Enhancing Talent Intelligence Evaluation with Improved Pre-training Models

Youren Chen[1,2], Yong Li[1], Ming Wen[2], Xiaohui Peng[2]*

[1] *College of Computer Science and Technology, Xinjiang Normal University, China*
[2] *Xinjiang Electronics Research Institute, China*
cyr0808@163.com, liyong@live.com, wmconet@126.com, 32332786@qq.com

Abstract

In the past few years, there has been a noticeable surge in the utilization of artificial intelligence techniques within the realm of talent evaluation, particularly in response to the exponential growth of big data. Conventional talent evaluation approaches, typically relying on the integration of expertise and system-based title appraisal, are plagued with challenges including errors in subjective judgment. The strategic importance of intelligent talent assessment cannot be overemphasized, especially for discovering and fostering high-potential talent across various domains, inclusive of research. This paper introduces an innovative method, termed TE-RCB, auxiliary applications in the evaluation of actual talent titles. The initial stage of this method's execution involves the meticulous construction of a talent evaluation dataset. This data is sourced from actual talent title applications and standardized scores provided by simulation expert engineers. The subsequent step is the application of the RoBERTa-WWM-large model, employed for the vectorization of pertinent information indispensable for talent evaluation. Further, the TextCNN model is utilized for the extraction of critical features from the talent-related information, succeeded by the application of the BiLSTM model to delve into deeper semantic correlations within the talent attributes. Comprehensive experiments are conducted on both the talent evaluation dataset and the publicly available dataset. The empirical outcomes decisively underline the superior accuracy of the proposed method, asserting its efficacy in addressing the pertinent challenges in the domain of intelligent talent assessment.

Keywords: Talent intelligence evaluation, Pre-training models, Deep learning

1 Introduction

In the context of globalization and informatization, the importance of talent is increasingly prominent, playing a significant role in the progress of a country and acting as a key factor in promoting economic development. Highly qualified and specialized talent can create value and enhance industry competitiveness, thereby driving economic development [1]. Talent constitutes the primary agent of innovation. It is through the presence of a cadre of high-caliber individuals that innovation can be propelled, enhancing innovative capabilities as an essential factor for both enterprise and national development. Talent represents the bedrock of societal progression. Exceptionally qualified individuals can make significant contributions across various societal domains, stimulating societal progress and elevating living standards [2].

The role of talent evaluation in talent development is unquestionable [3]. Talent evaluation is a tool for assessing and evaluating individual abilities and contributions, bearing profound significance [4]. Firstly, talent evaluation helps to increase individual development opportunities, enabling the discovery of individual strengths and weaknesses, and thus providing personalized training and development plans. Additionally, talent evaluation can promote the growth and expansion of an organization. Organizations need high-quality talents to support their development, and through evaluation, talents can be motivated to work more effectively for the organization, thereby improving its competitiveness and productivity. Finally, talent evaluation serves as an important benchmark for societal assessment of talent contributions [5]. Talent evaluation can assist organizations and countries in selecting and cultivating talents more scientifically, allowing talents to objectively understand their strengths and weaknesses, and thus formulate more reasonable career development plans. Concurrently, talent evaluation can stimulate talent competition, unlock talent potential, and promote the continuous development and progress of talents [6]. However, accurately evaluating talents has become a topic of great concern.

Different countries employ different methods and standards for talent evaluation [4], primarily utilizing the authoritative professional title review method of the government to evaluate talents. For example, Chinese university professional title reviews undertake a comprehensive evaluation based on academic achievements, technological innovation, and social services [7-8]. In contrast, the United States pays more attention to individual professional abilities, academic achievements, teaching quality, and management capabilities in the evaluation of talents in universities [9]. Different evaluation standards and methods can be applied to talents of different types and

*Corresponding Author: Yong Li; E-mail: liyong@live.com
DOI: https://doi.org/10.70003/160792642024112506011

backgrounds, to achieve fairness and scientific accuracy in talent evaluation [10]. In recent years, with the development of technologies such as artificial intelligence, intelligent methods have become helpful in assisting talent evaluation [11]. Data analysis can be carried out from multiple dimensions such as personal behavior trajectory, social network, and academic achievements, to predict future performance and potential, thereby providing a more accurate basis for talent selection and use [12]. However, despite these advantages, current traditional talent evaluation methods are still largely based on expert scoring and there are problems such as subjective judgment [13].

