Research Article

Research on Sentiment Analysis Model of Short Text Based on Deep Learning

Zhou Gui Zhou

Computer College of Huanggang Normal University, Huanggang, Hubei 438000, China

Correspondence should be addressed to Zhou Gui Zhou; hgsyzgz@hgnu.edu.cn

Received 25 April 2022; Revised 10 May 2022; Accepted 16 May 2022; Published 29 May 2022

Academic Editor: Jie Liu

Copyright © 2022 Zhou Gui Zhou. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the wide application of the Internet and the rapid development of network technology, microblogs and online shopping platforms are playing an increasingly important role in people's daily life, learning, and communication. The length of these information texts is usually relatively short, and the grammatical structure is not standardized, but it contains rich emotional tendencies of users. The features used by custumal machinery schooling methods are too sparse on the vector space model and lack the semantic information of short texts, which cannot well identify the semantic features and potential emotional features of short texts. In response to the above problems, this paper proposes a bidirectional long-term and short-term memory network model based on emotional multichannel, combining the attention mechanism and convolutional neural network features in deep learning and learning the short text by combining shallow learning and deep learning. The semantic information and potential emotional information of the short text can be improved to promote the effective expression of short-text emotional features and improve the short-text emotional classification effect. Finally, this paper compares the above models on multidomain classification data sets such as NLPIR and NLPCC2014. The accuracy and F1 value of the model proposed in this paper have achieved good improvement in the field of short-text sentiment analysis.

1. Introduction

These days, people want to check the latest current affairs, online shopping, news gossip, and financial stocks; people are no longer limited to reading newspapers or sitting in front of the TV to watch hot topics but have more ways to participate in the discussion of hot topics such as Weibo, Taobao, Douyin, Zhihu, and WeChat public platforms and other media. As an Internet platform, Weibo, Taobao, and Douyin, in another aspect, realize information sharing and dissemination by virtue of user relationships, attracting a large amount of individuals to participate, and are favored by people; on the other hand, a large amount of posts published by users are text mining and provide a huge amount of data. Sun et al. found that Weibo information can reflect changes in people's attention to hot spots [1] and can even infer the current emotional condition of users based on Weibo user information. In addition, Weibo sentiment analysis also provides reference opinions for some industries, such as stock trading decisions [2], movie box office predictions, and election predictions [3, 4]. Literature provides reference opinions for consumers to purchase products by mining online shopping platform product review information and establishing a review sentiment analysis model [5–7] and guidance for merchants to adjust production plans and product improvement and also promotes online shopping platforms. Users are provided with a more efficient quality of service. With the vigorous development of online social media and the rise of artificial intelligence, more and more experts, scholars, and scientific research institutions are now turning their attention to the analysis.

1.1. Research Status of Text Sentiment Analysis. Text sentiment analysis is an indispensable link in natural language handling. In the past, relatively large part experts and scholars have carried out research in the light of sentiment dictionary. Li and Hong reviewed sentiment analysis methods, respectively [8, 9].
1.2. Research Status of Short-Text Sentiment Analysis.

(1) In deep learning, the amount of training data is required by the model. When the scale of the model is large enough, the connection between short-text data can be fully captured. However, in most cases, facing the problem of short-text sentiment orientation classification and other classification problems, sufficient training data cannot be found, and a lot of data are manually annotated, and the model is

(2) In the method based on machine learning, Xie et al. combined sentiment dictionary, context features, and topic features [20] and proposed an SVM-based sentiment classification method. Li and Ji extracted features such as words, negative words, and special symbols to construct an SVM model and a CRF model to perform sentiment analysis on microblog data [21] and concluded that the appropriate choice should be made under different circumstances. Conclusions of the model.

(3) In the method based on deep learning, Zhou integrated part-of-speech features and word embedding features in the research on sentiment classification of product reviews [7]. The experiments are higher than the traditional text convolutional neural network. Chen proposed a microblog using part-of-speech features of emotional words and learning more hidden information [22]. The experiments verified the proposed model is robust to different data.

