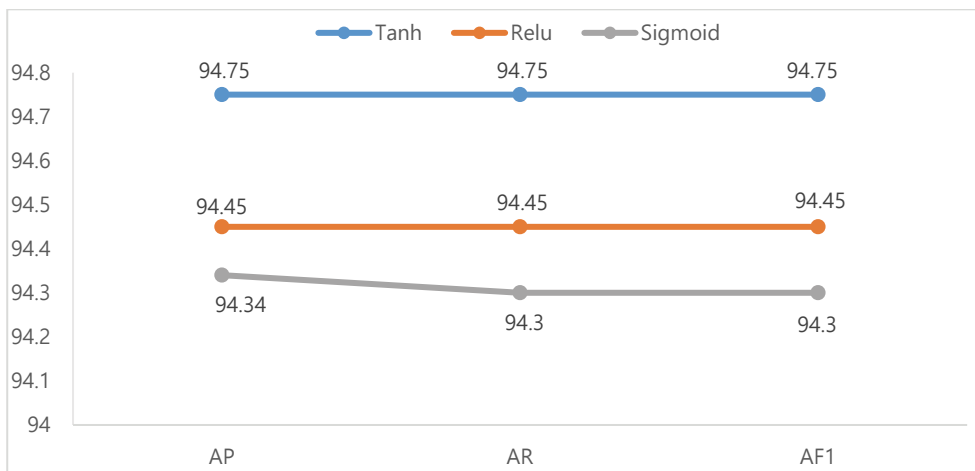


Fig. 12. The experimental results of TextCNN-BiLSTM model with different activation functions.

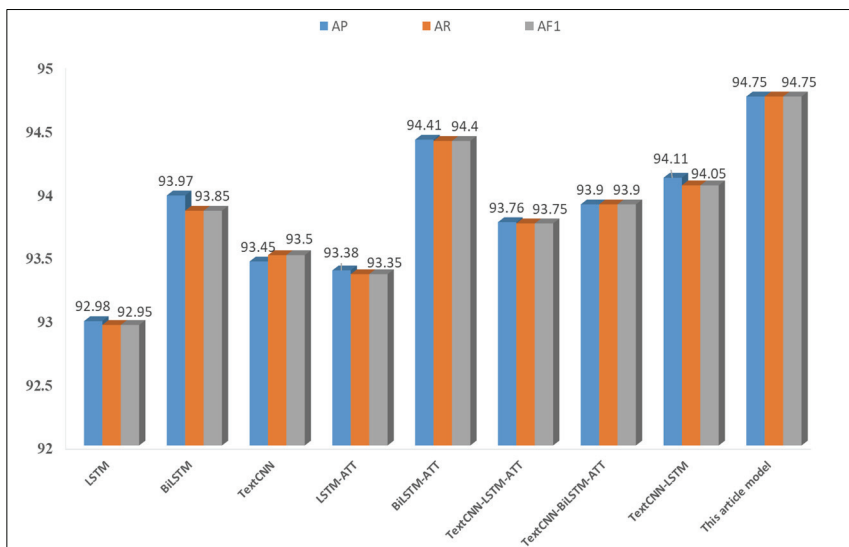
3) Comparison experiment of the classification model

To validate the effectiveness of our proposed improved hybrid model, we conducted tests using the same dataset. We continued to train word vectors with Word2vec, which were then used as inputs for the model. TextCNN, LSTM, BiLSTM, and LSTM-ATT models were employed for comparative testing. The experimental results of each model are presented in Table 3.