The talent professional title review is a process of grading and evaluating the professional technical positions of professionals, aimed at better managing and utilizing talent. This process involves multiple aspects of inspection and evaluation, including a comprehensive assessment of the candidate's educational background, work experience, professional skills, research outcomes, and social contributions.

Title reviews are typically conducted by professional review committees. These committees comprise peer experts who conduct reviews based on a set of established standards and procedures, primarily manually. To date, no intelligent evaluation methods have been applied in this field. The talent professional title review, led and issued by the government, serves as an important measure of talent capability assessment, and it is also the main application scenario of the intelligent talent evaluation proposed in this paper. As illustrated in Figure 1, a comparison is made between the traditional talent professional title review process and the intelligent talent professional title review process proposed herein.

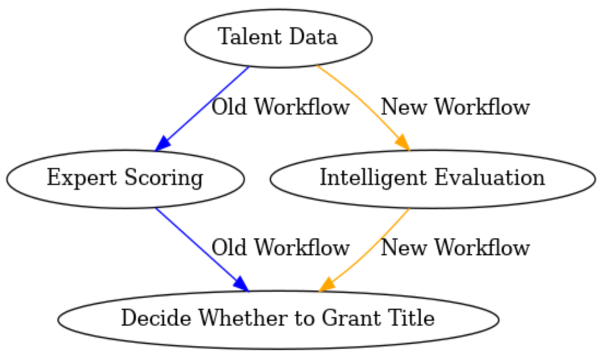

Figure 1. Comparison between this paper's title evaluation method and the traditional manual title evaluation method

Pretrained models have experienced swift progression, with exemplars such as GPT-4 [14] demonstrating remarkable efficacy in various natural language processing tasks, including automated Q&A and text dialogue. However, these models might not provide the optimal solution for evaluating talent in specific sectors, where it is crucial to assess talent fairly and impartially based on historical data.

In response, our study introduces a novel talent intelligence evaluation method. This approach leverages a combination of an enhanced pre-trained model, a convolutional neural network (CNN) [15], and a bidirectional long short-term memory network (BiLSTM) [16] to develop a new model, referred to as TE-RCB. This model aims to emulate expert talent evaluation, thereby providing an objective, accurate, and efficient tool for talent assessment in various professional fields.

This paper unfolds in a structured manner across five chapters. The first chapter introduces the background and presents an overview of talent, talent evaluation, and intelligent talent evaluation, thereby setting the stage for subsequent discourse. The second chapter explores related work on intelligent talent evaluation, text vectorization, and models used for extracting key textual information. The third chapter provides a detailed description of the construction process of the TE-RCB model, the primary intelligent talent evaluation method proposed in this study. The fourth chapter presents the experiments conducted and the corresponding results obtained from applying the proposed model to talent datasets and public datasets, followed by an analysis of these results. Finally, the fifth chapter summarizes the essence of the paper and outlines potential areas for future research in the increasingly important field of intelligent talent evaluation.

2 Related Work

Intelligent talent evaluation primarily involves the construction of models designed to learn the features of accumulated talent data. This learned information can then be used for the automated scoring of newly submitted talent data. This process paves the way for intelligent talent evaluation, making a revolutionary contribution to the field of professional title review. Currently, in the domain of natural language processing, pre-trained models have achieved superior performance in downstream tasks such as information retrieval and semantic understanding. These advancements provide robust support for intelligent talent evaluation, thus laying the foundation for the research presented in this paper [17].

2.1 Talent Evaluation

Talent evaluation encompasses an examination of multiple facets of an individual's capabilities, experiences, skills, and personal attributes. Numerous related works have been practically applied within the domain of talent assessment, such as professional title reviews. However, a majority of these efforts rely on manual labor, traditional data analysis systems, or a combination of both. To ascertain an individual's abilities with higher precision, a variety of evaluation techniques have been proposed, including interviews, questionnaires, simulated experiments, and work sample evaluations [11]. Yet, these methodologies often succumb to subjective influences and suffer from a lack of intelligence, thereby hindering the accurate and impartial evaluation of an individual's capabilities. This gap in talent evaluation can potentially be addressed by intelligent talent evaluation methods.