Because the short text is relatively short, it will bring about the problem of lack of text semantics, which brings challenges to the short-text sentiment analysis. Although the existing short-text sentiment analysis methods have done some feature extraction, feature selection, and model selection, many works still do not fully consider the context of short texts and deeply dissect semantic features. Some new words may not be recognized in the word segmentation stage. In order to improve the shortcomings of existing methods in feature selection, this paper extracts shallow learning features such as emotional part of speech, location information and dependencies of words from short texts, as well as deep learning features such as word vector features, convolutional neural network features, and emotional attention features. Learning features enrich the textual feature representation of short texts.

In the near future, deep learning has also been widely used in sentiment analysis of short texts. Its concept comes from the research of artificial neural network. Its purpose is to explain the feature information existing in the data by imitating the thinking structure and learning mechanism of the human brain and build a neural network for machine analysis and learning. Compared with the use of nonlinear network structure to make up for the shortcomings of the algorithm, it also has the following two shortcomings:

1) The medium in the light of the vocabulary for expressing emotions and emotional tendencies mainly judges the sentiment climate. It needs to manually construct the vocabulary for expressing emotions and emotional tendencies mainly or use the internal statistics Mutual Information (MI), Symmetric Conditional Probability (Symmetric Conditional Probability), external statistics (Branch Entropy and Access or Variety), and other methods to expand the sentiment dictionary. In text orientation analysis, well-known sentiment dictionaries are HowNet2, WordNet [10], and ConceptNet [11].

2) The feature-based method is to use statistical knowledge to screen features from a large quantity of corpora, use features to represent the entire text, and then use relatively unnoticeable algorithms in machine schooling to classify the text. This method requires high feature engineering and feature selection. The results directly affect the classification effect. For a long time, features have occupied an important position in text classification, the classification matrix has not been greatly improved, and it has indirectly led to the problem of overfitting, which is called the Hughes effect [12–14]. In reality, training a large number of features requires enough samples, where obtaining enough samples requires time and labor.

3) Based on deep learning methods, features such as words, sentences, and chapters can be mapped to high-dimensional spaces to learn deeper feature representations in text data. Wang added a self-attention mechanism (Attention Mechanism) after the output storey of the LSTM network [15] and obtained the context information of the LSTM output unit attention means for relevant automatic acquisition. The experimental results show that the attention mechanism can recognize the emotional information in the text [16, 17]. The model first inputs the word vector into the Bidirectional LSTM net to learn the textual content and emotional information, and then uses the self-look mechanism to extract emotional representations of monolingual and bilingual texts, respectively. Based on the above work, in the work of sentiment analysis, each word in the text has a multitudinous collision on the overall emotional climate of the text, especially some emotional words, which can often directly reflect the emotional climate of the text, and through the look mechanism, the importance of words can be obtained, and the potential representation information in the text can be learned.
2. Relevant Theoretical Basis and Technical Introduction

In recent years, the analysis and mining of time series data have been gradually applied to many fields. Natural language processing is a typical application [23–26]. In the research of many scholars in the scope of natural language processing, the text sentiment analysis task has two important steps. The first is to convert the text information into coded information that cannot be directly understood by the computer, and the second is to analyze the sentiment tendency of the text. Since the words in the text cannot be directly fed into the network model, the first task is to convert the text into a digital representation. The word vector is to map the words in the text into a digital vector representation. According to the different encoding methods, word vectors are mainly divided into discrete word vectors and distributed word vectors.

2.1. Text Representation. Natural language is a complex system that expresses a given intention and thought. It is generally composed of words and punctuation marks. One, two or more words are spliced into a word, and several words are connected to form a sentence. After continuous combination, it forms paragraphs and chapters. Unlike humans, machines cannot directly understand the emotional information in language but need to obtain the corresponding information in language by establishing certain rules or models [27], in which only one bit is 1 and the rest are 0. Under the one-hot rule, the words after word segmentation are discretized and mapped to the Euclidean space by the implementation of row vectors, but in large-scale data sets, the vocabulary size of a data set may reach tens of thousands of dimensions or even ten of thousands of dimensions, and vectors at this time will undoubtedly bring huge memory consumption. At the same time, the large-scale vocabulary makes the constructed word vector matrix too sparse, which brings great inconvenience to the feature schooling. In view of the fact that one-hot encoding will have problems such as dimensional disaster, word similarity, and poor model generalization ability in natural language modeling, Google proposed the word2vec model in 2013, also known as the word embedding model. Each position in the model-trained and the general value range is between –10 and 10. Taking “I appreciate one country, two systems” as an example, the 200-dimensional word vector representation is set as shown in Table 1.