Table 3. The experimental results

Model	AP (%)	AR (%)	AF1 (%)
LSTM	92.98	92.95	92.95
BiLSTM	93.97	93.85	93.85
TextCNN	93.45	93.50	93.50
LSTM-ATT	93.38	93.35	93.35
BiLSTM-ATT	94.41	94.40	94.40
TextCNN-LSTM-ATT	93.76	93.75	93.75
TextCNN-BiLSTM-ATT	93.90	93.90	93.90
TextCNN-LSTM	94.11	94.05	94.05
Proposed model	94.75	94.75	94.75

Upon examination of Table 3, it becomes evident that TextCNN, a variant of CNN, excels at extracting local features. However, due to its limited ability to capture long-distance dependencies between words, its classification effectiveness is relatively modest. LSTM and BiLSTM models demonstrate similar classification effects. Yet, the BiLSTM model, with its focus on context, exhibits a slightly enhanced classification effect compared to the LSTM model which solely emphasizes the preceding context. The BiLSTM-ATT model, leveraging bidirectional LSTM for context feature extraction followed by the application of an attention mechanism to weigh the extracted crucial features, achieves an outstanding classification effect. The TextCNN-BiLSTM-ATT model combines a convolutional neural network and bidirectional LSTM to extract features, followed by a weighted feature processing step. However, given that microblog comments are inherently brief, the accuracy of the extracted features diminishes after weighting. The hybrid TextCNN-BiLSTM model in this paper, by integrating TextCNN for local feature extraction and BiLSTM for context feature extraction, improves accuracy, recall rate, and F1 value by 1.21%, 1.25%, and 1.25%, respectively, compared with the standalone TextCNN model. Furthermore, it demonstrates an improvement of 0.78%, 0.9%, and 0.9% respectively when compared with the BiLSTM model. The comparative results for these models are illustrated in Fig. 13.

**Fig. 13.** The experimental results of different sentiment analysis models.

label	review
1	作者真有英国人严谨的风格, 提出观点、进行论述论证, 尽管本人对物理学了解不深, 但是仍然能感受到...
1	作者长篇大论借用详细报告数据处理工作和计算结果支持其新观点。为什么荷兰曾经具有欧洲最高的生产...
1	第一天晚上八点下单, 第二天下午3点多货物送到。xp驱动装F8vr系列的就行, 官网上就有下载, ...
1	散热好, 屏幕清晰, 键盘手感很好, 处理速度快, 硬盘声音小, 做工不错
1	酒店环境不错, 可房间一般。预订时一直强调要视野好一些的, 结果因先前入住的客人过了入住时间还没...
1	今天带糖果去超市, 她眼尖看到了健达的那个新款蛋形巧克力, 指着对我说, 妈妈这就是电视里广告的那...
1	蒙牛与伊利我就是喜欢伊利! 蒙牛就像明星, 宣传做的好, 但总有种太招摇的感觉...而且我发现它赞助的...
1	作为一台以高性价比的数码照相机0T735切入中端彩屏手机市场, 凭借其低价和内置摄像头的魅力...
0	我不是很受别人影响的人, 也不是此书最热的时候买的, , 事实上我也不知道此书, 在当当最近遇到之前...
0	我不明白, 为什么会有这么多的人追捧这本书呢, 我也是受了大众思想的影响, 一口气买了好多本, 用来...
0	运气超不好, 买来开箱向下方向键就是坏的, 要我打的去ACER维修站开检测单, 然后还要网上填表申...
0	散热不太好, 触摸板那里温度高。用超级兔子测试显示cpu温度60多。
0	我无话可说, 只觉得对不起近400元的房费, 附近有经济连锁酒店都比这家要性价比要好。这家酒店对...
0	房间装修实在是不敢恭维, 洗脸盆下面还漏水, 还四星级, 我看三粒星可能是花钱买的, 还不如锦江之星...
0	内屏是单色的, 现在哪个手机不是彩色的? 单位同事都在比谁的颜色更靓谁拍出的照片强, 可我用的还是...
0	晕倒, 我的小U坏了, 前些天只是偶尔开不了震动, 以为是用机太狠了, 所以没注意, 昨天正常使用时, ...

Fig. 14. Examples of e-commerce reviews.