Intelligent talent evaluation is an assessment methodology grounded in artificial intelligence technology. It scrutinizes practical experience, project experience, and other pertinent data related to an individual, aiming to appraise their key capabilities and assign appropriate labels [17]. Intelligent talent evaluation facilitates a more precise understanding of a candidate's capabilities and potential, laying the groundwork for sustainable societal and national development [6]. This method of evaluation brings us one step closer to a future where subjective bias is minimized, and talent assessment is carried out efficiently and impartially.

2.2 Pre-training Models

Pre-trained models like BERT [18] have made significant strides by employing transformer structures, demonstrating impressive results in text, visual, and multimodal data feature extraction tasks since their inception. Owing to their training on large-scale corpora, these models acquire abundant language knowledge, enhancing their efficacy in rendering text as word vectors. Within the realm of talent intelligence evaluation, pre-trained models can process and analyze the available talent information as word vectors, thereby evaluating their skills and potential. Furthermore, pre-trained models can be fused with other machine learning algorithms and data mining techniques to augment the precision and efficiency of talent evaluation.

Pre-trained models have increasingly become a popular research direction in previous studies. Models like BERT and GPT series stand as representative pre-trained models that have yielded commendable results in various fields through fine-tuning or enhancements. For instance, the study [19] utilized BERT-TextCNN in financial text analysis. Paper [20] deployed RoBERTa-wwm for fine-tuning and enhancements in Chinese semantic matching tasks, and paper [21] proposed the RoBERTa-wwm-TextCNN method to comprehend the semantic relationship between Chinese medical classification and text in a multi-classification task, which impressive performance.

Nevertheless, these models may not be universally apt for talent evaluation in specific domains as domain-specific terminologies and particular text structures could impact the model's performance. Therefore, the aim of this study is to investigate how to adapt pre-trained models and tailor them to construct talent intelligence evaluation models that are suitable for specific domains. This approach promises a more refined tool for talent assessment, better calibrated to the intricacies of specialized fields.

2.3 Sentence Semantic Feature Extraction Model

LSTM and TextCNN models have demonstrated commendable performance in comprehending and extracting key information from text. For instance, Zhang et al [22] proposed a technique using character embedding combined with a multi-scale convolutional neural network (CNN) to extract contextual information from question or answer sentences at different scales. This method was further enhanced by integrating a multi-head attention mechanism, promoting superior learning of semantic features in Chinese medical Q&A text. In another study [23], ALBERT combined with TextCNN was used for extracting local information for disaster emotion analysis, achieving satisfactory results. The LSTM-CNN-CBAM algorithm was put forth in another study [24] to extrapolate the long-term impact of gold prices using LSTM and to mine the deep features of gold price data using CNN. The CBAM was then employed to improve the network's feature extraction capability, yielding promising results in gold price prediction. In yet another study [25], the MGBA-LSTM-CNN model was suggested, which encoded the problem and candidate attributes at both the semantic and word levels, concatenated them into a new semantic vector, and finally used cosine distance to measure the similarity between two vectors to identify the candidate attribute most akin to the problem.

While these models have demonstrated promising performance in their respective domains, constructing precise models for talent intelligence evaluation in a specific domain continues to be a challenging task. Generally, existing literature and methodologies for talent evaluation still remain in the phase of data mining and manual evaluation [17]. On the other hand, there have been significant advancements in pre-trained models and semantic feature extraction models.

In this study, we introduce a method called TE-RCB, based on an enhanced pre-trained model, to build a more precise and trustworthy model that simulates expert rating through historical talent data, especially in the field of job title review. This method can provide intelligent assistance to experts and offer a fresh perspective for future research on talent intelligence evaluation.

3 TE-RCB Model

This paper presents a novel talent intelligence evaluation model named TE-RCB. The model's architecture is depicted in Figure 2. During the initial stage of the model, the RoBERTa-WWM-large model is utilized to extract sentence-level semantic features from the talent assessment dataset. The extracted feature representations are then fed into the TextCNN model, which conducts a convolutional operation to obtain deeper sentence-level features. The output vectors from the TextCNN model are subsequently processed through a Bidirectional Long Short-Term Memory (BiLSTM) network. The BiLSTM network leverages both past and future contextual information to generate comprehensive and rich contextual representations of the talent data. Finally, the similarity between talent characteristics and expert ratings is computed, with vector pooling and normalization techniques being applied during this stage to enhance the effectiveness of feature learning. It's worth noting that layers of the talent information and score with the same color in the figure share identical parameters of neural networks. This cohesive structure further bolsters the capability of the model, leading to a more robust and efficient talent intelligence evaluation.