2.2. Related Methods of Sentiment Analysis. Sentiment analysis is essentially a text classification problem. Deep learning can learn deep features in data and bring about brilliant bonanza in odd spheres. Therefore, profound learning has also been diffusely used in sentiment decomposition tasks in recent years. Neural networks commonly used in text sentiment decomposition include CNN, recurrent neural networks, and LSTM networks. These networks can better mine the latent information hidden in the text and outperform most custumal machinery schooling methods.

Since RNN adds a loop structure to the traditional neural network, the content of the loop body will be executed at each step, but when the number of historical nodes input to the RNN decreases, the RNN cannot memorize information far from the current node. LSTM is an improved model of RNN. The main improvement is the introduction of three phylums when memorizing information: input gate, forget gate, and output gate. Through these three gates, LSTM can bridge the information passing through the cell and can selectively add information or delete existing information according to the needs of the result, and specific LSTM network structure and internal structure are shown in Figure 1.

At the outset, LSTM used the forget gate to fix the message that the cell demands to discard. Forget phylum calculates a value among zero and one based on the historical message of the former condition and the current input communication as the condition of the cell information at the previous moment to determine what information to keep and discard. Among them, 0 means to discard all the communication of the historical condition, and 1 means to keep all the communication of the historical condition, where \( f(t) \) delegates the input value at time \( t \), \( h(t−1) \) delegates the worth of the concealed stratum at time \( t−1 \), and \( U(f), W(f), \) and \( b(f) \) are the degree of seriousness of the worth of LSTM of the concealed stratum in the forget phylum, the degree of seriousness of the current input, and the bias in the forget phylum, respectively, \( \sigma \) is the activation function of sigmod, and \( f(t) \) is the information discarded by cells. The specific calculation is shown in the following equation:

\[
f^{(t)} = \sigma \left( \sum h^{(t−1)} \cdot U^{(f)} + x^{(t)} \cdot W^{(f)} + b^{(f)} \right).
\]  

(1)

LSTM determines the information stored in the cell through the input phylum. The input phylum calculates a value from 0 to 1 by sigmod, the key in news to modernize the condition of the current node, that is, what information needs to be updated or stored. Among them, 0 means not accepting new information and 1 means accepting all the input information. LSTM generates a new memory \( C(t) \), which is the input memory, not the final memory. The condition is determined by the previous output and the current input, which are the worth of the concealed stratum of LSTM in the new memory, the degree of seriousness of the current input, and the degree of seriousness of the current input. The specific calculation of the bias in the new memory is shown in the following equation :

\[
\tilde{C}^{(t)} = \tanh \left( \sum h^{(t−1)} \cdot U^{(c)} + x^{(t)} \cdot W^{(c)} + b^{(c)} \right).
\]  

(2)
Table 1: Word2vec training word vector table.

| Words and phrases | Word embedding |
|------------------|---------------|
| Myself           | [-3.2554681, 3.0520194, -1.3131773, ..., 0.22637065] |
| Grateful         | [1.2681537, -0.09857514, 1.5265175, ..., 0.07370517] |
| One country, two systems | [-0.6349963, 1.6289423, -0.4314866, ..., -0.4279167] |

The final desired memory is generated by the part to be remembered and the part to be forgotten. Through the calculation of the previous two phylums, we already know the proportion \( i(t) \) of new information retained, the proportion of old information that needs to be forgotten, and the new memory and old memory \( f(t) \). In the final step, the output phylum is used to determine the upshot of the final cell output. The output phylum calculates a value between 0 and 1 by sigmoid—the input information to calculate the proportion of information that the cell finally outputs. Among them, 0 indicates that no information is output and 1 indicates that all the final memory results are output. The specific calculation is shown in the following equation:

\[
C^{(t)} = i^{(t)} \odot \tilde{C}^{(t)} + f^{(t)} \odot C^{(t-1)}. \tag{3}
\]

