

Research Article

The Application of Deep Learning Model in Recruitment Decision

Yang Wang¹ and Zhengjie Zhu²

¹*School of Economics and Trade, Anhui Vocational College of Finance and Trade, Hefei 230000, China*

²*School of Finance and Public Administration, Harbin University of Commerce, Harbin 150000, China*

Correspondence should be addressed to Yang Wang; hbzwendy@afc.edu.cn

Received 10 January 2022; Revised 25 January 2022; Accepted 10 February 2022; Published 2 March 2022

Academic Editor: Xin Ning

Copyright © 2022 Yang Wang and Zhengjie Zhu. 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 rapid development of social economy, the competition of human resources is becoming more and more fierce. Recruitment, as the main way for enterprises to obtain talents, determines the future development of enterprises to a great extent. Compared with Western advanced countries, the research on recruitment in China started late, the overall research level is relatively backward, and most of the relevant technical means and analytical methods are introduced from advanced countries. The existing literature is still relatively scattered, which is not conducive to the rapid development of recruitment direction research, and is not conducive to specific applications. Starting with the existing deep learning, from four models, that is, based on the traditional machine learning model, conditional random field (CRF), deep learning models Bi-LSTM-CRF, BERT, and BERT-Bi-LSTM-CRF identify and automatically extract recruitment entities and study recruitment accordingly; BERT-Bi-LSTM-CRF-BERT and BERT-BiLSTM-CRF are the models with the worst recognition effect. Although they have stronger text feature extraction ability and context information capture ability, they are limited by the small scale of information science recruitment corpus and the small number of entities, so their performance under this task is not brought into full play. Although CRF is relatively traditional, it can still achieve excellent results on some small-scale sparse datasets.

1. Introduction

In recent years, deep artificial neural networks have won numerous competitions in the fields of pattern recognition and machine learning [1]. The historical survey succinctly summarizes most of the related work from the previous millennium [2].

Nowadays, we know that deep learning tools, as deep learning algorithms that have attracted much attention in applied machine learning, are also a series of computing methods, which allow algorithms to program themselves by learning from a large number of examples showing required behaviors without specifying rules explicitly. Among them, the key challenge of face recognition can be solved by deep learning and using face recognition [3] and verification signals as supervision to reduce personal internal differences and expand personal differences at the same time; among them, convolution network has also become the preferred method for analyzing medical images [4, 5]; by taking MNIST as an example to show the uncertainty of

the model, the uncertainty of the model in Bayesian pipeline and deep reinforcement learning are regarded as a practical task [6]; “deep learning” technology determines sequence specificity from experimental data, which provides an scalable, flexible, and unified calculation method for pattern discovery [7, 8]; the inception module in convolution neural network [9] is interpreted as an intermediate step between conventional convolution and deep separable convolution operation. From this point of view, deep separable convolution can be understood as Inception module with the largest number of towers. According to these studies, a new deep convolution neural network architecture can be proposed; it can also be used to detect diabetic retinopathy for reference [10], how to apply the nonlinear semisupervised embedding algorithm popular in “shallow” learning technology to deep multilayer architecture [11]; deep learning can be applied to create an algorithm for automatically detecting diabetic macular edema in retinal fundus photos.

Deep learning can also be applied to recruitment. We know that recruitment, as the main way for enterprises to

obtain talents, largely determines the future development of enterprises; school-based research tailors recruitment strategies for specific subgroups through multilevel recruitment process [12]; the quality of recruitment will also directly affect the quality of enterprise human resources [13]; the timeliness of recruitment is one of the key factors that will affect the success of each activity [14] [15]. It is necessary to recruit enterprises or personnel to adopt scientific, flexible, and effective recruitment methods to ensure that enterprises attract young and fresh blood and make enterprises rejuvenate. Reasonable use of recruitment management system, mainly including recruitment principles, processes, channels, recruitment effect evaluation, and effectiveness promotion, ensures the benign development of enterprises.

2. Literature Review and Related Theoretical Basis

2.1. Literature Review

2.1.1. Research Status Abroad. Scientific recruitment process can greatly improve recruitment efficiency and promote the realization of recruitment objectives [16]. Among them, the free employment system is widely implemented in German enterprises. Through two-way selection, enterprise recruiters can recruit or dismiss employees according to the actual needs of enterprises, and employees will participate in recruitment according to their professional skills and professional demands. There is also personalized recruitment, that is, enterprises need to recruit according to their own needs for employees' skills and qualities.

In order to recruit talents to meet the needs of enterprise management and development, the formulation of recruitment requirements, recruitment links, screening (preliminary screening, written test, interview, psychological test, and evaluation center method), employment, and evaluation are several major links in the effective staff recruitment process. On the basis of the employee recruitment process, the three recruitment processes of preparation, implementation, and evaluation are summarized into recruitment procedures, such as recruitment, selection, employment, and evaluation.

Effectiveness is mainly reflected in the following aspects: first of all, analyze the results of recruitment to see if it can effectively achieve the expected goals, then evaluate the recruitment cost, compare it with the benefits it can bring to the enterprise, and decide whether to recruit or not; secondly, it depends on whether the recruited people meet the needs of enterprises; finally, the methods, channels, and personnel arrangement used in the recruitment process are summarized, and their effectiveness is evaluated. Based on the above viewpoints, a relatively complete and detailed recruitment evaluation system is constructed. Effective recruitment means that enterprises can make plans, organize, and coordinate personnel to maximize the recruitment results to meet the needs of enterprises and most conducive to the future development of enterprises.

2.1.2. Research Status in China. Domestic experts and scholars have also carried out a lot of research on recruit-

ment and achieved certain results. Hu Junchen interpreted and analyzed recruitment in his research. He believed that recruitment is the process of selecting and appointing talents from inside and outside in order to achieve better and faster development; the whole process from publishing recruitment information to completing recruitment; Ethan believes that enterprises should raise awareness, carry out recruitment through multiple channels, carry out scientific strategic planning, and improve the effectiveness of recruitment; Jin Minglan believes that for the recruitment department, formulating a good set of procedures can indeed improve the efficiency of recruitment, but it is necessary to fully realize that there is applicability, that is to say, different recruitment objects should have different procedures, and effective procedures for recruiting ordinary employees may not be applicable to the recruitment of senior talents, so appropriate adjustments should be made according to the actual situation; for enterprise recruitment, we should start from the overall process optimization, pay attention to its performance management, build a scientific and complete recruitment process, break the extensive mode for a long time, and change to a refined and scientific direction. Therefore, by paying attention to a specific process of recruitment management, the overall effectiveness can be improved.

2.1.3. Literature Review. To sum up, through the comprehensive research of domestic and foreign scholars on employee recruitment, such as the design of employee recruitment process and the choice of recruitment channels, as well as the premise of recruitment design, the evaluation of recruitment results, and the improvement of effectiveness, however, the existing literature is still relatively scattered, which is not conducive to the rapid development of recruitment direction research and is not conducive to specific application.

2.2. Theoretical Basis

2.2.1. Competency Theory. David McClelland, a famous psychologist, defines competency as an individual characteristic that can distinguish individual performance level in a specific job and organizational environment [16]. Competency is a targeted and dynamic new concept of ability, which has very strong post and professional characteristics. In different organizations and industries, similar jobs have different requirements for competency; in an organization, different jobs often have different requirements for competency.