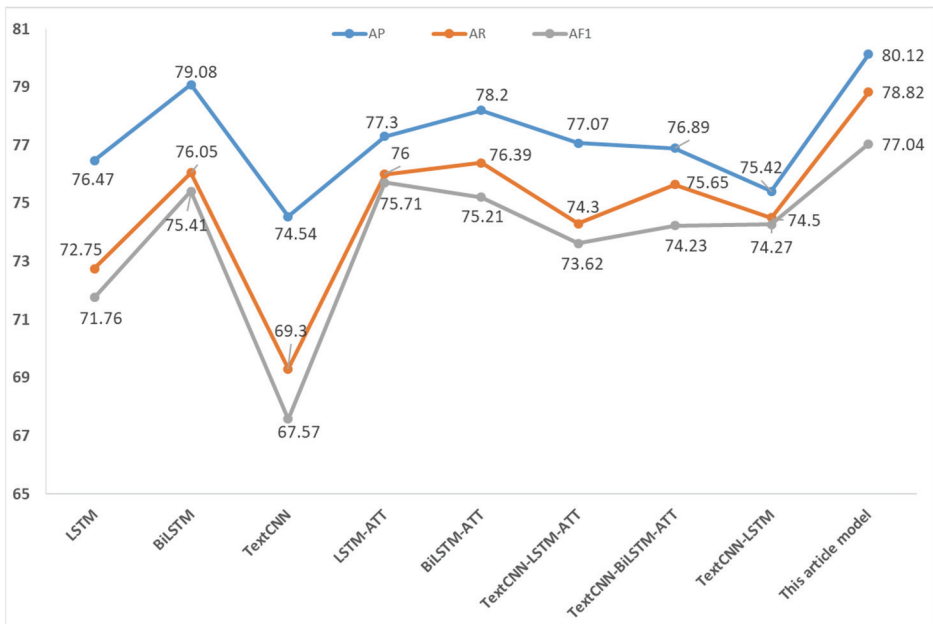


Fig. 15. The experimental result of different models in e-commerce reviews.

4) Verification Experiments with Different Datasets

The aforementioned experimental results indicate that the TextCNN-BiLSTM classification model excels in sentiment analysis of microblog comments. To corroborate the validity of this model, we selected another dataset for verification under the same experimental environment and parameters as in experiment (3).

The online_shopping_10_cats dataset encompasses e-commerce reviews for 10 categories of products, ranging from books and mobile phones to computers and hotels. Each category consists of positive and negative comments, with positive comments labeled as 1 and negative comments as 0. This dataset offers

broad coverage and is highly representative. We selected specific reviews from `online_shopping_10_cats` to form an experimental dataset, comprising 10 thousand positive reviews and 10 thousand negative reviews. A portion of the experimental dataset is displayed in Fig. 14.

Comparative experiment results demonstrate that the improved hybrid model continues to outperform other models in sentiment analysis of e-commerce reviews. The experimental results for these classification models are depicted in Fig. 15.

5. Conclusion

This paper introduces an enhanced mixed deep learning model for classifying emotions in microblog user comments. The methodology involves using Jieba word segmentation with a mixed stop-word list for data preprocessing, followed by Word2vec model utilization for word vectorization. Subsequently, the TextCNN model is employed for local feature extraction, while BiLSTM fully captures context features. The outcome is effective emotion classification. Experimental results demonstrate that the improved hybrid neural network model surpasses single models in sentiment analysis, achieving a precision of 94.75%. By employing a mixed stop-word list and Tanh activation function, the model demonstrates superiority over the unimproved version. Specifically, the results indicate that the model with the new stop-word list and Tanh activation function enhance accuracy rates by 4.1% and 0.4%, respectively. However, it should be noted that this study predominantly focuses on classifying emotions in Chinese texts, whereas microblog comments on major social platforms often include English content. Therefore, further research into sentiment classification of mixed Chinese and English texts is warranted in the future.

