
Natural Language Processing: Classification of Web Texts Combined with Deep Learning

Chenwen Zhang

*School of Teacher Education, Hubei University of Arts and Science, Xiangyang,
Hubei 441053, China
E-mail: zhangcw_cw@hotmail.com*

Received 16 December 2024; Accepted 01 May 2025

Abstract

With the increasing number of web texts, the classification of web texts has become an important task. In this paper, the text word vector representation method is first analyzed, and bidirectional encoder representations from transformers (BERT) are selected to extract the word vector. The bidirectional gated recurrent unit (BiGRU), convolutional neural network (CNN), and attention mechanism are combined to obtain the context and local features of the text, respectively. Experiments were carried out using the THUCNews dataset. The results showed that in the comparison between word-to-vector (Word2vec), Glove, and BERT, the BERT obtained the best classification result. In the classification of different types of text, the average accuracy and F1 value of the BERT-BGCA method reached 0.9521 and 0.9436, respectively, which were superior to other deep learning methods such as TextCNN. The results suggest that the BERT-BGCA method is effective in classifying web texts and can be applied in practice.

Keywords: Natural language processing, deep learning, web text, text classification.

Journal of ICT Standardization, Vol. 13_1, 25–40.

doi: 10.13052/jicts2245-800X.1312

© 2025 River Publishers

1 Introduction

With the continuous progress of the Internet, more and more data are generated in daily life and work in the form of images, speech, and texts. Text data is quite an important component of web data, which may come from chat records, movie reviews, and network news, and contains a lot of useful information. Text classification is a technology of natural language processing (NLP) [1], which refers to the use of information extracted from text data to divide text into different types [2]. It can provide support for information retrieval, news filtering, medical diagnosis, etc. [3]. As the data generation speed of web text is so fast, the manual annotation method makes it more and more difficult to meet the growing demand for text classification. In the case of the continuous development of computers, the automatic classification of text by computers has become a key issue [4]. Li et al. [5] designed a dual-view graph convolutional network (GCN) for classifying multi-label texts and found through experiments that the method successfully balanced efficiency and performance. Ramadhani et al. [6] combined the term frequency-inverse document frequency method and support vector machine to classify Instagram captions and obtained high accuracy. Min et al. [7] designed a topic-aware cosine graph convolutional neural network (CNN) to classify short texts. Experiments on eight STC datasets found that the method had better accuracy and macro-F1 score. Shaikh et al. [8] compared different machine learning methods using a dataset containing 11.8k patient descriptions of the 20 most common diseases in text classification in evidence-based medicine. They found that the random forest tree had the highest overall accuracy of 83%. Compared with ordinary texts, web text has a large amount of data, wide sources, and various forms. Therefore, the classification of web text is more complex, and the performance of existing text classification methods still needs to be improved. Deep learning can automatically extract features and is widely used in NLP tasks. Therefore, this paper mainly analyzes how to use deep learning methods to design a high-performance web text classification method. In the first section of this paper, the designed text classification method is introduced. Bidirectional encoder representations from transformers (BERT) word vector representation, bidirectional encoder representations from transformers (BiGRU), and CNN were combined to establish a BERT-BGCA model. The second section is the experiment and analysis part. The classification performance of the model was compared and verified on the THUCNews dataset, which highlighted the classification performance of the model. The third section is the discussion and conclusion

part. A simple discussion and summary of the experimental results are carried out, and the research limitations and future research directions are pointed out. The research in this paper combines several deep learning methods to improve the classification performance and optimize the traditional deep learning method, providing some theoretical basis for the further application of deep learning in network text classification and offering some references for the performance optimization of deep learning models.

2 Classification of Web Text Based on Deep Learning

2.1 Word Vector Representation

Word-to-vector (Word2vec) is a classical word vector representation approach [9] that has extensive applications in feature extraction of NLP tasks [10]. It includes continuous bag-of-words (CBOW) and skip-gram, which can effectively reflect the semantic relationship between words. However, it cannot solve the polysemy problem, nor can it express contextual information.