Through the competency iceberg model proposed by McLelland, this paper demonstrates the relationship between individual and post job performance from two levels of quality and ability, and can predict future job performance from this. As shown in Figure 1, surface competence floats on the water, knowledge, or skills; deep competency is hidden underwater, such as self-concept, self-awareness, characteristics, and motivation. Whether knowledge and skills can have outstanding performance fundamentally depends on whether the underwater part is firm or not. Therefore, the underwater part is the most critical.

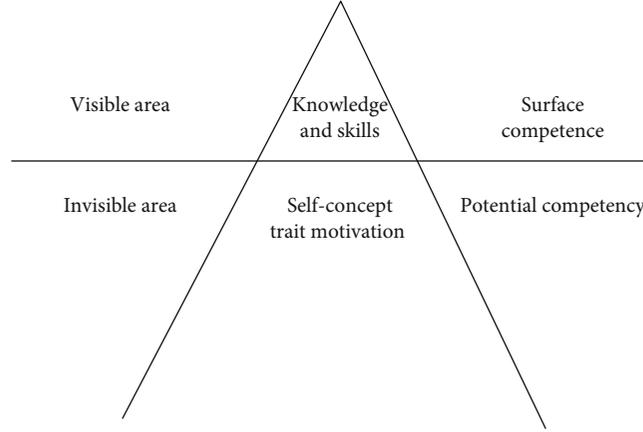


FIGURE 1: Iceberg model.

Richard Boyz's onion model has also made an important contribution to the development of competency model theory. The onion model is taken as an example through the section diagram of onion. As shown in the following Figure 2, the outer knowledge is the easiest part to cultivate; the inner layer is motivation and personality, which generally does not change, but is the core of the whole competency.

2.2.2. Person-Job Matching Theory. Recruitment should follow the matching principle, which means that in the recruitment process, candidates should be considered to match with occupations, teams, organizations, and jobs, and finally find suitable employees. Person-job matching means that the recruited employees should be able to meet the requirements of the post, so that the recruited employees can match the post requirements.

Holland studied the relationship between employees' personality and work and thought that employees' personality and job-seeking tendency are closely related. Under the same ability conditions, employees with different personalities may have completely different behaviors for the same job. Therefore, it is necessary to further develop and tap the potential of employees by matching their personalities and positions, so that they can highlight their attitudes and values and then gain professional identity. Enterprises can set the personality of the required employees according to the post types, and then carry out targeted recruitment work.

Williams' research focuses on the matching of employees' personality, ability, and interest with the specific requirements of jobs, which also proves the importance of the theory of person-job matching to a great extent.

3. Getting Started with Deep Learning

3.1. Overview of Convolution Neural Networks (CNN). Deep learning convolution neural network (CNN), which consists of convolution layer and sampling layer and alternately forms network topology, is a deep learning algorithm that can realize fast response to data. The feature extraction process can be expressed as a fractional function $S(x_i, w, b)$. The cross loss function of classification error of i samples (x_i, y_i) is defined as [17, 18]

$$L_i = -\ln e^{S_{y_i}} + \ln \sum e^{S_i}. \quad (1)$$

In the expression, S_{y_i} represents the true classification score of the i th sample of the training set, and S_i represents the ratio of the index of the current element to the sum of the indexes of all elements. The output of sample (x_i, y_i) after passing through the network is $f(x)$, and the corresponding sample loss value is

$$L_i(f(x), y) = -\ln f(x)_y. \quad (2)$$

The error sensitive term l of CNN output layer of deep learning layer is

$$\delta_l = \frac{\partial L}{\partial a^l} = \nabla_a^l(x) - \ln f(x)_y = f(x) - y, \quad (3)$$

where a^l represents the input of layer l . Finally, the weight of each neuron is updated by using the back propagation rule of CNN, so that the overall error function of the model will be continuously reduced. The convolution process is defined as follows:

$$x_j^l = f \left(\sum_{i \in M_j} x_i^{l-1} \times k_{ij}^l + b_j^l \right), \quad (4)$$

where l is the convolution layer number of the model, k_{ij}^l is the convolution kernel number, b_j^l is the additive bias, f is the activation function, and M_j is the input image. Because of the alternate setting of convolution layer and sampling layer of CNN, the learning ability of network will be greatly affected by input characteristics.

3.2. Overview of Cyclic Neural Network

3.2.1. Cyclic Neural Network. Cyclic neural network is a kind of neural network model for processing sequence data [19]. Compared with the traditional neural network, which limits the input to a fixed size, RNN network model can accept any length sequence as input, which makes RNN

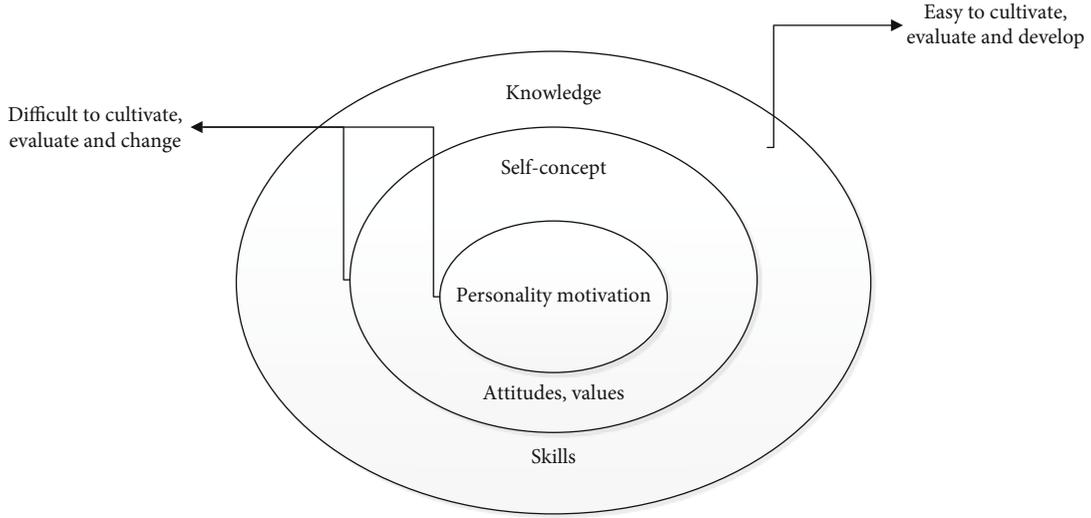


FIGURE 2: Onion model.

model analyze emotion based on context. Nowadays, RNN network model has been successfully and widely used in sequence class modeling.

RNN network model is based on the idea of unfolding computational graph, which makes the parameters in deep neural network shared. In RNN network model, the hidden state of the model at a certain time and the output of the model at the previous time jointly determine the input of the model at the next time, and it can also be regarded as the result of the joint action of the input of the model at that time and the historical information accumulated by the model, so it is regarded as a neural network with memory function.

As shown in the figure, it is the structure diagram of the standard RNN network model. The standard RNN takes the following form: given a sequence (x_1, \dots, x_T) of input vectors, RNN calculates a series of hidden states (h_1, \dots, h_T) and a series of outputs (y_1, \dots, y_T) and iterates to $t = 1$ by following the equation when $t = T$:

$$\begin{aligned} h_t &= \tan h(W_{hx}x_t + W_{hh}h_{t-1} + b_h), \\ y_t &= W_{yh}h_t + b_y, \end{aligned} \quad (5)$$

where W_{hh} is a hidden-to-hidden weight matrix, W_{hx} is an input-to-hidden weight matrix, W_{yh} is a hidden-to-output weight matrix, and b_h and b_y are deviations. When $t = 1$, $W_{hh}h_{t-1}$ is replaced by a special initial deviation vector h_{init} , and the nonlinearity of Tanh function is applied in coordinate mode, as shown in Figures 3 and 4.