In a classical recurrent network, the condition and output are always transmitted from front to back. However, in some problems, the transmission and output of the condition are not only related to the previous condition but also related to the subsequent condition such as prediction. A missing word is not only related to the preceding text but also to the following text. Among them, BiLSTM is a common neural network model that considers the contextual relationship. It is based on Bi-RNN. One LSTM module propaphylums from front to back, and another LSTM module propaphylums from back to front. BiLSTM builds a double-storey network model. The input is forwarded to the LSTM and the reverse LSTM, respectively, and the final output result is the vector superposition of the LSTM output results in two different directions. Finally, the entire output result is fully connected, and then, the sigmoid output is the final output result, where BiLSTM is trained in the same way as LSTM [28–30].

3. Sentiment Classification Model Based on Deep Learning

Considering that the custumal machinery schooling methods and profound learning methods are insufficient in feature representation, this paper makes two improvements to the custumal machinery schooling methods and profound learning methods: one is to use shallow learning features as one of the input features of profound learning. Add a look mechanism to the network layer to allow the model to better learn the underlying semantic features in the text.

3.1. A Bidirectional Long- and Short-Term Memory Network Model Based on Emotional Multichannel. So as to make full use of the unique emotional resource information in the text sentiment decomposition task, this paper proposes BM-ATT-BiLSTM based on multichannel sentiment. The learning storey builds multiple channels to improve sentiment classification performance. BM-ATT-BiLSTM is a left-to-right multistorey neural network structure mainly composed of 5 parts: input layer, semantic learning layer, emotional attention layer, merging storey, and sentiment classification output storey. The input storey inputs are composed of features and shallow features (emotional part-of-speech features of words, location information features, and dependency features). The traditional LSTM model can obtain the forward semantic information in the text but ignore the reverse semantic information of the text. In response to this paper, let LSTM learn backward directions of the sequence. Feed the data in both directions into the BiLSTM model, and the calculation method is relatively simple, by calculating the emotional weight of each word. If the previous storey of the BiLSTM network is multiplied, it will be normalized and fused batch. Use the fully connected layer of the model to output the feature matrix, use softmax to normalize the feature matrix, and finally get the classification result.

3.2. Experimental Parameter Setting. Different hyperparameters may have different effects on the experimental results. Although parameter tuning itself is not the main research content of this paper, for the sake of fairness, this paper considers the overall effect of the experiment and
finally determines the hyperparameters of BM-ATT-BiLSTM. Some hyperparameters of the BM-ATT-BiLSTM model are shown in Table 2.

3.3. Experimental Environment and Comparative Experiments. The computer hardware configuration used in the experiment in this paper is as follows: the CPU is Intel Core processor i5-9400f, and the GPU series is NVIDIA GTX; the operating system is configured as Windows 10; the programming software used is PyCharm, the programming language used is Python, and the support library used in profound learning is pytorch10, keras11, and gensim12.

3.4. Experimental Results and Analysis. Since the NLPIR and NLPCC2014 data sets do not contain many training samples, too many iterations may lead to overfitting problems, and if the number of iterations is insufficient, it is troublesome for the matrix to be taught effective features. Therefore, 20% of the data set is divided into the validation set, 16% is divided into the test set, and 64% is divided into the drilling battery. Precision, Recall, and F1-measure are selected as evaluation indicators. So as to make the experiment more fair, the evaluation index results in this paper take the average of 50 experimental results. The detailed experimental results are shown in Table 3.

(1) Compared with LSTM, the overall effect of CNN is weaker than that of LSTM. In the process of extracting features, CNN mainly captures multiple different N-grams of text, and there are many different convolution kernels for one N-gram. Useful information is extracted from different angles, but the experimental data sets are all short texts, and the amount of data is not large, so the use of CNN to extract features may lead to insufficient features and adds emotional attention to CNN, and the effect is obviously promoted.