References

- [1] R. Zeng and X. Xu, "A study on early warning mechanism and index for network opinion," *Journal of Information*, vol. 28, no. 11, pp. 52-54, 2009.
- [2] R. Li, Z. Lin, H. Lin, W. Wang, and D. Meng, "Text emotion analysis: a survey," *Journal of Computer Research and Development*, vol. 55, no. 1, pp. 30-52, 2018. <https://doi.org/10.7544/issn1000-1239.2018.20170055>
- [3] J. Wu, K. Lu, and S. B. Wang, "Sentiment analysis of film review based on multiple sentiment dictionary and SVM," *Journal of Fuyang Normal University (Natural Science edition)*, vol. 36, no. 2, pp. 68-72, 2019.
- [4] Z. Cheng and L. Wang, "Sentiment analysis method of online comments based on support vector machine," *Electronic Technology & Software Engineering*, vol. 2019, no. 16, pp. 3-4, 2019.
- [5] W. J. Cui, "Deep learning-based text emotion analysis," Ph.D. dissertation, Jilin University, Changchun, China, 2018.
- [6] L. Dang and L. Zhang, "Method of discriminant for Chinese sentence sentiment orientation based on HowNet," *Application Research of Computers*, vol. 27, no. 4, pp. 1370-1372, 2010.
- [7] K. Wang and R. Xia, "A survey on automatic construction methods of sentiment lexicons," *Acta Automatica Sinica*, vol. 42, no. 4, pp. 495-511., 2016 <https://doi.org/10.16383/j.aas.2016.c150585>
- [8] Y. M. Zhou, J. L. Yang, and A. M. Yang, "A method on building Chinese sentiment lexicon for text sentiment analysis," *Journal of Shandong University (Engineering Science)*, vol. 43, no. 6, pp. 27-33, 2013.