3.2.2. Multiplicative Cyclic Neural Network. In the process of character-level language modeling, the standard RNN network can be regarded as a borderless tree model, in which each node is a hidden state vector, and each edge is marked by a character, which determines how the parent node produces children. If the hidden state vectors representing the two history records “type” and “walk” share a generic representation indicating that this may be a verb stem, it can be

performed with the character “i.” Generate a hidden state of emotional analysis “n.” In this paper, the stemming representation of a previously hidden verb is combined with the character “i” to make it more emotionally analytic. Make the weight matrix from its hidden state to hidden state a function of the learning of the current input x_t .

Make the weight matrix from its hidden state to hidden state a function of the current input learning:

$$\begin{aligned} h_t &= \tan h\left(W_{hx}x_t + W_{hh}^{(x_t)}h_{t-1} + b_h\right), \\ y_t &= W_{yh}h_t + b_y. \end{aligned} \quad (6)$$

Compared with the standard RNN network, $W *$ in the expression replaces $W a$ in the expression. $W_{hh}^{(x_t)}$ allows each character to specify a different hidden weight matrix. RNN of this architecture uses tensor to define $W_{hh}^{(x_t)}$, called tensor RNN (tensor RNN), which allows each possible input to use a different transition matrix.

When input x_t is the encoding of a character 1-of- N , each character has an associated individually learned weight matrix, and the conversion matrix x_t assigned to $W_{hh}^{(x_t)}$ representing the character is defined as follows:

$$W_{hh}^{(x_t)} = \sum_{n=1}^N W_{hh}^{(n)} x_t^{(n)}. \quad (7)$$

The superscript is used to represent the index in the array, $x_t^{(n)}$ is the n th coordinate of x_t , N is the dimension of x_t , and $W_{hh}^{(x_t)}$ stores N matrices $(W_{hh}^{(1)}, \dots, W_{hh}^{(N)})$.

A large number of parameters in tensor RNN require huge storage space, which makes it difficult to deal with most problems. Multiplicative cyclic neural network is a network structure specially designed to allow flexible transition depending on input, and this problem can be solved by decomposing tensor $W_{hh}^{(x_t)}$. MRNN introduces three

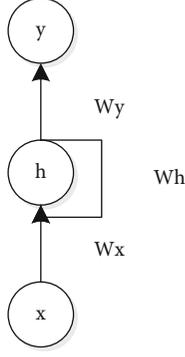


FIGURE 3: Standard RNN network model framework diagram.

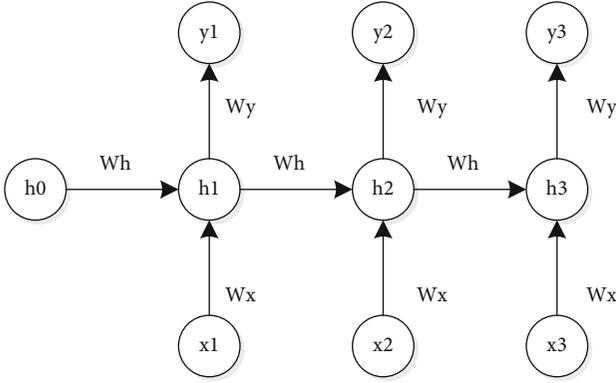


FIGURE 4: Standard RNN network model expansion calculation diagram.

matrices W_{hm} , W_{mh} , and W_{mx} and will be reparameterized $W_{hh}^{(x_t)}$ as follows:

$$W_{hh}^{(x_t)} = W_{hm} \times \text{diag}(W_{mx}x_t) \times W_{mh}. \quad (8)$$

MRNN can be regarded as the shared parameter approximation of tensor RNN, which uses the factorized hidden-to-hidden transformation matrix instead of the normal RNN hidden-to-hidden matrix $W_{hh}^{(x_t)}$, and has an input-dependent intermediate diagonal matrix $\text{diag}(W_{mx}x_t)$. When the dimension of $W_{mx}x_t$ is large, factorization has the same expressive ability as the original tensor; when $W_{mx}x_t$ dimension is small, fewer parameters are needed, while retaining most of the expressive ability of tensor.

MRNN can be regarded as an extension of tensor RNN, that is, an extension of expression.

The MRNN calculates the hidden state sequence (h_1, \dots, h_T) , the factor state sequence (m_1, \dots, m_T) , and the output sequence (y_1, \dots, y_T) through the following iterative equations.

$$m_t = \text{diag}(W_{mx}x_t) \times W_{mh}h_{t-1}, \quad (9)$$

$$h_t = \tan h(W_{hm}m_t + W_{hx}x_t),$$

$$y_t = W_{yh}h_t + b_y. \quad (10)$$

As shown in Figure 5 and Figure 6, the input of each time step of MRNN is obtained by nonlinear processing of two hidden states, the factor state m_t multiplication unit and the hidden state of the previous step, and the multiplication unit and input. The output of the product of these two linear filters is output. In the MRNN structure, each input can integrate its own hidden to hidden weight matrix by measuring the benefits of all factors. Each hidden-to-hidden weight matrix is represented by a rank-one matrix defined by the external product of input-to-hide and hide-to-output.

In the character-level language modeling project, MRNN is improved on the standard RNN. Because of the “disappearance/explosion gradient problem” in model training, it may be difficult to retain long-term information.

3.2.3. Low Molecular Weight Heparin Circulating Neural Network. Long-short-term memory (LSTM) cyclic neural network is a special RNN structure [20]. LSTM controls how information flows into and out of the network through a series of logic gates, solves the “gradient dispersion/explosion problem” in RNN network training, and effectively learns the long-term dependencies in input sequences.

In addition to the external RNN loop in the LSTM network structure diagram, there is also a linear self-loop inside, as shown in the following figure.

In addition to input and output, each unit of LSTM has additional parameters and gating units to control information transfer and forgetting. In order to better deal with historical information and control the flow of information, the design of LSTM network adds state unit c , input gate i , forgetting gate f , and output gate y . The weight of the forgetting gate f_t control state unit linear self-cycle is calculated from the output at time $t-1$ and the input at time t as follows:

$$f_t = \sigma(W_{fx}x_t + W_{fh}h_{t-1} + b_f), \quad (11)$$

where x_t is the input of the network, h_{t-1} is the hidden layer vector at the previous moment, W_{fx} is the forgetting-input weight matrix, W_{fh} is the forgetting-hiding weight matrix, and b_f is the offset. σ is the Sigmoid function, and the expression is as follows:

$$\sigma(x) = \frac{1}{(1 + e^{-x})}. \quad (12)$$

At time t , the input gate i_t , the status unit temporary value \tilde{c}_t , and the status unit update value c_t are calculated as follows:

$$i_t = \sigma(W_{ix}x_t + W_{ih}h_{t-1} + b_i),$$

$$\tilde{c}_t = \tan h(W_{cx}x_t + W_{ch}h_{t-1} + b_c), \quad (13)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t, \quad (14)$$

where W_{ix} is the input weight matrix, W_{ih} is the input-hidden weight matrix, W_{cx} is the state-input weight

