Proceedings of Machine Learning Research vol 278:1–11, 2025 2025 2nd International Conference on Machine Learning and Intelligent Computing

Research on the Intelligent Screening Algorithm for College Faculty Recruitment Based on BiGRU-attention

Shirui Gai 18943900739@163.COM
Human Resources Development and Management Center, Jilin Animation Institute, Changchun, China

Editors: Nianyin Zeng, Ram Bilas Pachori and Dongshu Wang

Abstract

With the rapid development of higher education, university faculty recruitment is facing increasing pressure to resume screening. Traditional screening methods are inefficient and highly subjective, making it difficult to meet the needs of universities for outstanding talent. This study aimed to construct an intelligent screening model based on the Bidirectional Gated Recurrent Unit (BiGRU) and attention mechanism to improve the efficiency and accuracy of university faculty recruitment. First, a large amount of university faculty recruitment resume data were collected and preprocessed to construct a high-quality dataset. Subsequently, the BiGRU model was introduced to deeply mine the text features of the resumes. By taking advantage of its ability to effectively process sequential data and capture contextual information, the model enhances the ability to extract key information from resumes. Simultaneously, combined with the attention mechanism, the model could focus on important features, further improving the screening accuracy. The experimental results showed that the constructed BiGRU-attention model performs excellently in the task of screening university faculty recruitment resumes. Compared with traditional methods, it significantly improved indicators such as the accuracy and recall rate. It could provide more efficient and intelligent decision support for university recruitment work, help universities select outstanding teaching talents that better meet job requirements, and promote the construction of teaching staff in higher education.
Keywords: Bidirectional gated recurrent unit; Attention mechanism; Screening

1. Introduction

In the current era of the rapid development of higher education, the quality of the university faculty team is directly linked to a university's teaching standards, research capabilities, and the quality of talent cultivation. With the continuous expansion of university enrollment and the deepening of educational reforms, the number of application materials received in university faculty recruitment each year has been growing exponentially. The traditional screening method for faculty recruitment mainly depends on manual resume review. This method is not only inefficient but also highly susceptible to subjective factors. As a result, it is extremely challenging to swiftly and precisely select outstanding talents who truly meet the job requirements from the massive amount of application information.

At the same time, the rapid development of artificial intelligence technology has provided new ideas for solving this difficult problem. As an important algorithm in the field of deep learning, the convolutional neural network have achieved remarkable results in many fields, such as image recognition, natural language processing, and data mining. In the existing research on intelligent screening of university faculty recruitment, the in-depth mining and utilization of data are insufficient, and the accuracy and intelligence level of the models still need to be improved. This study conducted a matching calculation of unstructured texts in resumes based on the BiGRU-attention algorithm, broke through the limitations of traditional screening methods, realized the intelligence and refinement of recruitment screening, improved the accuracy and scientific nature of recruitment, selected more outstanding and suitable teachers for universities, and thus promoted the optimization and development of the university faculty team.

2. Related work

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

In the process of university faculty recruitment, the construction of an intelligent screening model is an important means to improve the efficiency and accuracy of recruitment. In recent years, with the development of deep learning technology, techniques such as BiGRU and the attention mechanism have been widely applied in the fields of natural language processing and data mining. These technologies excel in handling complex data and extracting key features, and thus have been introduced into the intelligent screening systems for faculty recruitment.