Glove carries out word vector representation based on word co-occurrence frequency [11]. Compared with Word2vec, Glove can better capture the global semantic information of words. It is efficient in large-scale corpus training. BERT is a popular word vector representation model [12]. Based on a transformer model, it can effectively capture context information and sentence relationships and train richer text representation, with superior performance in NLP tasks. Therefore, BERT is also chosen in this paper to implement the word vector representation before the classification of web text. For text $X = \{x_1, x_2, \dots, x_n\}$, the word embedding representation is obtained by BERT training:

$$S = BERT(X). \quad (1)$$

Then, the word vector representation containing context information, i.e., the final hidden layer output of the sentence, is obtained through the CLS vector:

$$C = BERT(CLS). \quad (2)$$

2.2 Feature Extraction

Among deep learning methods, the recurrent neural network (RNN) has a favorable effect in processing sequence data [13], especially in mining the

timing and semantic information of data, and has extensive applications in NLP tasks [14]. The gated recurrent unit (GRU) is a variant of RNN [15], which has a better effect than RNN in capturing long-term dependencies. GRUs employ two gating mechanisms:

Reset gate: this determines how much information needs to be forgotten:

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]). \quad (3)$$

Update gate: this controls the ratio of newly input information to past status information:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (4)$$

where h_{t-1} is the hidden state of the previous moment, x_t is the input of the current moment, and W_r and W_z are weights. According to r_t and z_t , candidate hiding state \tilde{h}_t can be calculated:

$$\tilde{h}_t = \tanh(W_{\tilde{h}} x_t + r_t \odot U_{\tilde{h}} h_{t-1}), \quad (5)$$

where $W_{\tilde{h}}$ and $U_{\tilde{h}}$ are the weight of x_t and h_{t-1} , and \odot refers to the multiplication of elements. The final hidden state is:

$$h_t = z_t \odot \tilde{h}_t + (1 - z_t) \odot h_{t-1}. \quad (6)$$

BiGRU adds a reverse structure on the basis of the forward structure of GRU, so that the model can learn the information of the previous and following texts and obtain better feature extraction effect [16]. The forward structure calculates from $t = 1$, and the hidden state obtained is:

$$h_t^f = GRU(h_{t-1}^f, x_t). \quad (7)$$

The reverse structure calculates from $t = T$, and the hidden state obtained is:

$$h_t^b = GRU(h_{t+1}^b, x_t). \quad (8)$$

BiGRU concatenates the two hidden states to obtain the output at time t :

$$h_t = [h_t^f, h_t^b]. \quad (9)$$

The final output of BiGRU is:

$$y_t = W_y h_t, \quad (10)$$

where W_y is the weight.

What BiGRU learns is the global sequence feature of text, and the phrase-level text feature also plays an essential role in the classification of web text. Therefore, this paper chooses a CNN to learn the local feature of word vectors [17]. CNN uses convolution kernels to extract features. It is assumed that there is input vector $T \in R^{d \times m}$ and convolution kernel $K \in k^{d \times h}$ (d is the vector dimension and h is the width of the convolution kernel). The output convolution feature is:

$$g_i = \tanh(\langle K, T \rangle + b), \quad (11)$$

where $\langle \cdot \rangle$ denotes the convolution calculation process and b is the bias. After convolution, the output of CNN is obtained through pooling operation:

$$y_i = \max_i g_i. \quad (12)$$

2.3 Attention Mechanism and Classification Model

For the features learned from BiGRU and CNN, in order to highlight the role of key information, this paper adds the attention mechanism after both BiGRU and CNN [18]. The attention mechanism mainly includes query vector Q , key vector K , and value vector V . In calculation, the similarity between Q and each K is calculated:

$$Sim(Q, K_i) = \frac{Q \cdot K_i}{\|Q\| \cdot \|K_i\|}. \quad (13)$$

The softmax function is used for normalization processing to obtain the attention weight:

$$a_i = softmax(Sim) = \frac{e^{Sim_i}}{\sum_{i=1}^{L_s} e^{Sim_i}}. \quad (14)$$