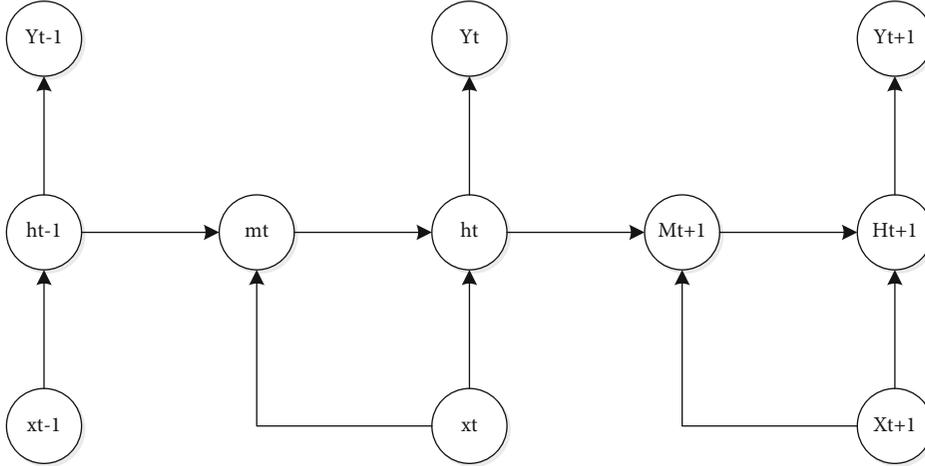


FIGURE 5: Schematic diagram of multiplication recurrent neural network.

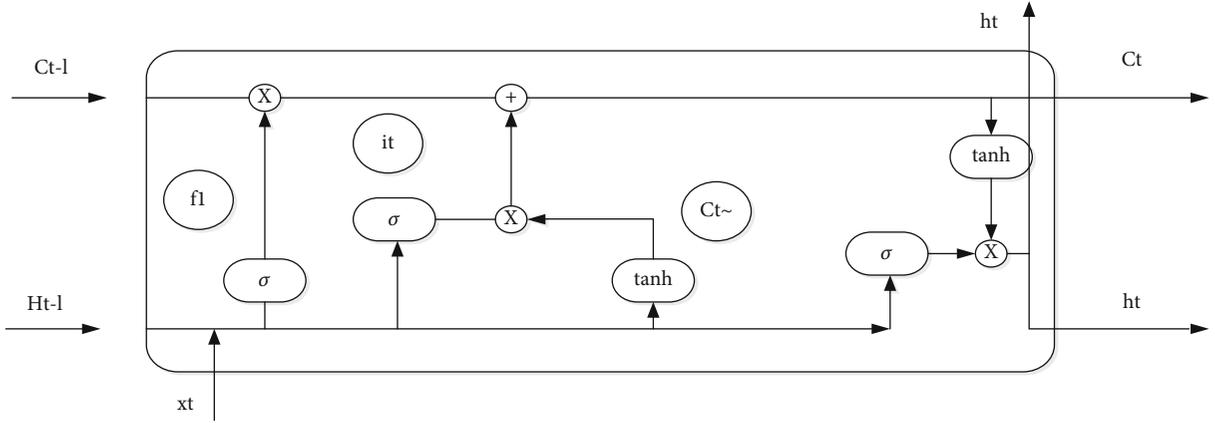


FIGURE 6: LSTM network structure diagram.

matrix, W_{ch} is the state-hidden weight matrix, b_i and b_c are the offset, and Tanh is the hyperbolic tangent function. The input gates i and f_t forgetting gates are both acted upon by the Sigmoid function, which sets the weights to between 0 and 1. When $f_t = 1$, the state unit c transmits all the information at time $t-1$ to time t ; when $f_t = 0$, the state unit c_t forgets all the information at time $t-1$. When $i_t = 1$, the state unit c_t will accept all the information of the temporary value of the state unit $c_t^~$ at time t ; when $i_t = 0$, the state unit c_t forgets all the information of the temporary value of the state unit $c_t^~$ at time t . At time t , the calculation formula of output gate y and hidden state h_t is as follows:

$$\begin{aligned} y_t &= \sigma(W_{yx}x_t + W_{yh}h_{t-1} + b_y), \\ h_t &= \tan h(c_t) \cdot y_t. \end{aligned} \quad (15)$$

Both LSTM and MRNN network architectures have multiplication units. The gating unit of LSTM is designed to control the information flow through the network; while MRNN's gating unit is designed to allow the change of conversion function between inputs. The LSTM gating unit receives input from the input unit and the hidden

unit, thereby allowing multiplicative interaction between the hidden units and limiting the degree of input-hidden multiplicative interaction to a certain extent. LSTM gating through Sigmoid function makes them compressed to between 0 and 1, which makes LSTM easier to control, but its performance is not as good as linear gating of MRNN. The linear gate of MRNN is directly learned from each input and can be placed between the products of two dense matrices, which provides greater flexibility for the final weight matrix product.

4. Design and Establishment of Model

4.1. Brief Introduction of Model. Because there are some problems in the research, such as insufficient data to support analysis, untimely update of content, backward analysis, and mining methods, it is difficult to objectively and fairly reflect the current information science-related curriculum setting and the talent demand of jobs as a whole, so named entity recognition will be used for analysis.

Based on the traditional machine learning model conditional random field (CRF) [21] and the deep learning models Bi-LSTM-CRF [22], BERT, and BERT-Bi-LSTM-CRF, the recruitment entities [23] are identified and automatically

TABLE 1: Signs and meanings of recruitment entities.

Serial number	Entity tag	Meaning of entity	Quantity
1	xl	Requirements for experience	698
2	nl	Requirements for ability	3691
3	zy	Requirements for character	2008
4	xg	Requirements for academic qualifications	528
5	jy	Professional requirements	299

extracted by computer. By using the method of computer automatic information extraction, the information of information science-related posts such as “requirements for academic qualifications,” “requirements for majors,” “requirements for experience,” “requirements for abilities,” and “requirements for personality” are extracted from massive job advertisements or announcements, and the common needs of each post are excavated from these information. The recognition and extraction of recruitment entities essentially belongs to the category of Chinese named entity recognition. Named entity recognition refers to the process of recognizing specific contents such as person name, place name, organization name, and time from text, which is the basis of natural language processing research in deep-level applications such as chunk annotation, automatic syntactic analysis, information retrieval, and machine translation. This paper also marks the relevant entities of the collected data. The following Table 1 is the definition of the marks and meanings of the recruitment entities [24, 25].

Among them, BERT model and BERT-Bi-ISTM-CRF model use open-source model as initialization word vector, and Bi-LSTM-CRF model uses word2vec tool to embed each character into a 128-dimensional word vector.

4.2. Introduction of Performance Evaluation Index of Recruitment Entity Identification. In this paper, we use three evaluation indexes: precision, recall, and F -score to measure the related entity recognition performance of the four models in this paper. The calculation methods of each index are as follows, where denotes the number of correctly identified recruiting entities, the number of incorrectly identified recruiting entities, and the number of unidentified recruiting entities, respectively.

$$\text{Precision rate : } P = \frac{A}{A+B} \times 100\%, \quad (16)$$

$$\text{Recall rate : } R = \frac{A}{A+C} \times 100\%,$$

$$\text{F value : } F = \frac{2 \times P \times R}{P+R} \times 100\%. \quad (17)$$

4.3. Model Test. In order to make the distribution probability of samples consistent in the dataset, this paper will use tenfold cross-validation to carry out experiments. Through the previous definition of educational background, ability, speciality, character, and experience, B is the beginning word

TABLE 2: Entity recognition effect.