Shi (2022) used the BiGRU model to understand the semantic and contextual relationships in the texts of resumes and application materials for teacher recruitment. The introduction of the attention mechanism further enhances the performance of the model. The attention mechanism can dynamically assign the importance of different features, enabling the model to focus on more crucial information. Kang (2022) introduced the attention mechanism in the scenario of teacher recruitment to identify the most important parts in applicants' resumes, such as educational background, research achievements, and teaching experience, thus improving the accuracy of the screening process. Nie et al. (2022); Zhao et al. (2020) combined deep - learning and artificial intelligence technologies to develop an intelligent screening system to optimize the recruitment process. By analyzing a large amount of historical recruitment data, the system can identify the key factors that influence recruitment decisions and adjust the screening criteria and strategies accordingly. This data - driven approach not only improves the efficiency of recruitment but also helps universities better identify and attract outstanding talents. In practical applications, the construction of an intelligent screening system needs to take into account various factors, including data pre - processing, model training and optimization, as well as system deployment and maintenance. Zhang et al. (2024) designed and implemented a sentiment analysis model for Weibo comment texts based on the Bi-LSTM (Bidirectional Long Short-Term Memory) and the attention mechanism. The Bi-LSTM model learns the context features, and the attention mechanism learns the key features. Xu et al. (2023) utilized the BiGRU-Attention (Bidirectional Gated Recurrent Unit-Attention) model to analyze the emotions of graduate scholars in online academic forums and to excavate the hidden academic sentiments. Han et al. (2023) proposed the XLNet-BiGRU-Att model for text sentiment recognition. This model combines XLNet, the bidirectional recurrent unit, and the attention mechanism. Zhen and Xiaoxuan (2020); Jeng et al. (2020) showed that through reasonable design and implementation, these systems can significantly improve the efficiency and quality of university teacher recruitment and provide strong support for talent selection in educational institutions.

In conclusion, the intelligent screening model based on BiGRU-attention has broad application prospects in university faculty recruitment. By combining advanced deep learning technologies and big data analysis methods, these models can effectively enhance the intelligence level of the recruitment process and provide important technical support for universities in talent introduction (Cheng et al., 2022).

3. Unstructured text matching calculation based on BiGRU-attention

Resumes contain a great deal of textual data information, such as name, phone number, email address, job application intention, address, work experience, teaching experience, etc. The recruitment information module also contains a large amount of textual information, such as company name, nature of work, work location, name of the position being recruited, job requirements, job description, etc. From this, it can be seen that the textual information in resumes and recruitment can be divided into three categories: (1) Individual information text: name, email address, phone number, company name, etc. (2) Structured text: the location of the job seeker, type of job position applied for, expected salary, expected nature of work, educational background, years of work experience, major, work location, type of position being recruited, salary, nature of work, educational requirements, work experience requirements, major requirements, etc. (3) Unstructured text: work experience, project experience, job description, job requirements, etc.

In this paper, the BiGRU-Attention model was constructed to process the unstructured text data such as work experience and project experience in resumes, as well as job descriptions and job requirements in recruitment information. The corpus of this paper was jointly composed of the work experience and project experience in resume texts, and the job descriptions and job requirements in recruitment information texts. After performing data preprocessing operations such as word segmentation and stop word removal on them, word vectors were trained through the word embedding technology in Word2vec. Then, the dataset of vectorized text data pairs of resumes and recruitment information of positive and negative samples was fed into the BiGRU-Attention matching model for training and iteration. Finally, the resume and recruitment information matching pairs in the test set were fed into the model for matching calculation to obtain the corresponding matching values.

3.1. Method for matching the similarity of resumes

In the resumes and recruitment information, the most important unstructured texts are the work experience and professional skills in resume information texts, as well as the job descriptions and job requirements in recruitment information texts. However, these unstructured texts could not be directly used for matching. Therefore, after further processing using the word embedding technology of Word2vec, the calculation of the model matching value was carried out. Before entering the Word2vec word embedding training, operations of word segmentation and stop word removal were also required. In this paper, the jieba word segmentation technology in the Python library and the stop word list of Harbin Institute of Technology were used to preprocess the above unstructured texts. In this way, the sentences in the unstructured texts could be represented in the form of words and separated by spaces. After such processing, the words in the unstructured texts could be highlighted, and the density of keywords could be increased. In the subsequent training, the word vectors were more accurate without being interfered by stop words.