Then, the weighted summation is performed:

$$Attention = \sum_{i=1}^{L_s} a_i \cdot V_i. \quad (15)$$

The attention mechanism is added after BiGRU:

$$M_g = \tanh(h_t), \quad (16)$$

$$a_g = softmax(W_g M_g), \quad (17)$$

$$S_g = h_t a_g, \quad (18)$$

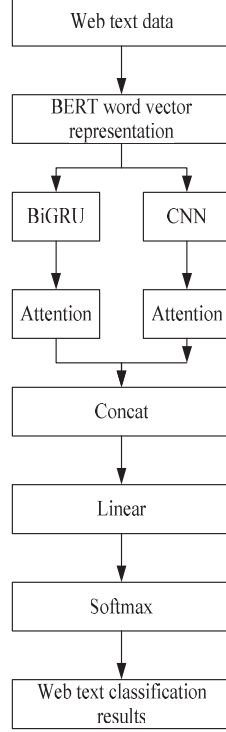


Figure 1 BERT-BGCA web text classification model.

where h_t is the hidden output of BiGRU, W_g is the weight, a_g is the attention weight coefficient, and S_g is the output after the attention mechanism.

Similarly, a CNN is added after the attention mechanism:

$$M_c = \tanh(y_i), \quad (19)$$

$$a_c = \text{softmax}(W_c M_c), \quad (20)$$

$$S_c = y_i a_c, \quad (21)$$

where y_i is the output of CNN, W_c is the weight, a_c is the attention weight coefficient, and S_c is the output after the attention mechanism.

The structure of the BERT-BGCA web text classification model designed in this paper is shown in Figure 1.

As shown in Figure 1, the word vectors obtained after BERT training are input into the BiGRU-attention and CNN-attention respectively for feature

extraction, and the obtained features are concatenated:

$$S = \text{cat}(S_g, S_c). \quad (22)$$

Then, linear transformation is performed through a linear layer:

$$O = W^T S + b_O, \quad (23)$$

where W^T is the training parameter matrix and b_O is the bias. Finally, the obtained output vector O is input into the softmax layer to obtain the classification result of the web text.

3 Experiment and Analysis

3.1 Experimental Setup

The experimental environment is presented in Table 1.

The experimental dataset was from THUCNews [19], and the texts were from the historical data of Sina news. Ten types of news titles (entertainment, games, sports, education, real estate, stocks, current politics, science and technology, society, finance and economics) were selected for the experiment. Some text examples are shown in Table 2. Five thousand titles were selected for each category, totaling 50,000 titles. The training, verification, and test sets were divided into 8:1:1.

After repeated experiments, the parameters of the BERT-BGCA model were determined (Table 3).

3.2 Evaluation Indicators

Based on a confusion matrix, the accuracy and F1 value were used to evaluate the classification effect of web texts:

- (1) Accuracy (the proportion of samples that are correctly classified, e.g., the text belonging to the category of entertainment is correctly classified

Table 1 Experimental environment

| | |
|-------------------------------------|---------------------|
| Operating system | Windows 11 |
| Central processing unit | I5-11260H@2.60GHz |
| Graphics processing unit | GeForce RTX 2080 Ti |
| Memory | 16 GB |
| Programming language | Python3.9 |
| Deep learning framework | PyTorch1.11.0 |
| Compute unified device architecture | 11.4 |

Table 2 Examples of texts in the THUCNews dataset

| Category | Example |
|-----------|---|
| Reality | Haitang Community near the East 5th Ring 230–290 m ² two-bedroom near-ready apartments 2% discount available |
| Finance | Golden certificate consultant: What does roller-coaster market volatility signify? |
| Education | The registration for the CET-4 and CET-6 exams in Shanghai in the first half of 2010 should be completed before 8 April. |
| Science | “Mobile Wallet” made its debut at the Science and Technology Expo. |
| Society | Reporters paid a return visit to the “Coke Boy” in the earthquake: He will be invited to visit the United States. |
| Game | The smoke of war has risen again in the siege warfare and vassal warfare of “Red Cliff Online”. |