Model	Precision rate (%)	Recall rate (%)	F value (%)
CRF	35.24	34.12	33.21
Bi-LSTM-CRF	36.78	42.12	37.28
BERT	26.76	39.16	32.98
BERT- Bi-LSTM-CRF	29.32	42.87	34.72

TABLE 3: Recruitment entity extraction and evaluation based on CRF model.

Serial number	Precision rate (%)	Recall rate (%)	F value (%)
1	90.22	74.67	82.12
2	86.23	72.18	83.12
3	88.12	78.23	82.15
4	90.26	76.29	82.81
5	88.42	76.11	80.29
6	87.28	78.27	80.56
7	88.27	74.09	78.89
8	88.28	71.72	78.29
9	89.27	71.67	78.28
10	84.80	78.19	81.28
Average	88.12	75.14	80.78

of entity, I is the middle word of entity, E is the end word of entity, and 0 is all the middle words of nonentity.

The following are specific tag sets:

$$\{B-xl, I-xl, E-xl, B-nl, I-nl, E-nl, B-xg, I-xg, E-xg, B-zy, I-zy, E-zy, B-jy, I-jy, E-jy, 0\} \quad (18)$$

Table 2 below shows the results of the four models.

From the table analysis, the accuracy rate, recall rate, and F value are all small and do not exceed, so we will analyze the reasons why F value is too small for F value as follows:

- (1) The amount of data used to support model training is insufficient. The size and quality of datasets largely determine the performance of deep learning models
- (2) There are too many types of recruiting entities. It is easy to interfere with the experimental results

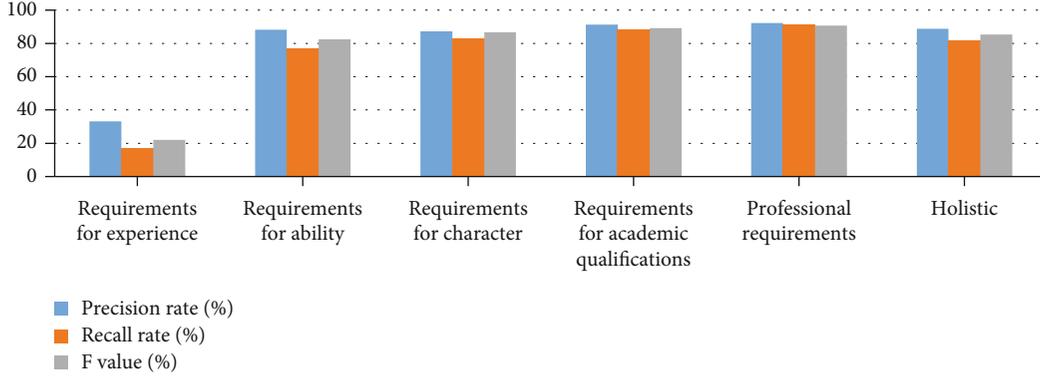


FIGURE 7: The best-performing CRF model entity recognition effect evaluation of each category.

- (3) The recruitment entity with long length occupies too much. It interferes with the experimental process to some extent

To sum up, under the premise of not reducing the entity category, we have improved the recruitment entity identification experiment appropriately and conducted the experiment under the condition that only the recruitment entities whose entity length is no more than 6 are considered. By deleting entities with too long length and too low frequency, the common features of entities are more obvious, and the recognition ability of machine learning and deep learning models for entities is enhanced.

5. Experimental Results and Analysis

Based on the above data collection and collation, four models of CRF, BERT, BERT-Bi-ISTM-CRF, and Bi-LSTM-CRF were tested and analyzed.

5.1. Results and Analysis of CRF Experiment. The specific performance of tenfold cross-validation by CRF model is shown in Table 3.

It can be seen from Table 3 that in the tenfold cross-validation, the experimental evaluation results of each group are 50% or above. Among them, the second group is the best, with the best F value of 83.12%, the accuracy rate of 86.23%, and the recall rate of 72.18%. The F value difference between the best group 2 model and the worst group 9 model is 4.84%, which shows that the performance of CRF is easily affected by the distribution of datasets.

From the mean value, the average value of F is 80.78%. Compared with CRF accuracy, recall, and F value in table model test, the data are all improved, so the entity recognition performance of the model has been improved to a certain extent. On the one hand, the results show that the improvement of recruitment entity identification experiment in this paper is effective; on the other hand, it also shows that CRF basically meets the requirements of this study.

Figure 7 below lists the detailed evaluation results of the second group in the tenfold cross experiment, that is, the group with the highest F value, and shows us in detail the

TABLE 4: Recruitment entity extraction evaluation based on Bi-LSTM-CRF model.

Serial number	Precision rate (%)	Recall rate (%)	F value (%)
1	81.78	74.29	76.58
2	78.38	76.19	79.81
3	82.30	73.28	77.44
4	29.32	74.28	77.72
5	81.90	74.32	78.36
6	81.58	73.38	77.09
7	84.86	70.68	77.08
8	82.69	78.09	79.80
9	76.48	70.73	73.48
10	77.58	78.85	74.21
Average	80.93	74.25	77.39

recognition performance of five recruitment entities: “requirements for academic qualifications,” “requirements for majors,” “requirements for experience,” “requirements for abilities,” and “requirements for personality.”

CRF model has a good recognition effect on four entities: “requirements for academic qualifications,” “requirements for majors,” “requirements for abilities,” and “requirements for personalities.” However, the recognition effect of “professional requirements” entity is the best, with the accuracy rate reaching 92.17%, and the recall rate and F value exceeding 90%, which shows that CRF has reached an applicable level in recognizing professional requirements of related posts. Among them, the recognition effect of “requiring experience” entity is the worst, and the F value is only 22.12%, which shows that the characteristics of “requiring experience” entity are relatively insignificant. Through the statistics of the distribution of various entities in the training dataset, it is found that this category of entities only accounts for 4.91% of the total number of entities, while the number of entities with the highest proportion of “ability requirements” is 54.16% of the total number of entities, accounting for half. This also shows that the scarce training corpus is also one of the reasons for the poor recognition effect. In this group of experiments, the F value of three entities, namely, “requirement for ability,” “requirement for character,” and

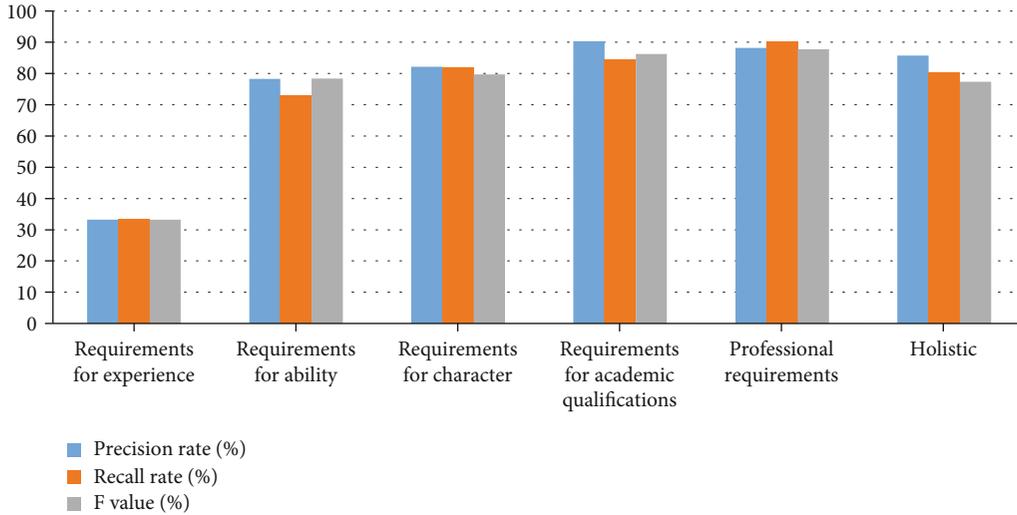


FIGURE 8: The best-performing Bi-ISTM-CRF model entity recognition effect evaluation of each category.