In the training of word vectors, this paper used the Word2vec model in the gensim. models library of Python to train word vectors. The Skip-gram algorithm was adopted for training in this

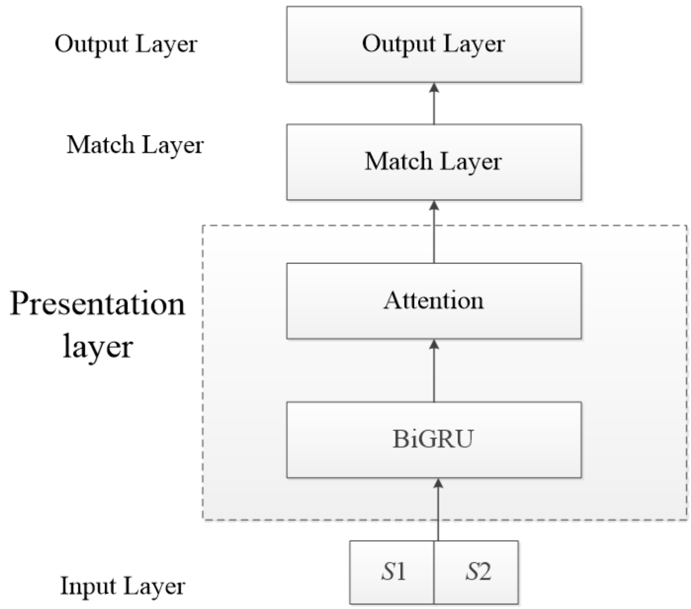

Figure 1: Network Model Diagram of BiGRU-attention

paper, and the window size was set to 5. Among the methods of hierarchical *Softmax* and negative sampling, hierarchical *Softmax* training was adopted. The dimension of the word vectors was set to 200.

3.2. Text matching algorithm based on BiGRU-attention matching model

The text matching algorithm based on the BiGRU-attention model was a deep learning model that combined the Bidirectional Gated Recurrent Unit (BiGRU) and the attention mechanism. It was used to address text matching tasks, such as question-answering matching, semantic similarity calculation, information retrieval, and so on. This model captured the contextual information of the text and utilizes the attention mechanism to focus on the key parts, thereby improving the accuracy of matching.

Therefore, in this paper, the BiGRU-attention model, which combines the BiGRU model and the Attention mechanism, was adopted for semantic training of the text. Finally, the matching value between the resume and the recruitment text was obtained through the similarity calculation in the matching layer. The structure of the matching model was shown in Figure 1.

(1) Input layer

The input consists of two text sequences to be matched, which were usually represented as sequences of word vectors. Suppose the input texts were $S1 = \{w_{11}, w_{12}, \ldots, w_{1n}\}$ and $\mathrm{S2} = \{w_{21}, w_{22}, \ldots, w_{2\mathrm{m}}\}$, where w_{ij} was the word vector of the j-th word in the i-th text.

(2) Bidirectional GRU layer

The Bidirectional Gated Recurrent Unit (BiGRU) was used to encode the input texts and capture the contextual information of the texts. For each text sequence, the BiGRU generated hidden states in two directions. Forward GRU: It processed the sequence from left to right and generates hidden states $\vec{h}_{ij}$. Backward GRU: It processed the sequence from right to left and generates hidden states

$\overleftarrow{h}_{ij}$. The forward and backward hidden states are concatenated to obtain the final hidden state at each time step.

$$h_{ij} = [\overrightarrow{h}_{ij}; \overleftarrow{h}_{ij}] \tag{1}$$

Finally, the hidden states of texts S_1 and S_2 were $H1 = \{h_{11}, h_{12}, \ldots, h_{1\mathrm{n}}\}$ and H2 $= \{h_{21}, h_{22}, \ldots, h_{2\mathrm{m}}\}$.

(3)Attention mechanism

The attention mechanism was employed to capture the interactive information between two texts and calculate the contribution of each word to the other text. For texts S_1 and S_2, an attention weight matrix A was computed, where A_{ij} represented the correlation between the i-th word of S_1 and the j-th word of S_2. The calculation of attention weights typically utilized dot-product attention or additive attention:

$$A_{ij} = softmax(h_{1i}^T W h_{2j}) \tag{2}$$

where W was a learnable weight matrix.