Table 3 BERT-BGCA parameter settings

| Parameter | Value |
|----------------------|--------|
| Hidden size of BiGRU | 128 |
| Embedding_dim | 300 |
| Learning_rate | 0.0001 |
| Batch_size | 128 |
| Dropout | 0.5 |
| Optimizer | Adam |

to the category of entertainment by the model):

$$Acc = \frac{TP + TN}{TP + FN + FP + TN}. \quad (24)$$

(2) F1 value (the harmonic average of recall rate and precision)

$$F_1 = F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (25)$$

$$Precision = \frac{TP}{TP + FP} \quad (26)$$

$$Recall = \frac{TP}{TP + FN} \quad (27)$$

where TP refers positive samples that are classified as positive, TN refers to negative samples that are classified as negative, FP refers to negative samples that are classified as positive, and FN refers to positive samples that are classified as negative.

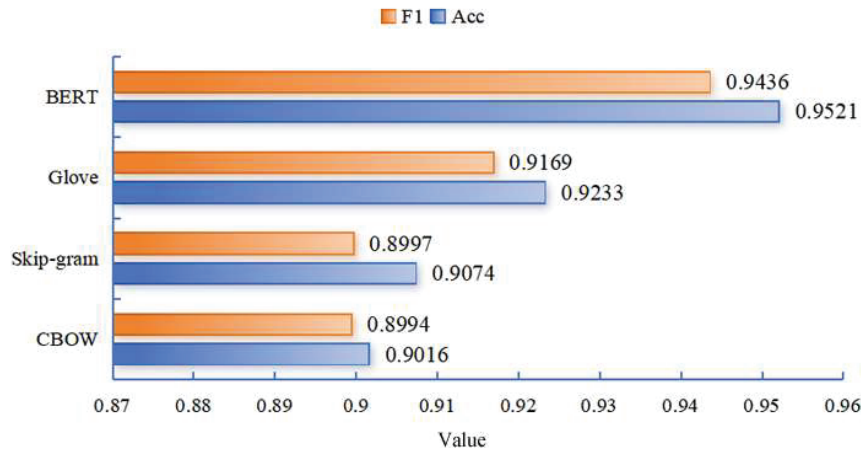


Figure 2 Comparison of different word vector representation approaches.

3.3 Analysis of Results

The selected word vector representation method was verified. The BGCA model was used to classify web texts, and the classification effect of various word vector representation approaches was compared (Figure 2).

It can be found that there was a small gap between the two methods in Word2vec in web text classification. The Acc values of CBOW and skip-gram were 0.9016 and 0.9074, respectively, and the F1 values were 0.8994 and 0.8997, respectively. In comparison, the results of web text classification obtained by Glove were significantly improved, with an Acc value of 0.9233 and an F1 value of 0.9169, indicating that Glove was superior to the traditional Word2vec method in word vector representation. Moreover, the Acc value obtained by BERT was 0.9521, which increased by 0.0288 compared with Glove, and the F1 value was 0.9436, which increased by 0.0267 compared with Glove. The results suggested the superiority of BERT selected as the word vector representation method.

It can be found from Table 4 that, in terms of text classification of different categories, the BERT-BGCA method achieved 0.9 higher Acc and F1 values. The average Acc and F1 values were 0.9521 and 0.9436, respectively. It achieved the best classification results for sports, with an Acc value of 0.9864 and an F1 value of 0.9832. It achieved the worst classification results for stocks, with an Acc value of 0.9123 and an F1 value of 0.9112. The different classification effects may be caused by differences in wording and the style of texts of different categories.