- [9] Y. Zhao, B. Qin, Q. Shi, and T. Liu, "Large-scale sentiment lexicon collection and its application in sentiment classification," *Journal of Chinese Information Science*, vol. 31, no. 2, pp. 187-193, 2017.
- [10] C. Jiang, Y. Guo, and Y. Liu, "Constructing a domain sentiment lexicon based on Chinese social media text," *Data analysis and Knowledge Discovery*, vol. 3, no. 2, pp. 98-107, 2019. <https://doi.org/10.11925/infotech.2096-3467.2018.0578>
- [11] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? Sentiment classification using machine learning techniques," in *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Philadelphia, PA, USA, 2002, pp. 78-86.
- [12] J. Li, Y. Rao, F. Jin, H. Chen, and X. Xiang, "Multi-label maximum entropy model for social emotion classification over short text," *Neurocomputing*, vol. 210, pp. 247-256, 2016. <https://doi.org/10.1016/j.neucom.2016.03.088>
- [13] S. Kaur, G. Sikka, and L. K. Awasthi, "Sentiment analysis approach based on N-gram and KNN classifier," in *Proceedings of 2018 1st International Conference on Secure Cyber Computing and Communication (ICSCCC)*, Jalandhar, India, 2018, pp. 1-4.
- [14] L. Zhang, Y. Tan, L. Zhu, and W. Dong, "Analyzing the features of negative sentiment microblog," *Intelligence Theory and Practice*, vol. 42, no. 7, pp. 132-137, 2019. <http://www.itapress.cn/CN/Y2019/V42/I7/132>
- [15] M. Sun, Y. Li, Z. Zhuang, and T. Qian, "Sentiment analysis based on BGRU and self-attention mechanism," *Journal of Jiangnan University (Natural Science Edition)*, vol. 48, no. 4, pp. 80-89, 2020. <https://doi.org/10.16389/j.cnki.cn42-1737/n.2020.04.011>
- [16] H. Zhao, L. Wang, and W. Wang, "Text sentiment analysis based on serial hybrid model of bi-directional long short-term memory and convolutional neural network," *Journal of Computer Applications*, vol. 40, no. 1, pp. 16-22, 2020. <https://doi.org/10.11772/j.issn.1001-9081.2019060968>
- [17] Y. Miao, Y. Ji, S. Zhang, W. Cheng, and E. Peng, "Application of CNN-BiGRU model in Chinese short text sentiment analysis," *Information Science*, vol. 39, no. 4, pp. 85-91, 2021.
- [18] Q. Yang, Y. Zhang, J. Zhu, and T. Wu, "Text sentiment analysis based on fusion of attention mechanism and BiGRU," *Computer Science*, vol. 48, no. 11, pp. 307-311, 2021. <https://doi.org/10.11896/jsjcx.201000075>
- [19] L. Yan, X. Zhu, and X. Chen, "Emotional classification algorithm of comment text based on two-channel fusion and BiLSTM-attention," *Journal of University of Shanghai for Science and Technology*, vol. 43, no. 6, pp. 597-605, 2021. <https://doi.org/10.13255/j.cnki.jusst.20210102001>
- [20] H. Fan and P. F. Li, "Sentiment analysis of short text based on FastText word vector and bidirectional GRU recurrent neural network: take the microblog comment text as an example," *Information Science*, vol. 39, no. 4, pp. 15-22, 2021. <https://lib.cqvip.com/Qikan/Article/Detail?id=7104517819>
- [21] X. Yang, M. Guo, H. Hou, J. Yuan, X. Li, K. Li, W. Wang, S. He, and Z. Luo, "Improved BiLSTM-CNN+ Attention sentiment classification algorithm fused with sentiment dictionary," *Science Technology and Engineering*, vol. 22, no. 20, pp. 8761-8770, 2022. <http://www.stae.com.cn/jsygc/article/abstract/2112609?st=alljournals>
- [22] B. Shen, X. Yan, L. Zhou, G. Xu, and Y. Liu, "Microblog sentiment analysis based on ERNIE and dual attention mechanism," *Journal of Yunnan University (Natural Science Edition)*, vol. 44, no. 3, pp. 480-489, 2022. <https://doi.org/10.7540/j.ynu.20210263>
- [23] M. H. Ali Al-Abyadh, M. A. Iesa, H. A. Hafeez Abdel Azeem, D. P. Singh, P. Kumar, M. Abdulmir, and A. Jalali, "Deep sentiment analysis of twitter data using a hybrid ghost convolution neural network Model," *Computational Intelligence and Neuroscience*, vol. 2022, article no. 6595799, 2022. <https://doi.org/10.1155/2022/6595799>
- [24] Y. Hu, T. Tong, X. Zhang, and J. Peng, "Self-attention-based BGRU and CNN for sentiment analysis," *Computer Science*, vol. 49, no. 1, pp. 252-258, 2022. <https://doi.org/10.11896/jsjcx.210600063>
- [25] L. Khan, A. Amjad, N. Ashraf, and H. T. Chang, "Multi-class sentiment analysis of Urdu text using multilingual BERT," *Scientific Reports*, vol. 12, no. 1, article no. 5436, 2022. <https://doi.org/10.1038/s41598-022-09381-9>

- [26] O. Wu, T. Yang, M. Li, and M. Li, "Two-level LSTM for sentiment analysis with lexicon embedding and polar flipping," *IEEE Transactions on Cybernetics*, vol. 52, no. 5, pp. 3867-3879, 2022. <https://doi.org/10.1109/TCYB.2020.3017378>
- [27] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, 1998. <https://doi.org/10.1109/5.726791>
- [28] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 1997. <https://doi.org/10.1162/neco.1997.9.8.1735>



Haiqin Tang <https://orcid.org/0000-0002-5485-9596>

She received B.S. degree in Financial Management from Southwest Petroleum University in 2020. She is currently an M.S. candidate in Business and Tourism School from Sichuan Agricultural University, Chengdu, China. Her current research interests include natural language processing and deep learning.



Ruirui Zhang <https://orcid.org/0000-0003-1898-1487>

She received B.S., M.S., and Ph.D. degrees in School of Computer Science from Sichuan University in 2004, 2007, and 2012, respectively. She is a lecturer at the School of Business and Tourism, Sichuan Agricultural University, China. Her current research interests contain network security, wireless sensor networks, intrusion detection and artificial immune systems.