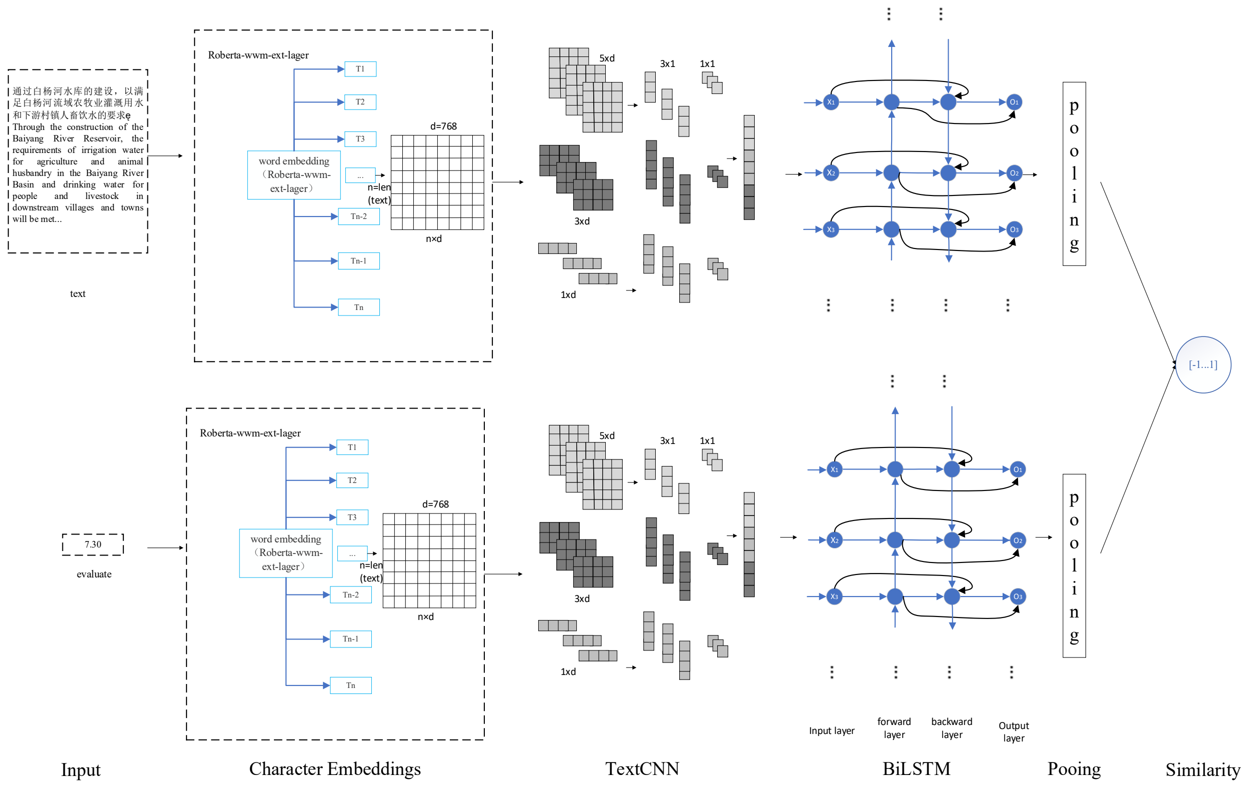

Figure 2. TE-RCB model

3.1 Embedding Vector Representation Layer

The embedding vector representation layer converts text sentences into two-dimensional vectors for recognition by neural network models. Traditional word embedding methods such as Word2Vec [26] and GloVe [27] provide a fixed vector representation for each word but cannot solve the problem of polysemy. In contrast, the RoBERTa-WWM [28] model provides a sentence-level feature vector representation for each sentence based on the context of a word in different sentence contexts, addressing the issue of polysemy. The RoBERTa-WWM-large model, with its increased number of parameters, possesses enhanced text representation capabilities. Hence, in this study, RoBERTa-WWM-large is selected as the word vector representation model.