(2) The results of CNN and CNN + SVM show that using SVM instead of softmax can improve the classification effect. The reason is that the loss function of SVM can get faster convergence on the three data sets of text, and the output of softmax is only a probability of compressed data, and the probability distribution of softmax is deviated from the actual result.

(3) Among all the LSTM memory networks involved in the comparative experiments, BM-ATT-BiLSTM has the best effect. An important reason is that the emotional attention mechanism is added, and more potential emotions can indeed be learned from the text through attention. According to Table 3, the BM-ATT-BiLSTM method outperforms other models in terms of precision and recall. This effect is essentially due to the time series characteristics of natural language. Cells in the LSTM model can effectively record the time series information in the text. The BiLSTM model structure is used to learn the semantic information in the text, in order to enhance the capability of the mold to learn the reverse text, and strengthen the capability to seize the news of the text context. Combining the emotional parts of speech, location information of words, as well as word embedding features, emotional attention features, and features extracted by convolutional neural networks, this paper proposes three neural network models, which are the foundations of emotional multichannel features—BiLSTM classification model, convolutional neural network in the model mechanism increases the attention, and multikernel convolutional neural network in the model mechanism increases the attention. Tentatives show that the BM-ATT-BiLSTM recommended in this paper has the best performance; therefore, it can be concluded that adding the above shallow learning features and profound schooling features to the short-version sentiment decomposition can improve the classification manifestation of the mold.

4. Conclusion and Outlook

Combining the emotional part of speech features, location information features, and dependency features of words, as well as word embedding features, emotional attention features, and features extracted by convolutional neural networks, this paper proposes three neural network models, which are emotional multichannel features—BiLSTM classification model, convolutional neural network in the model mechanism increases the attention, and multikernel convolutional neural network in the model mechanism increases the attention. Tentatives show that the BM-ATT-BiLSTM recommended in this paper has the best performance, so it can be concluded that adding the above shallow

Table 2: BM-ATT-BiLSTM model hyperparameter settings.

| Hyperparameters   | Characterization | NLPIR | NLPCC2014 |
|-------------------|------------------|-------|------------|
| η                 | Learning rate    | 0.01  | 0.01       |
| Epoch             | Number of iterations | 10  | 10         |
| Hidden            | Number of hidden units | 128 | 200        |
| P dropout         | Dropout          | 0.5   | 0.5        |
| λ                 | L2 coefficient   | le-3  | le-3       |
| Batch             | Batch size       | 64    | 64         |
| D                 | Word vector size | 200   | 200        |
| Dimension         | Word vector dimension | 200 | 200        |

Table 3: Comparison of experimental results.

| Model            | Pre  | Rec  | F1   | Pre  | Rec  | F1   |
|------------------|------|------|------|------|------|------|
| CNN              | 78.66| 94.44| 86.65| 70.40| 66.31| 68.29|
| CNN + SVM        | 86.32| 88.11| 87.21| 73.13| 71.30| 72.20|
| LSTM             | 81.94| 92.92| 87.09| 72.23| 72.56| 72.39|
| BiLSTM           | 87.60| 89.31| 88.45| 73.92| 71.17| 72.52|
| BM-ATT-BiLSTM    | 87.02| 90.57| 88.76| 75.57| 70.69| 73.05|

Scientific Programming 5
learning features and profound learning features to the short text sentiment analysis.

4.1. Summary. In a bidirectional LSTM based on emotional multichannel, shallow learning features and word embeddings and profound learning features of emotional attention are fused. Experiments show that the model integrates shallow learning features such as emotional part of speech, location information, and word dependence, and the classification effect is better than CNN and other models. In the profound learning model based on convolutional neural network, word embedding, emotional attention feature, and convolutional neural network feature are fused. Experiments show that feeding profound learning features into SVM can improve the performance of classification, and the model works best compared to other convolutional neural networks.

There are two main contributions of this paper.

(1) This paper puts forward a bidirectional LSTM memory network model based on emotional multichannel, which integrates shallow learning features and profound learning features at the feature level, so the model can heighten the expression ability of text semantic information and learn the potential emotion of the text. The test outcome is also multifeature prefusion, which is very helpful for sentiment climate analysis.