Through the attention weight matrix, the weighted representations of texts and were generated:

$$\widetilde{H}_1 = AH_2, \widetilde{H}_2 = A^T H_1 \tag{3}$$

(4)Match layer

Concatenate the original hidden state and the weighted representation to obtain the enhanced text representation:

$$M_1 = [H_1; \widetilde{H}_1], M2 = [H_2; \widetilde{H}_2] \tag{4}$$

Perform pooling operations on M1 and M2to obtain fixed-length vector representations v1 and v2. For the resume information vJ and the recruitment information vZ obtained in this article, then calculate the matching value between the two through the cosine similarity, and the calculation is as follows:

$$sim_{un}(J, Z) = \cos ine(v_J, v_Z) = \frac{v_J^T \cdot v_Z}{\|v_J\| \times \|v_Z\|} \tag{5}$$

Combine it with the matching value of the structured text to obtain the comprehensive similarity calculation, which is expressed as:

$$u = \beta_1 sim_{un} + \beta_2 sim_{st} \tag{6}$$

Among them: β_1 and β_2 represent the weight relationships of the unstructured text sim_{un} and the structured text sim_{st} respectively.

(5)Output layer

Obtain the final output through the fully connected layer and the softmax function:

$$y = \ softmax(w_s u + b_s) \tag{7}$$

Among, w_s is the weight matrix and b_s is the bias term. The probabilities of the calculation results obtained by using the softmax function in the above formula are compared with the labels of the original sample data to optimize the loss function, and its loss value Loss is as follows:

$$Loss = -\sum_j \hat{y}^{(j)} \log y^{(j)} \tag{8}$$

From this, it can be known that the unstructured texts in the resume and recruitment information are respectively processed by BiGRU for extracting the semantic features of words. Then, through the attention layer, the keywords are weighted to form the word vector features. Finally, the similarity matching between the resume and the recruitment information is calculated.

4. Experimental results and data analysis

4.1. Descriptive statistics of data

(1) Analysis of the Basic Information of Applicants This study analyzes the data of 674 applicants for college teaching positions collected by Jilin Animation Institute in 2024. In terms of educational background, those with a doctoral degree account for 5%, those with a master's degree account for 87%, and those with a bachelor's degree account for 8%. There is a trend that highly educated talents account for a relatively large proportion, which is consistent with the actual situation that universities have relatively high requirements for the educational background of teachers. In terms of the distribution of majors, applicants from science and engineering majors account for 31%, and those from liberal arts majors account for 69%. Among them, applicants from majors such as art and education are relatively concentrated, reflecting that universities have a greater demand for talents in these subject areas. Regarding work experience, applicants with more than 5 years of teaching or research experience account for 23%, those with 3 to 5 years of experience account for 47%, and those with less than 3 years of experience account for 30%. This indicates that the work experience of applicants is widely distributed, but there are relatively few experienced talents.

Through the analysis of the distribution characteristics of these data, we can initially understand the overall situation of the applicant group, providing a reference basis for feature selection and weight allocation in subsequent model training. For example, during model training, according to the distribution of educational backgrounds and majors, different weights can be assigned to applicants with different educational levels and professional backgrounds to improve the accuracy of the model in screening different types of talents.

(2) Analysis of Recruitment Position Requirements

By analyzing the information of the 23 collected recruitment positions, it is found that there are significant differences in the responsibilities and job requirements of different positions. Teaching and research positions require applicants to have a solid foundation of professional knowledge and high research capabilities, be able to undertake curriculum teaching tasks, and publish a certain number of research achievements in related fields. For example, the teaching and research position in the field of computer science requires applicants to have a doctoral degree, conduct in-depth research in areas such as artificial intelligence and big data, publish at least 3 SCI papers as the first author in the past 5 years, and possess rich teaching experience. On the other hand, administrative teaching assistant positions place more emphasis on the applicants' communication and coordination abilities, organizational and management capabilities, and service awareness. For instance, the administrative teaching assistant position in the Academic Affairs Office requires applicants to have a bachelor's degree or above, be proficient in office software, have more than 2 years of relevant work experience, and be able to handle teaching affairs and communication between teachers and students efficiently.