“requirement for educational background,” also exceeded 80%, which basically met the task requirements.

5.2. Experimental Results and Analysis of Bi-LSTM-CRF. The following are the specific values of the tenfold cross-validation evaluation indicators of Bi-LSTM-CRF model, as shown in Table 4.

As can be seen from Table 4, the overall effect of Bi-LSTM-CRF is lower than that of CRF. The average accuracy rate of 10 groups of experiments is 80.93%, the average recall rate is 74.25%, and the average F value is 77.39%. The average F value of CRF is 3.39% lower than that of CRF model. Specifically, the highest accuracy rate, recall rate, and F value were 82.30% in the third group, 78.85% in the tenth group, and 79.81% in the second group, respectively. Similar to CRF model, Bi-LSTM-CRF model has the highest overall score for the recognition effect of recruitment entities. Figure 8 also lists the evaluation scores of the recognition performance of five recruitment entities in this group of experiments: “requirements for academic qualifications,” “requirements for majors,” “requirements for experience,” “requirements for abilities,” and “requirements for personality.”

In Bi-LSTM-CRF model, “professional requirements” entity and “experience requirements” entity are also two categories with the highest F score and the lowest F score, and the ranking of the scores of the five entities is the same as that of CRF. The recognition effects of the four entities, namely, “requirements for education,” “requirements for specialty,” “requirements for ability,” and “requirements for personality,” are all better, while the recognition effects of the entity, “requirements for experience,” are the worst. This shows that the distribution of entities in the corpus does have a certain impact on the final recognition results of the model. Compared with CRF model, the F value of Bi-LSTM-CRF model is 33.23%, which is 11.11% higher than that of CRF model. This result shows that the neural network model based on sequence-to-sequence architecture has stronger feature learning ability on entities with lower occurrence frequency.

TABLE 5: Recruitment entity extraction evaluation based on the BERT model.

Serial number	Precision rate (%)	Recall rate (%)	F value (%)
1	77.08	77.56	77.33
2	72.78	82.42	80.18
3	81.54	78.16	79.87
4	80.05	77.47	78.78
5	75.17	79.08	77.05
6	68.93	79.98	74.01
7	75.94	77.09	76.00
8	77.52	83.02	76.91
9	76.80	76.81	76.81
10	69.72	82.87	75.52
Average	75.49	79.32	77.28

5.3. Results and Analysis of BERT Experiment. In the recruitment entity identification experiment based on BERT model, the evaluation index results obtained by 10% cross-validation are shown in Table 5.

We know that BERT, as a new model in the field of natural language processing, has achieved the best results in 11 different NLP tests.

In this experiment, the results of BERT models in 10 groups are quite different, the highest F value in the second group is 80.18%, and the lowest F value in the sixth group is 74.01%, which shows that the distribution of corpus has a great influence on the entity recognition results of BERT models. The average accuracy, recall, and F values of these 10 groups were 75.49%, 79.32%, and 77.28%, respectively. Although the average F value of BERT model is slightly lower than that of Bi-LSTM-CRF model by 0.11%, the F value of the second group of experiments with the best effect is higher than that of the latter by 0.37%, and the score of F value of the group of experiments with the worst effect of BERT model is higher than that of Bi-LSTM-CRF model,

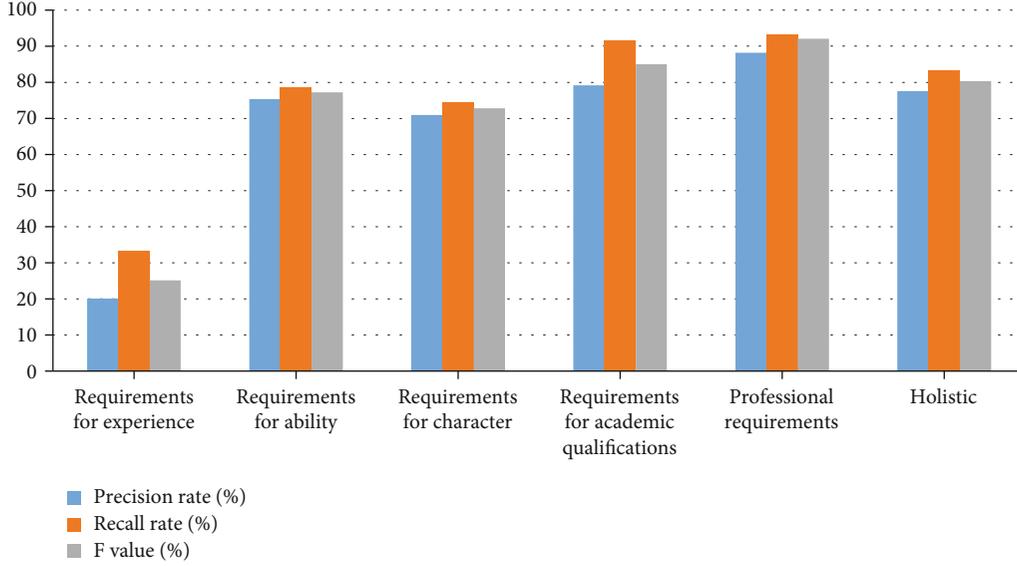


FIGURE 9: The best-performing BERT model entity recognition effect evaluation of each category.

which shows that the two models have basically the same ability to identify recruitment entities.

Figure 9 also lists the experimental evaluation results of the recognition performance of five recruitment entities in this group of experiments in BERT, namely, “requirements for education,” “requirements for major,” “requirements for experience,” “requirements for ability,” and “requirements for personality.”

The recognition F value of BERT model for “professional requirements” entities is as high as 92.11%, which is better than that of CRF model for “professional requirements” entities (92.17%) and Bi-LSTM-CRF for “professional requirements” entities (87.78%) under the same conditions, and its recall rate is as high as 93.28%. Similarly, the BERT model is also lacking in recognition effect for entities with low frequency of “experience requirements.” In addition, unlike CRF model and Bi-LSTM-CRF model, the performance of BERT model is better than that of Bi-LSTM-CRF model in the category of “entity for personality,” while BERT model is better in the category of “entity for ability.”

5.4. Experimental Results and Analysis of BERT-Bi-LSTM-CRF. BERT-Bi-LSTM-CRF model maps recruitment corpus to word vectors through BERT deep text representation model and then extracts features and learns context information through Bi-LSTM-CRF neural network. The final tenfold cross-validation results are shown in Table 6.