In this study, we utilized the RWL (RoBERTa-WWM-large) [29] model to obtain sentence-level word vector representations and fine-tuned the pre-trained model parameters according to the dataset during the training process. The RWL model has four characteristics: (1) the sample generation strategy in the pre-training phase is changed to a whole-word masking strategy, but dynamic masking is not used; (2) NSP is removed to improve model efficiency; (3) the training dataset is larger, with longer training steps and longer training sequences; (4) the larger version of the parameters has stronger content and word vectorization capabilities.

The RWL model structure, depicted in Figure 3, initially accepts a text statement as input, symbolized by $W1, W2, \ldots, Wn$. The sum of the word vector, segment vector, and position vector of the statement serves as the model's input, represented by $E1$, $E2$, ..., En. The middle layer embodies the 12-layer bidirectional Transformer feature extractor, with $T1$, $T2$, ..., Tn being the output vectors of the model. Comprising 12 transformers, the word vector dimension of the RoBERTa-WWM-large model is 768.

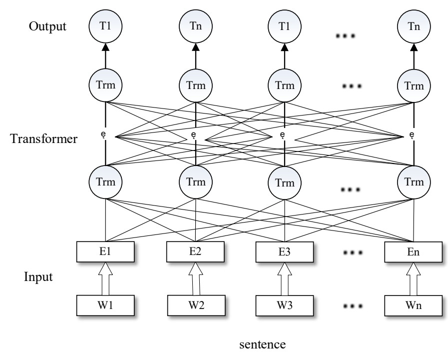

Figure 3. RWL model

$$\text{Svec} = [\text{RWL}([\text{S}])_{\text{xector\_out}} \quad \textbf{(1)}$$

where and are respectively the initial feature vectors of the text S: Svec, xector_out is out.

3.2 Semantic Acquisition Layer

The TextCNN model [23] utilizes multiple convolutional filters of varying sizes to extract crucial information from sentences, enabling the capture of local correlations and obtaining deeper semantic insights. In this study, the TextCNN model is employed, as depicted in Figure 4, which consists of embedding, convolutional, and pooling layers.

The semantic information acquisition layer is utilized to obtain deeper semantic and contextual information based on the sentence-level word embedding matrix obtained from the previous layer. For the word embedding matrix, TextCNN performs one-dimensional convolutional operations to extract multiple types of local features from the text for obtaining more comprehensive context information. To effectively extract deeper features, this paper sets three different sizes of convolutional kernels, 1, 3, and 5, and sets the padding value to the same to ensure that the dimension of the output vector of the convolutional layer is consistent. The pooling layer uses max pooling to extract more significant features from the convolutional layer output, and the pooled feature vectors are concatenated, as shown in equation (5).

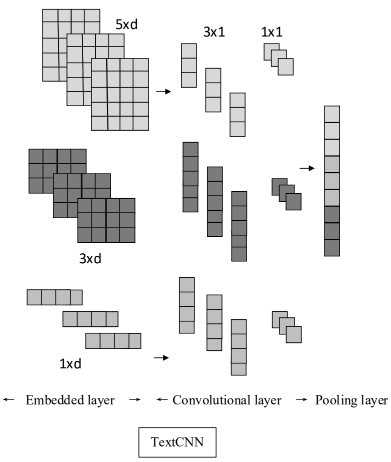

Figure 4. TextCNN model

This module employs three convolutional kernels with different sizes to extract local feature vectors of attribute vectors, similar to the n-gram approach, but avoiding the problems of data sparsity and the exponential increase in feature space size caused by increasing n. Then, the feature vectors obtained through the convolution operation are subjected to maximum pooling to obtain the key feature representation of the text, as shown in equation (3). Finally, the results of the maximum pooling operation are concatenated to obtain the feature vector, as shown in equation (4).

$$\text{pool}^{c_i} = \text{Maxpool}\left(\text{conv}\left(\widetilde{x^s}\right)\right) \quad \textbf{(2)}$$

$$\widetilde{z^s} = \text{pool}^{c_1} \oplus \text{pool}^{c_2} \oplus \text{pool}^{c_3} \quad \textbf{(3)}$$

where c_i denotes the convolution kernel of the size, $c_{\_i}\times n$ denotes the size of the convolution kernel set to 1,3,5 in this paper, and n is the number of attributes, where conv denotes the convolutional operation.