During the model training and screening process, we extract key features as model inputs according to the responsibilities and job requirements of different positions. For example, for teaching and research positions, the number of research achievements, the level of published papers, teaching experience, etc. are regarded as important features; for administrative teaching assistant positions, work experience, communication skills, proficiency in office software, etc. are taken as key features. By extracting and analyzing these features in a targeted manner, the accuracy of the model in screening applicants for different positions is improved.

4.2. Model construction

In this paper, the BiGRU-attention matching model was built through Keras. The 674 collected resume samples were divided into a training set and a test set at a ratio of 4:1. The dimension of the word vector was the same as that of the previous word2vec training. The batch training method was adopted. The Loss was continuously optimized through the categorical_crossentropy function, which was a classification cross-entropy function. During the iterative optimization process, the rmsprop optimizer was used to accelerate the entire training iteration process. In the process of training the model, this paper also used the Dropout function to randomly discard a small amount of data to prevent the phenomenon of model overfitting. Some parameter information during the training of the BiGRU-attention matching model was shown in the table 1.

Table 1: Parameter Information Table of the BiGRU-attention Matching Model

Training Parameters	**Parameter Value**
Dimension of Word Vector	250
Learning Rate	0.05
Number of Iterations	15
Batch Training Size	128
Loss Function	*categorical_crossentropy*
Optimizer	rmsprop
Metric	acc
Proportion of Validation Set	0.2
Regularization	0.2

4.3. Model performance analysis

In order to verify the effectiveness of the matching model in this paper, the following groups of models are added for comparative experiments:

(1) After training the word vectors of the text through the Word2vec model first, directly use the LSTM model for semantic training to directly obtain the LSTM-based matching model.

(2) First, train the word vectors in the text through the Word2vec model, then combine the semantic training of the LSTM model and use the attention mechanism to obtain the LSTM-Attention-based matching model.

(3) First, train the word vectors in the text through the Word2vec model, then combine the semantic training of the BiLSTM model and use the attention mechanism to increase the weights of the key words, and then train the BiLSTM-attention-based matching model.

All the above comparative experiments and the experimental scheme of this paper can be divided into a three-layer network structure, namely the input layer, the representation layer, and the matching layer. This paper adopts the commonly used indicators in text classification tasks, such as precision, recall, and the harmonic mean (F1 score), to measure the effectiveness of the model. Precision is the proportion of samples correctly identified as positive examples by the model among all the samples identified as positive examples. Recall is the proportion of samples correctly identified as positive examples by the model among all the actual positive examples. The F1 score is the harmonic mean of precision and recall, which is used to comprehensively consider precision and recall.

(1) F1 value

In this paper, the recommendations of TOP-4, TOP-8, TOP-12, TOP-16 and TOP-20 are selected to verify the F1 value. The relationships between the F1 value, the MAP value and TOP-N were shown in Figure 2 and Figure 3.