Table 4 Classification results of BERT-BGCA for different categories of texts

| | Acc | F1 |
|---------------|--------|--------|
| Finance | 0.9334 | 0.9312 |
| Realty | 0.9762 | 0.9564 |
| Stocks | 0.9123 | 0.9112 |
| Education | 0.9771 | 0.9701 |
| Science | 0.9226 | 0.9169 |
| Society | 0.9455 | 0.9433 |
| Politics | 0.9221 | 0.9251 |
| Sports | 0.9864 | 0.9832 |
| Game | 0.9737 | 0.9625 |
| Entertainment | 0.9717 | 0.9361 |
| Mean | 0.9521 | 0.9436 |

In order to further demonstrate the performance of the proposed method, it was compared with other deep learning-based text classification methods, including:

- (1) TextCNN [20]: Use convolution kernels of different sizes to extract N-gram features and then implement text classification through the fully connected layer,
- (2) TextRNN [21]: Text modelling is implemented using an RNN, and it performs well in the processing of long text sequences,
- (3) TextGCN [22]: View the text as a graph structure, with word vectors as nodes and semantic relations as edges, and use a graph convolutional network (GCN) to implement text modelling,
- (4) BiGRU [23]: Vectorize the text and learn the context features in the BiGRU network,
- (5) BiGRU-Att [24]: The text features are extracted using BiGRU, the weights are assigned in combination with the attention layer, and the text classification is performed in the softmax layer.

It can be seen from Figure 3 that the TextCNN, TextRNN, and TextGCN methods all exhibited poor performance in web text classification, and the Acc and F1 values were both below 0.9, indicating that these models were relatively simple in learning text features. Neither TextCNN nor TextRNN learned the long-term dependency of text. The semantic information covered by TextGCN was also insufficient. In other words, context dependence and the extraction of local information were insufficient. In the comparison of the BiGRU-based models, it was found that adding the attention mechanism

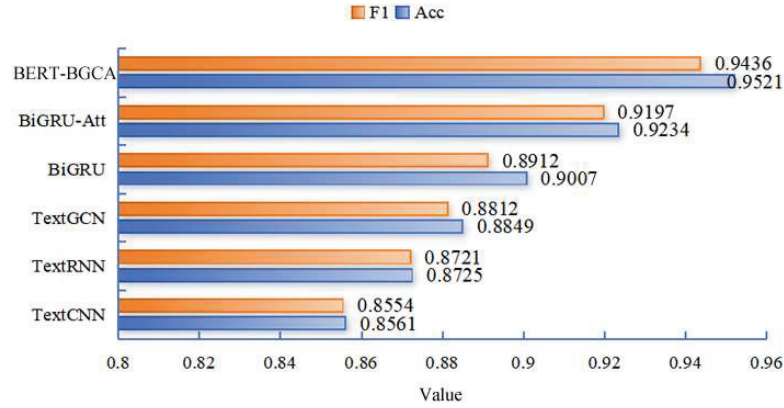


Figure 3 Comparison with other deep learning methods.

could effectively enhance the classification ability of the model. The Acc value of the BiGRU-Att method was 0.9234, which increased by 0.0227 compared with the BiGRU method, and the F1 value was 0.9197, which increased by 0.0285 compared with the BiGRU method. These results verified the importance of the attention mechanism. Then, in the comparison between the BERT-BGCA and BiGRU-Att methods, the CNN-attention added in the proposed method further extracted the local features of the text. The obtained Acc value was 0.9521, which was 0.0287 higher than that of the BiGRU-Att method, and the F1 value was 0.9436, which was 0.0239 higher than that of the BiGRU-Att method. These results verified the reliability of the BERT-BGCA method in the web text classification.

4 Discussion and Conclusion

Text classification is not only an important part of NLP tasks but also has certain practical significance. It has application value in scenarios such as information filtering and news recommendation. The rapid development of the network has accumulated a large amount of online text. The classification of online text is of great significance for network content management and information retrieval recommendation. Therefore, this paper conducted research on the classification of web text.

Compared with traditional machine learning methods, deep learning models have better performance in NLP tasks. Therefore, this paper chose deep learning to design the classification method of web text, combined

BERT, BiGRU, and CNN to establish the BERT-BGCA model, and verified the classification performance of this model using the THUCNews dataset. Compared with word2vec and Glove, when BERT was used as the word vector representation method, the BGCA model obtained better Acc and F1 values in text classification, which highlighted the reliability of using BERT; in the text classification of different categories, the Acc and F1 values of the BERT-BGCA model were both above 0.9, indicating that the model can handle different types of text well. Then, compared with other classification models, the BERT-BGCA model also showed the best results. The BERT-BGCA model used both BiGRU and CNN to extract features and was added with the attention mechanism, making full use of the global and local features of the text, thus achieving good performance in the classification of web text.