Among the 10 groups of evaluation data in Table 6, the second group has the best evaluation result, in which the best F value is 81.03%, which is 2.09% lower than the best group in BERT. From the average point of view, the accuracy rate, recall rate, and F value are 74.43%, 80.09%, and 76.69%, respectively. In BERT-Bi-LSTM-CRF model, the difference between group 2 with the highest F value of 81.03% and group 9 with the lowest F value of 75.41% is 4.84%, which once again shows that corpus quality has great influence

TABLE 6: Recruitment entity extraction evaluation based on the BERT-Bi-LSTM-CRF model.

Serial number	Precision rate (%)	Recall rate (%)	F value (%)
1	76.39	77.75	77.03
2	71.99	80.79	81.03
3	80.59	77.28	78.89
4	78.02	78.69	78.32
5	74.08	79.88	76.81
6	67.88	79.66	73.29
7	73.24	75.31	74.28
8	77.37	82.87	76.45
9	73.90	76.97	75.41
10	70.64	81.93	75.89
Average	74.43	80.09	76.69

on BERT performance. Figure 10 shows the experimental evaluation results of the recognition performance of five recruitment entities in the experiment of the model with the highest F value in the second group, namely, “requirements for academic qualifications,” “requirements for majors,” “requirements for experience,” “requirements for abilities,” and “requirements for personality.”

BERT-Bi-LSTM-CRF model also performs poorly in recognizing entities with “requirements for experience,” which shows that both traditional machine learning model and deep learning model cannot solve the problem of low-frequency entity feature learning well, and often perform poorly in recognizing low-frequency categories when facing training sets with uneven category distribution. Whether it is CRF model, Bi-LSTM-CRF model, or even BERT model and BERT-Bi-LSTM-CRF model, they all achieve the best performance on the second group of corpus in the tenfold cross experiment of recruitment entity recognition, which shows that the distribution of training set and test set has

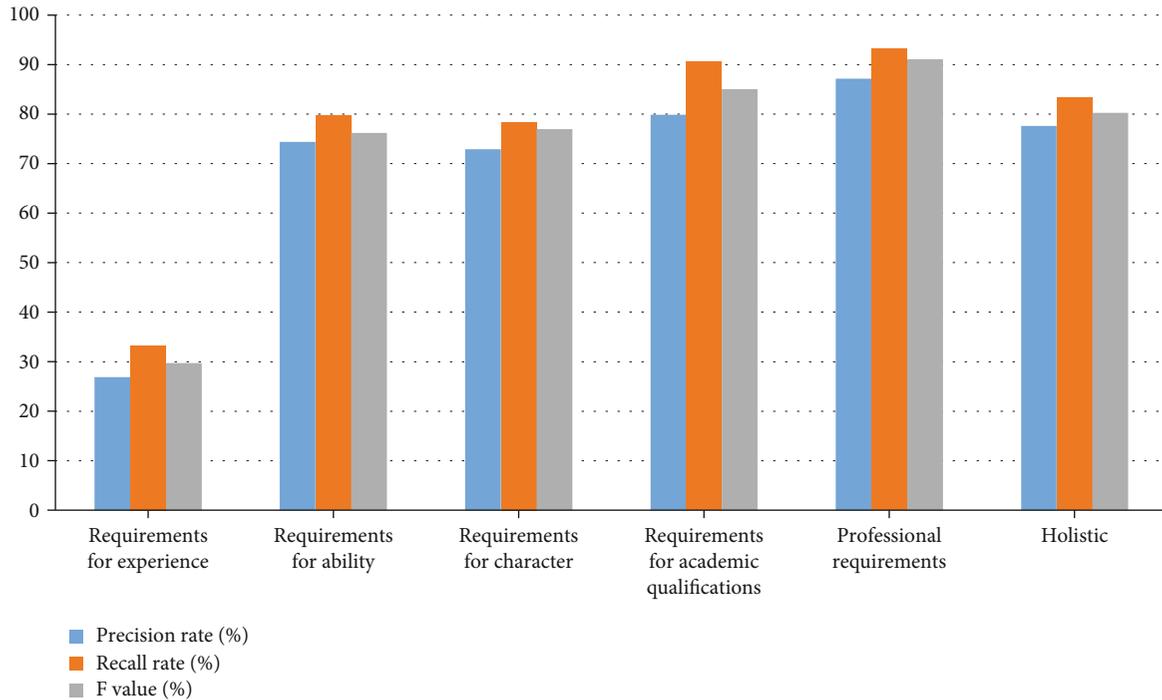


FIGURE 10: The best-performing BERT-Bi-LSTM-CRF model entity recognition effect evaluation of each category.

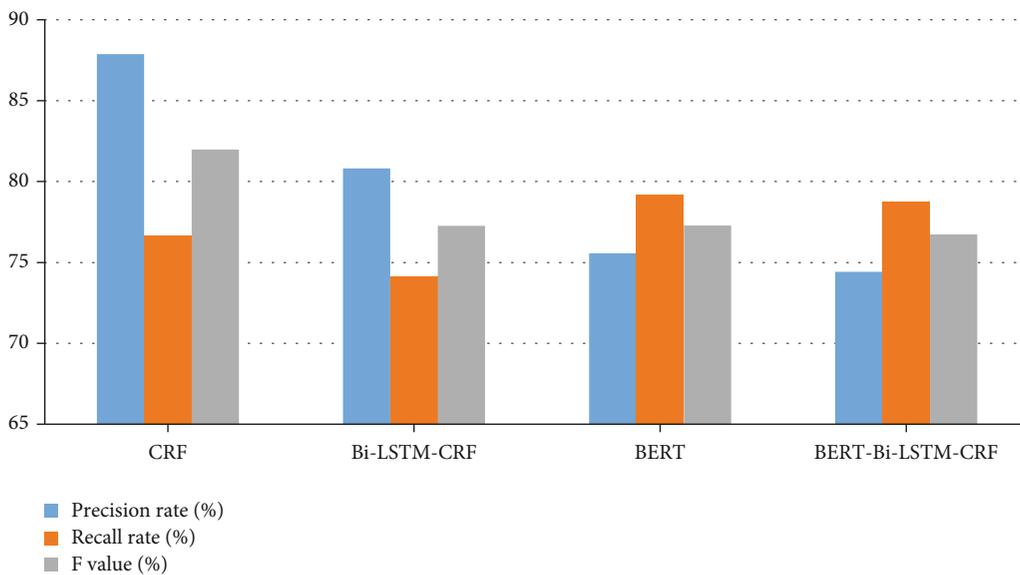


FIGURE 11: Comparison of evaluation values of all machine learning models.

an impact on the feature extraction ability and generalization ability of machine learning and deep learning models. It is also proved that the method of cross-validation can eliminate the random error caused by the uneven distribution of data and ensure the accuracy and authenticity of the experimental evaluation results.

5.5. Overall Performance Comparison. Figure 11 shows the numerical comparison of accuracy rate, recall rate, and F

value of 10-fold cross-validation of CRF, BERT, BERT-Bi-LSTM-CRF, and Bi-LSTM-CRF4 machine learning models and deep learning models. It can be found that except CRF model, the F value of other models is lower than 80%. Bi-LSTM-CRF model and BERT model ranked second and third, respectively. The models with the worst recognition effect are BERT-Bi-LSTM-CRF-BERT model and BERT-BiLSTM-CRF model. Although they have stronger ability to extract text features and capture context information, they

are limited by the small scale of recruitment corpus and the small number of entities, so their performance under this task is not maximized. Although CRF model is relatively traditional, it can still achieve excellent results on some small-scale sparse datasets.