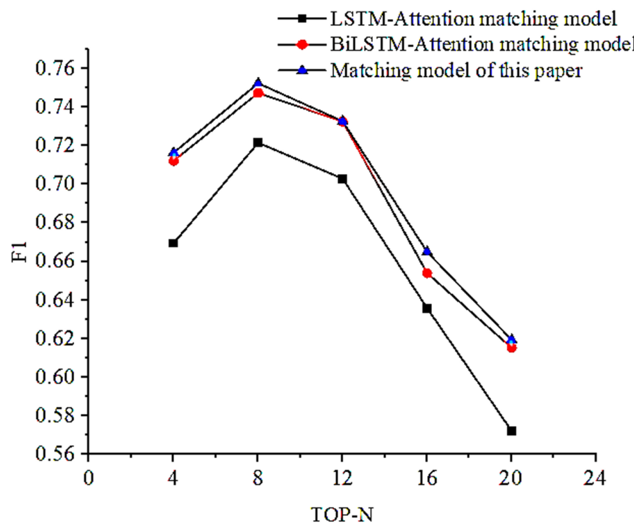

Figure 2: The Relationship between the F1 Value and TOP-N

As can be seen from Figure 2 and Figure 3, the F1 value of the scheme in this paper is significantly better than that of the LSTM-Attention matching model, and slightly better than that of the BiLSTM-attention matching model. Due to the relationship of the experimental data samples, it is obvious that the F1 value shows the best performance when TOP-8. Compared with the MAP value, the matching model of this text scheme is also significantly better than the LSTM-Attention matching model, and also slightly better than the Bi-LSTM-Attention matching model. Compared with the LSTM-Attention matching model, the average MAP of each item of TOP-N in the scheme of this paper is approximately 7.1% better. Compared with the BiLSTM-attention matching model, the average MAP value of each item of TOP-N is increased by approximately 0.4%. This indicates that when screening applicants for college teaching positions, this model can more accurately identify talents who meet the position requirements, and at the same time, it can more comprehensively cover potential suitable candidates, effectively improving the quality and efficiency of the screening.

(2) Precision-Recall curve

When verifying the performance of the algorithm, simply looking at the precision and recall values of the sampling points is not sufficient to fully illustrate the superiority or inferiority of the model's performance. Usually, it is necessary to plot the overall precision and recall values into a

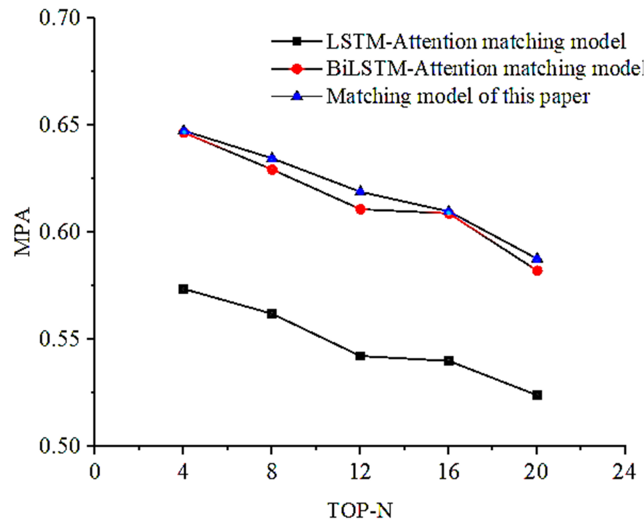

Figure 3: The Relationship between the MAP Value and TOP-N

P-R (Precision-Recall) curve, which can clearly compare the performance of various models. The results of each prediction are plotted into a P-R curve as shown in Figure 4.

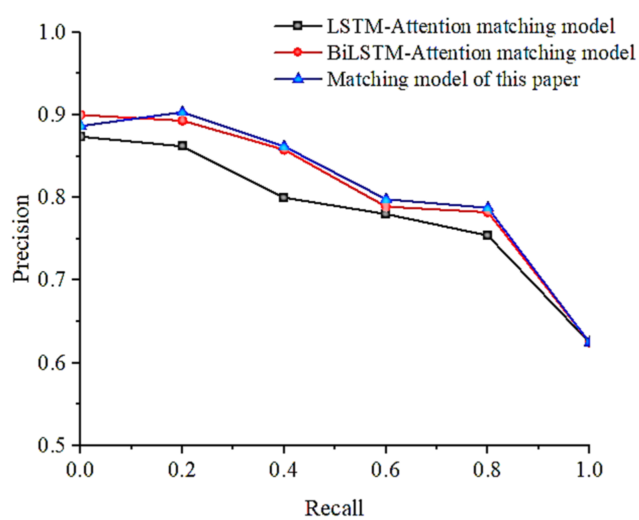

Figure 4: Precision-Recall curve

By plotting the Precision-Recall (P-R) curve based on the overall precision value and recall value, it can be seen that the model in this paper outperforms the LSTM-attention matching model and the BiLSTM-attention matching model. From this, it can be known that the matching scheme in this paper is superior to the model predictions of other matching schemes, indicating that the prediction results of the model are relatively close to the actual situation and have high reliability. Through the comparison with traditional screening methods on multiple evaluation indicators, the effectiveness and superiority of the intelligent screening model for college teacher recruitment based on the convolutional neural network are fully verified.