Ultimately, a linear transformation of ($\widetilde{z^s}$) is obtained through a fully connected layer, and then the extracted feature vector is obtained using the relu activation function.

$$Tvec = [\text{TextCNN}\ ([Svec])]_{\text{ector\_out}} \quad \textbf{(4)}$$

where Tvec is the initial feature vector of the text S, and ector_out is the output.

BiLSTM advantage in text feature learning lies in its ability to effectively capture talent contextual information from both preceding and succeeding words, enabling a comprehensive understanding of the text's context [30]. For the data in this paper, the input is the output $Tvec$ (X_1,X_2,X_3) of the TextCNN, the output is the $Bvec$ (O_1,O_2,O_3), The BiLSTM model structure is shown in Figure 5.

The computational formulas in the BiLSTM framework are central to the functioning of input gates, forget gates, output gates, and cell states. Suppose at time step t, the input is X_t in our model, this would be $Tvec(X1,X2,X3)$, the hidden state from the previous time step is $ht-1$, and the current cell state is ct. The computation for the BiLSTM can be detailed as follows:

Input gate (i_t): This gate controls the amount of information from the current time step input entering the cell state, calculated by the formula:

$$i_t = (w_i \cdot [h_{t-1}, x_t] + b_i) \quad \textbf{(5)}$$

Here, Wi and bi denote the weights and bias of the input gate respectively, while σ signifies the sigmoid function.

Forget gate (f_t): This gate regulates how much of the cell state from the previous time step is retained in the current time step. It is computed as follows:

$$f_t = \delta(w_f \times [h_{t-1}, x_t] + b_f) \quad \textbf{(6)}$$

Here, Wf and bf represent the weights and bias of the forget gate respectively.

Cell state (c_t): This computes the cell state for the current time step, which is jointly determined by the current time step input controlled by the input gate and the previous time step cell state regulated by the forget gate. The calculation is given by:

$$\tilde{c} = \tanh(w_c \cdot [h_{t-1}, x_t] + b_c) \quad \textbf{(7)}$$

$$c_t = f_t * c_{t-1} + i_t * \tilde{c} \quad \textbf{(8)}$$

Here, Wc and bc signify the weights and bias of the cell state respectively, and * represents the Hadamard product (element-wise multiplication).

Output gate (o_t): This gate controls the hidden state of the current time step, computed as follows:

$$o_t = \partial(w_o [h_{t-1}, x_t] + b_o) \quad \textbf{(9)}$$

Here, Wo and bo denote the weights and bias of the output gate respectively.

Hidden state (h_t): This computes the hidden state of the current time step, which is determined by the current time step cell state regulated by the output gate. It is calculated as:

$$h_t = o_t * \tanh(c_t) \tag{10}$$

$$Bvec = \text{BiLSTM}(TCvec) \tag{11}$$

The vector is fed into a BiLSTM layer to process the sentence simultaneously, to obtain comprehensive contextual information for the sentence. The input dimension of the BiLSTM network is consistent with the output dimension of the TextCNN layer, and the output is finally pooled, as shown in equation (11).

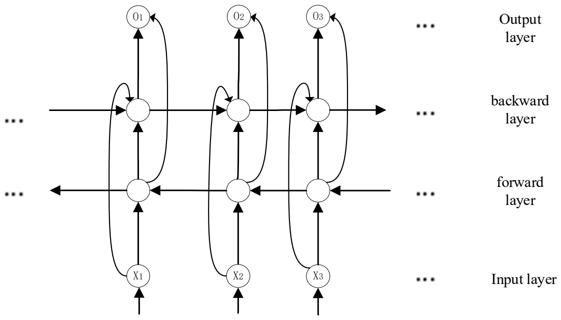

Figure 5. BiLSTM model

3.3 Normalization

Text vector normalization is a technique used in natural language processing (NLP) to scale down the magnitude of the vectors representing textual data. In NLP, each word or document is represented as a high-dimensional vector in a vector space, where each dimension represents a particular feature or attribute of the text. This layer normalizes embeddings to unit length, as shown in equation (12).

$$\text{normalized} = x / \|x\| \tag{12}$$

Here, $||x||$ represents the Euclidean norm or magnitude of the embedding vector x.