(2) This paper adopts CS-ATT-CNN and CS-ATT-TCNN to solve the problem of sentiment climate analysis of short texts. The model effectively combines machine learning and profound learning, the training time is short, and the model is better than the traditional convolution—Neural network classification model and TCNN model.

Theoretically, the semantic information of the LSTM model text is good, but from the practical point of view, the LSTM model still has deficiencies, and the BM-ATT-BiLSTM method in this paper combines the emotional part-of-speech features, location information features, and dependency features of words. As well as from the emotional attention feature, we can learn the underlying emotional regularity in the text so as to capture the important information that affects the emotional climate of the text. The BM-ATT-BiLSTM proposed in this paper is good, but there is one more factor that should be considered in the forget phylum, and it takes a lot of time to detect these meaningless communications can achieve better experimental results in a short time.

4.2. Prospect. Judging from the current situation and the experiments of this paper, there is still a long way to go in the analysis of sentiment climate of short texts. There are many challenges and opportunities, and there are many problems that deserve further study and improvement. The following two points can be used as the next step for research.

(1) This paper only mines text data and does not consider the author’s user attributes. If user attributes, user’s Weibo or online shopping comments, and posting time and other information are fully considered, the accuracy of text orientation analysis may be greatly improved. In the experiment, this paper only uses the data of Wikipedia as the corpus of the training word vector, and the data of Weibo and product reviews can be added as the expanded corpus of the training word vector in the future. Due to the complexity and diversity of the network environment and Chinese, some new words may not be recognized in the word segmentation stage, and there is a polysemy problem in Chinese. In the future, a new word dictionary can be constructed in combination with the context.

(2) There are some sarcastic sentences in Weibo and product reviews. This paper cannot identify the emotional polarity of such sentences well. Therefore, we can add sarcastic sentences to transform the model of this paper. If you combine the knowledge graph to do sentiment decomposition, consider each word as an entity in the knowledge graph, and form a many-to-many relationship between entities and entities, and it is very possible to dig out the emotional connection between words and words. In the next step, we will combine some actual Internet projects to verify more combined models to reflect their socio-economic value.

Data Availability

The data set can be accessed upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] X. Sun, C. Zhang, and F. J. Ren, “User emotion modelling and anomaly detection based on social media,” Chinese Journal of Information, vol. 32, no. 4, pp. 120–129, 2018.
[2] Y. Wang, “Stock market forecasting with financial micro-blog based on sentiment and time series analysis,” Journal of Shanghai Juotong University (Science), vol. 22, no. 2, pp. 173–179, 2017.
[3] C. Y. Liu, Research on sentiment mining technology of social network comments in box office prediction, Nanjing University of Posts and Telecommunications, Nanjing, China, 2019.
[4] W. Budiharto and M. Meiliana, "Prediction and analysis of Indonesia Presidential election from Twitter using sentiment analysis," Journal of Big Data, vol. 5, no. 1, p. 51, 2018.
[5] J. B. Bao, Sentiment analysis of online shopping platform product reviews based on machine learning, Harbin Institute of Technology, Harbin, China, 2019.
[6] K. Li, Analysis and research on the sentiment climate of commodity reviews based on expression skills, China University of Mining and Technology, Beijing, China, 2018.
[7] D. Y. Zhou, Research on sentiment analysis of commodity reviews based on convolutional neural network, Nanjing University of Posts and Telecommunications, Nanjing, China, 2019.
[8] R. Li, S. Sahu, and M. Schachner, "Phenelzine, a cell adhesion molecule L1 mimic small organic compound, promotes functional recovery and axonal regrowth in spinal cord-injured zebrafish," *Pharmacology Biochemistry and Behavior*, vol. 171, no. 1, pp. 30–38, 2018.

[9] W. Hong and M. Li, "A survey of research on text sentiment analysis methods," *Computer Engineering and Science*, vol. 41, no. 4, pp. 750–757, 2019.

[10] C. Fellbaum, "An electronic lexical database," *Library Quarterly Information Community Policy*, vol. 25, no. 2, pp. 292–296, 1998.