5. Conclusion

Aiming at the problems of low processing efficiency of a large amount of unstructured text data and strong subjectivity in screening during the recruitment of college teachers, this paper proposes an intelligent screening model based on BiGRU-attention. The combination of BiGRU and the attention mechanism is applied to the college recruitment scenario. Through the bidirectional gated recurrent unit, long-distance text dependencies are captured, and combined with the dynamic attention weight allocation, the semantic representation ability of unstructured texts is significantly improved. Experiments have verified that the BiGRU-attention model performs excellently in the unstructured text matching task. Its F1 value reaches a peak (86.7%) at the TOP-8 recommendation position, which is 7.1% higher than that of the LSTM-attention model and 0.4% higher than that of the BiLSTM-attention model. The analysis of the P-R curve shows that the model is superior to the comparative models in terms of the balance between precision and recall, proving that the synergistic effect of the bidirectional context capture and the attention weighting mechanism can effectively extract key semantic features.

References

Y. Cheng, J. Zhang, and Y. Liu. The impact of enterprise management elements on college students' entrepreneurial behavior by complex adaptive system theory. *Frontiers in Psychology*, 12:769481, 2022. doi: 10.3389/fpsyg.2021.769481.

T. Han, Z. Zhang, M. Ren, and et al. Text emotion recognition based on xlnet-bigru-att. *Electronics*, 12(12):2704, 2023. doi: 10.3390/electronics12122704.

Y.L. Jeng, C.F. Lai, S.B. Huang, and et al. To cultivate creativity and a maker mindset through an internet-of-things programming course. *Frontiers in Psychology*, 11, 2020. doi: 10.3389/fpsyg.2020.01572.

Zhonghui Kang. Artificial intelligence network embedding, entrepreneurial intention, and behavior analysis for college students' rural tourism entrepreneurship. *Frontiers in Psychology*, 13:843679, 2022. doi: 10.3389/fpsyg.2022.843679.

Y. Nie, H. Wang, B. Liu, and S. Yang. Research on construction of high-quality application-oriented talent cultivation system for internet of things engineering: Based on educational psychology. *Frontiers in Psychology*, 13:921840, 2022. doi: 10.3389/fpsyg.2022.921840.

Gaoyan Shi. Design and implementation of iot data-driven intelligent law classroom teaching system. *Computational Intelligence and Neuroscience*, 2022:8003909, 2022. doi: 10.1155/2022/8003909.

Q. Xu, S. Chen, Y. Xu, and et al. Detection and analysis of graduate students' academic emotions in the online academic forum based on text mining with a deep learning approach. *Frontiers in Psychology*, 14, 2023. doi: 10.3389/fpsyg.2023.1107080.

Y. Zhang, N. Xu, D. Feng, and et al. Design and implementation of sentiment analysis of weibo comments based on deep learning. In *2024 5th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT)*, pages 1579–1582, 2024. doi: 10.1109/AINIT61980.2024.10581844.

D. Zhao, H. Zhong, Y. Wu, and Q. Zhou. A study of the impact of internet-based instruction integrated innovation education on university student entrepreneurial team collaboration and strategic innovation. *Frontiers in Psychology*, 11:1264, 2020. doi: 10.3389/fpsyg.2020.01264.

C. Zhen and Y. Xiaoxuan. Adoption of human personality development theory combined with deep neural network in entrepreneurship education of college students. *Frontiers in Psychology*, 11: 1346, 2020. doi: 10.3389/fpsyg.2020.01346.