6. Conclusion

Facing the talent training and recruitment in the new era, the recruitment entity knowledge base is established by using deep learning technology. The mining and analysis of recruitment entities not only present the real needs of recruitment for each post but also provide a reference for the improvement of posttraining system, which can help relevant workers to improve their ability to cope with work. Based on four machine learning and deep learning models, CRF, Bi-LSTM-CRF, BERT, and BERT-Bi-ISTM-CRF, this paper realizes the automatic identification and extraction of recruitment entities; it can be found that CRF can achieve excellent results on some small-scale sparse datasets. It is hoped that in the future research, we can continue to expand the scale of recruitment corpus to meet the needs of large-scale deep learning model training and choose more advanced deep learning models to mine the needs of various recruitment positions. It is also hoped that in the future, a learning model dedicated to recruitment needs can be created, which is more helpful for recruitment decision-making.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declared that they have no conflicts of interest regarding this work.

Acknowledgments

This work was supported by the Anhui Vocational College of Finance and Trade to improve the quality of the establishment project-based on the perceived value of the new generation of employees to leave the problem research-take chain catering enterprises in Anhui Province as an example (tzpyxj130). Anhui Vocational College of Finance and Trade, Quality Improvement and Quality Training College undertakes types of projects—How does the hybrid teaching model contribute to the classroom revolution of “Chain Store Operation Management” under the background of the Three Education Reform (tzpysj176). The National Natural Science Foundation of China supported the project “Research on the Influencing Factors of Energy-saving Behaviors of Tourism Enterprises and Low-Carbon Guiding Policies under Carbon Emission Constraints” (41801137). This work is also supported by the Heilongjiang Province Philosophy and Social Science Research Planning Project

“Research on Sustainable Development Countermeasures of Rural Tourism in Heilongjiang Province under the Background of Rural Revitalization”(20JYE276)

References

- [1] S. Ramakrishnan, “Fundamentals of machine learning for predictive data analytics: algorithms, worked examples, and case studies,” *Computing Reviews*, vol. 57, no. 6, pp. 346–346, 2016.
- [2] J. Schmidhuber, “Deep learning in neural networks: an overview,” *Neural Networks*, vol. 61, pp. 85–117, 2015.
- [3] S. Yi, X. Wang, and X. Tang, “Deep learning face representation by joint identification-verification,” *Advances in Neural Information Processing Systems*, vol. 27, 2014.
- [4] A. A. Adegun, S. Viriri, and R. O. Ogundokun, “Deep Learning Approach for Medical Image Analysis,” *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 6215281, 9 pages, 2021.
- [5] L. Geert, K. Thijs, E. Babak, A. Bejnordi, and A. Arindra, “A survey on deep learning in medical image analysis,” *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.
- [6] B. Ghoshal, A. Tucker, B. Sanghera, and W. L. Wong, “Estimating uncertainty in deep learning for reporting confidence to clinicians when segmenting nuclei image data,” in *2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 318–324, IEEE, Cordoba, Spain, 2019.
- [7] G. Chen, Q. Pei, and M. M. Kamruzzaman, “Remote sensing image quality evaluation based on deep support value learning networks,” *Signal Processing Image Communication*, vol. 83, article 115783, 2020.
- [8] G. Sun and S. Bin, “Router-level internet topology evolution model based on multi-subnet composited complex network model,” *Journal of Internet Technology*, vol. 18, no. 6, pp. 1275–1283, 2017.
- [9] F. Chollet, “Xception: deep learning with depthwise separable convolutions,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1800–1807, IEEE, Honolulu, HI, USA, 2017.
- [10] V. Gulshan, L. Peng, M. Coram et al., “Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs,” *JAMA*, vol. 316, no. 22, pp. 2402–2410, 2016.
- [11] J. Weston, F. Ratle, H. Mobahi, and R. Collobert, “Deep learning via semi-supervised embedding,” in *Neural Networks: Tricks of the Trade. Lecture Notes in Computer Science*, vol. 7700, G. Montavon, G. B. Orr, and K. R. Müller, Eds., pp. 639–655, Springer, Berlin, Heidelberg, 2012.
- [12] K. F. Harrington, D. Binkley, K. D. Reynolds et al., “Recruitment issues in school-based research: lessons learned from the High 5 Alabama Project,” *Journal of School Health*, vol. 67, no. 10, pp. 415–421, 1997.
- [13] S. Paramasivan, R. Huddart, E. Hall, R. Lewis, A. Birtle, and J. L. Donovan, “Key issues in recruitment to randomised controlled trials with very different interventions: a qualitative investigation of recruitment to the SPARE trial (CRUK/07/011),” *Trials*, vol. 12, no. 1, pp. 1–15, 2011.
- [14] E. J. Jackson and C. P. Parks, “Recruitment and training issues from selected lay health advisor programs among African Americans: a 20-year perspective,” *Health Education & Behavior*, vol. 24, no. 4, pp. 418–431, 1997.

- [15] L. Marks and E. Power, "Using technology to address recruitment issues in the clinical trial process," *Trends in Biotechnology*, vol. 20, no. 3, pp. 105–109, 2002.
- [16] Y. Wan-su and Z. Chi, "Research on the problems and countermeasures of SME recruitment from the perspective of demand theory," *Technology and Industry*, vol. 10, no. 4, pp. 36–39, 2010.
- [17] Y. J. Kim, T. W. Kim, J. S. Yoon, and M. K. Kim, "Study of the construction of a coastal disaster prevention system using deep learning," *Engineering and Technology*, vol. 33, no. 6, pp. 590–596, 2019.
- [18] Y. Alsulami, "An evaluation of a fourth level English textbook used in secondary schools in Riyadh City," *English Language Teaching*, vol. 14, no. 2, pp. 68–83, 2021.
- [19] G. B. Chen, L. Wang, and M. M. Alam, "Intelligent group prediction algorithm of GPS trajectory based on vehicle communication," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 3987–3996, 2021.
- [20] H. Haotian, W. Dongbo, D. Sanhong, S. Tianrui, and Y. Wenhao, "Analysis of information science education and personnel training based on information science recruitment entity mining," *Information Theory and Practice*, vol. 44, no. 1, pp. 8–17, 2021.
- [21] J. Lafferty, A. McCallum, and F. C. Pereira, "Conditional random fields: probabilistic models for segmenting and labeling sequence data," in *Proceedings of the Eighteenth International Conference on Machine Learning*, pp. 282–289, San Francisco USA, 2001.
- [22] G. Lample, M. Ballesteros, and S. Subramanian, "Neural architectures for named entity recognition," in *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, San Diego, California, 2016.
- [23] H. Shuiqing, W. Dongbo, and H. Lin, "Research on automatic recognition model of ancient Chinese place names based on pre-Qin corpus," *Library and Information Work*, vol. 59, no. 12, pp. 135–140, 2015.
- [24] H. Wei, H. Jianqiao, and L. Yuefeng, "Research on entity recognition model of terrorism-related information based on BiLSTM-CRF," *Intelligence Journal*, vol. 38, no. 12, article 149156, 2019.
- [25] D. Shengchun, F. Zhen, and W. Nan, "Named entity recognition in business domain based on BiLSTM-CRF," *Modern Intelligence*, vol. 40, no. 3, pp. 103–110, 2020.