[11] H. Liu and P. Singh, "ConceptNet—a practical commonsense reasoning tool-kit," *BT Technology Journal*, vol. 22, no. 4, pp. 211–226, 2004.

[12] B. M. Shahshahani and D. A. Landgrebe, "The effect of unlabeled samples in reducing the small sample size problem and mitigating the Hughes phenomenon," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, no. 5, pp. 1087–1095, 1994.

[13] J. Tang, Z. Lu, and J. Su, "Progressive self-supervised attention learning for aspect-level sentiment analysis," in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pp. 557–566, Florence, Italy, July 2019.

[14] G. A. Ruz, P. A. Henríquez, and A. Mascaréño, "Sentiment analysis of Twitter data during critical events through Bayesian networks classifiers," *Future Generation Computer Systems*, vol. 106, no. C, pp. 92–104, 2020.

[15] Y. Wang, M. Huang, X. Zhu, and L. Zhao, "Attention-based LSTM for aspect-level sentiment classification," in *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pp. 606–615, Austin, Texas, November 2016.

[16] D. M. Hu, C. W. Zhu, and C. Hu, "Multilingual text sentiment analysis method with attention mechanism based on pre-training model," *Small microcomputer system*, vol. 41, no. 2, pp. 278–284, 2020.

[17] A. Onan, "Classifier and feature set ensembles for web page classification," *Journal of Information Science*, vol. 42, no. 2, pp. 150–165, 2016.

[18] Y. Xiao, X. Ding, and R. J. He, "Sentiment analysis of Chinese Weibo based on domain sentiment dictionary," *Electronic Design Engineering*, vol. 23, no. 12, pp. 18–21, 2015.

[19] X. D. Chen, *Analysis of Chinese Weibo sentiment climate based on sentiment dictionary*, Huazhong University of Science and Technology, Wuhan, China, 2012.

[20] L. X. Xie, M. Zhou, and M. S. Sun, "Multi-strategy Chinese microblog sentiment analysis and feature extraction based on hierarchical structure," *Chinese Journal of Information*, vol. 46, no. 1, pp. 73–84, 2012.

[21] T. T. Li and D. H. Ji, "Sentiment analysis of microblog based on multi-feature combination of SVM and CRF," *Research on Computer Application*, vol. 32, no. 4, pp. 978–981, 2015.

[22] K. Chen, B. Liang, and W. D. Ke, "Chinese microblog sentiment analysis based on multi-channel convolutional neural network," *Computer Research and Development*, vol. 55, no. 5, pp. 945–957, 2018.

[23] Z. Jia, Y. Lin, Y. Liu, and Z. J Jiao, "Refined nonuniform embedding for coupling detection in multivariate time series," *Physical Review*, vol. 101, no. 6, 2020.

[24] Z. Jia, J. Junyu, X. Zhou, and Y. Zhou, "Hybrid spiking neural network for sleep EEG encoding," *Science China Information Sciences*, vol. 65, 2022.

[25] Z. Jia, X. Cai, and Z. Jiao, "Multi-modal physiological signals based squeeze-and-excitation network with domain adversarial learning for sleep staging," *IEEE Sensors Journal*, vol. 22, 2022.

[26] Z. Jia, M. Gao, Z. Zhang et al., "Community environment and physical activity influence on rural residents' mental health in the COVID-19 containment," *International Journal of Environmental Health Research*, vol. 29, pp. 1–12, 2022.

[27] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Proceedings of the Advances in Neural Information Processing Systems*, pp. 3111–3119, Red Hook, NY, USA, December 2013.

[28] G. Paltoglou and M. Thelwall, "Twitter, MySpace, d," *ACM Transactions on Intelligent Systems and Technology*, vol. 3, no. 4, pp. 1–19, 2012.

[29] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: continual prediction with LSTM," *Neural Computation*, vol. 12, no. 10, pp. 2451–2471, 2000.

[30] Z. Jia, Y. Lin, J. Wang et al., "Multi-view spatial-temporal graph convolutional networks with domain generalization for sleep stage classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 1977–1986, 2021.