According to the comprehensive experimental results, the BERT-BGCA model showed excellent classification performance for the THUCNews dataset. It can classify different types of texts relatively accurately and be further applied in actual web text classification, providing a new method for application scenarios such as public opinion analysis and automatic summarization.

However, this study also has some limitations. For instance, experiments were only conducted on Chinese texts, i.e., the used datasets were not rich enough, and only the accuracy rate of the BERT-BGCA model was considered, without an in-depth understanding of the efficiency. Therefore, in future work, the quantity and diversity of datasets will be further expanded, and the operational efficiency of the model will be studied.

Funding

This study was supported by Chinese National Education Science Plan 2022 Key project of the Ministry of Education “Multidimensional Classroom Teaching Evaluation Research Based on Artificial Intelligence” (No. DCA220444).

References

- [1] C. W. Chen, S. P. Tseng, J. F. Wang, ‘Outpatient Text Classification System Using LSTM’, *J. Inf. Sci. Eng.*, vol. 37, pp. 365–379, 2021. DOI: 10.6688/JISE.202103_37(2).0006.

- [2] T. Zhou, Y. Wang, X. Zheng, ‘Chinese text classification method using FastText and term frequency-inverse document frequency optimization’, *J. Phys.: Conf. Ser.*, vol. 1693, no. 1, pp. 1–6, 2020. DOI: 10.1088/1742-6596/1693/1/012121.
- [3] A. Kaddour, N. Zellal, L. Sayad, ‘Improving text classification using text summarization’, *2022 2nd International Conference on New Technologies of Information and Communication (NTIC)*, pp. 1–8, 2022. DOI: 10.1109/NTIC55069.2022.10100492.
- [4] H. Zhou, ‘Research of Text Classification Based on TF-IDF and CNN-LSTM’, *J. Phys.: Conf. Ser.*, vol. 2171, no. 1, pp. 1–8, 2022.
- [5] X. Li, B. You, Q. Peng, S. Feng, ‘Dual-view graph convolutional network for multi-label text classification’, *Appl. Intell.*, vol. 54, no. 19, pp. 9363–9380, 2024. DOI: 10.1007/s10489-024-05666-w.
- [6] P. P. Ramadhani, S. Hadi, ‘Text classification on the Instagram caption using support vector machine’, *J. Phys.: Conf. Ser.*, vol. 1722, no. 1, pp. 1–7, 2021. DOI: 10.1088/1742-6596/1722/1/012023.
- [7] C. Min, Y. Chu, H. Lin, B. Wang, L. Yang, B. Xu, ‘Topic-aware cosine graph convolutional neural network for short text classification’, *Soft Comput.*, vol. 28, no. 13–14, pp. 8119–8132, 2024. DOI: 10.1007/s00500-024-09679-y.
- [8] S. Shaikh, M. Y. Khan, M. S. Nizami, ‘Using Patient Descriptions of 20 Most Common Diseases in Text Classification for Evidence-based Medicine’, *2021 Mohammad Ali Jinnah University International Conference on Computing (MAJICC)*, Karachi, Pakistan, pp. 1–8, 2021. DOI: 10.1109/MAJICC53071.2021.9526252.
- [9] S. Paliwal, A. K. Mishra, M. N. N. Senthilkumar, ‘XGBRS Framework Integrated with Word2Vec Sentiment Analysis for Augmented Drug Recommendation’, *Comput. Mater. Con.*, vol. 72, no. 3 Pt.2, pp. 5345–5362, 2022.
- [10] R. Indira, W. Maharani, ‘Personality Detection on Social Media Twitter Using Long Short-Term Memory with Word2Vec’, *2021 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT)*, pp. 64–69, 2021. DOI: 10.1109/COMNETSAT53002.2021.9530820.
- [11] P. Gupta, I. Roy, G. Batra, A. K. Dubey, ‘Decoding Emotions in Text Using GloVe Embeddings’, *2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*, pp. 36–40, 2021. DOI: 10.1109/ICCCIS51004.2021.9397132.

- [12] T. Saha, S. Ramesh Jayashree, S. Saha, P. Bhattacharyya, ‘BERT-Caps: A Transformer-Based Capsule Network for Tweet Act Classification’, *IEEE T. Comput. Soc. Sy.*, vol. 7, no. 5, pp. 1168–1179, 2020. DOI: 10.1109/TCSS.2020.3014128.
- [13] R. Sabitha, P. Poonkodi, M. S. Karthik, S. Karthik, ‘Premature Infant Cry Classification via Deep Convolutional Recurrent Neural Network Based on Multi-class Features’, *Circ. Syst. Signal Pr.*, vol. 42, no. 12, pp. 7529–7548, 2023.
- [14] J. T. Oh, L. S. Yong, ‘A Fuzzy-AHP-based Movie Recommendation System with the Bidirectional Recurrent Neural Network Language Model’, *J. Digit. Converg.*, vol. 18, pp. 525–531, 2020. DOI: 10.14400/JDC.2020.18.12.525.
- [15] Y. Yevnin, S. Chorev, I. Dukan, Y. Toledo, ‘Short-term wave forecasts using gated recurrent unit model’, *Ocean Eng.*, vol. 268, pp. 1–8, 2023. DOI: 10.1016/j.oceaneng.2022.113389.
- [16] Q. Qian, J. Yu, H. Zhan, R. Wang, ‘A novel DL-BiGRU multi-feature fusion and deep transfer learning based modeling approach for quality prediction of injection molded products using small-sample datasets’, *J. Manuf. Process.*, vol. 120, pp. 272–285, 2024. DOI: 10.1016/j.jmapro.2024.04.030.
- [17] S. Lee, J. S. Lee, ‘Experimental evaluation of convolutional neural network-based inter-crystal scattering recovery for high-resolution PET detectors’, *Phys. Med. Biol.*, vol. 68, 2023. DOI: 10.1088/1361-6560/acacb.
- [18] G. L. De la Peña Sarracén, P. Rosso, ‘Correction to: Offensive keyword extraction based on the attention mechanism of BERT and the eigenvector centrality using a graph representation’, *Pers. Ubiquit. Comput.*, vol. 28, no. 2, pp. 443–444, 2024. DOI: 10.1007/s00779-024-01791-y.
- [19] M. S. Sun, J. Y. Li, Z. P. Guo, Y. Zhao, Y. B. Zheng, X. C. Si, Z. Y. Liu, ‘THUCTC: an efficient Chinese text classification toolkit’, 2016.
- [20] Y. Kim, ‘Convolutional Neural Networks for Sentence Classification’, *Conference on Empirical Methods in Natural Language Processing*, vol. 2014, pp. 1746–1751, 2014. DOI: 10.3115/v1/D14-1181.
- [21] P. Liu, X. Qiu, X. Huang, ‘Recurrent Neural Network for Text Classification with Multi-Task Learning’, *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, vol. 2016, pp. 2873–2879, 2016. DOI: 10.48550/arXiv.1605.05101.

- [22] L. Yao, C. Mao, Y. Luo, ‘Graph convolutional networks for text classification’, *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, pp. 7370–7377, 2019.
- [23] Y. Liu, C. Liu, L. Wang, Z. Chen, ‘Chinese Event Subject Extraction in the Financial Field Integrated with BiGRU and Multi-head Attention’, *J. Phys.: Conf. Ser.*, vol. 1828, no. 1, pp. 1–8, 2021. DOI: 10.1088/1742-6596/1828/1/012032.
- [24] W. Wang, Y. X. Sun, Q. J. Qi, X. F. Meng, ‘Text sentiment classification model based on BiGRU-attention neural network’, *Appl. Res. Comput.*, vol. 36, no. 12, pp. 3558–3564, 2019.

Biography



Chenwen Zhang holds a Master’s degree of science. He is an associate professor. He graduated from Central China Normal University in 2009. He is working in Hubei University of Arts and Science. His research interests include digital learning and teacher career development.