3.4 Loss Learning Layer

The loss function employed in this paper, referenced as equation [31], is particularly effective in cases where only positive pairs exist, such as paraphrase pairs, duplicate question pairs, (query, response) pairs.

To approximate the conditional probability $P(y|x)$, we leverage a set of K potential responses, comprising one correct response and K-1 randomly selected negative responses. For enhanced efficiency and simplicity, negative responses are selected from other examples within a stochastic gradient descent training batch. Specifically, for a batch of size K, we present K input emails $x=(x_1,x_k)$ and their corresponding responses $y=(y_1,y_k)$. In this setup, each response y_j serves as a negative candidate for the input x if i≠j. Owing to the nature of stochastic gradient descent, which shuffles the training data at each pass, the K-1 negative examples for each x vary in each iteration, as depicted in equation (13).

$$\begin{aligned} &\mathcal{J}(x,y,\theta) \\ &= -\frac{T}{K}\sum_{i=1}^{K}\log P\left(y_i \cdot x_i\right) \\ &= -\frac{I}{K}\sum_{i=1}^{K}\left[S\left(x_i,y_i\right) - \log\sum_{j=t}^{K}e^{s\left(x_i,y_j\right)}\right] \end{aligned} \tag{13}$$

Here, θ represents the word embeddings and neural network parameters used to calculate S. It is important to note that this loss function remains invariant to the addition of any function f(x) to S(x,y), thus, S(x,y) is learned up to an additive term that does not impact the max over y, which is performed during the inference time search.

4 Experimental Analyses

4.1 Experimental Environment and Parameters

The experimental setup employed in this study made use of Sentence-transformer in conjunction with the PyTorch deep learning framework, the computational tasks were executed on featuring Tesla P100 GPUs.

Key model parameters were defined as follows: the maximum sequence length, denoted as 'max_seq_length', was established at 64, while 'train_batch_size' was set to 16. In addition, 'predict_batch_size' was configured at 8, and the number of training epochs, represented as 'num_train_epochs', was set to 1.0.

4.2 Data Source

To make the model training and testing more realistic, this paper selects four main parts of data from the talent title evaluation process in China, which mainly relies on manual expert evaluation. These data are real data uploaded by talents to the evaluation platform in the talent title evaluation process of the Xinjiang government in China.

In the experimental section of this paper, we primarily use data submitted by talent, such as the four parts of Work Summary, Project content, Practical skills, Project results. Currently, the professional title review largely relies on expert experience for talent evaluation. By using the intelligent assessment model proposed in this paper, we aim to resolve issues such as subjective judgment, and time-consuming and labor-intensive expert evaluations.

The length of the talent data varies, but the scores are determined by expert evaluation. Finally, based on the scores, talents are ranked and determined whether to be awarded titles such as professor or researcher.

This study employed historical data from professionals with distinct technical expertise, sourced from the Xinjiang Professional and Technical Platform. The data was anonymized to ensure privacy. The total sample comprised 4,000 individuals, primarily concentrated within engineering and technical domains, which notably constituted a significant proportion of job summary data in title evaluations. The data were stratified based on the current representation of intermediate titles among professionals and technicians, offering a comparative study across high, intermediate, and entry-level titles at a ratio of 1:3:6, respectively. The selected dataset incorporated more than 12,000 sentences associated with work summaries and project descriptions—areas that are notoriously subjective and time-consuming for experts to evaluate, thereby leading to a high rate of misjudgments. Each sentence was assigned a score within the range of 0-10, based on the assessment by simulated experts, as detailed in Table 1, Table 2, Table 3, Table 4, Table 5.

To further validate the effectiveness of the model, this paper selects two distinctive datasets: a publicly available Douban question-and-answer dataset with short question-answer texts, and a long-text dataset where both questions and answers have an average length exceeding 50 characters cMedQA2 [22] datasets. These datasets are used to further verify the model's strong semantic understanding ability for short and long texts and its strong domain adaptability.

Table 1. Data source eg work summary

No.	Work summary	Scores
1	绿城玉园一期的样板段 1#、2# 楼地上三层石材施工过程核算中，发现石材线条复杂并且繁多，存着浪费资源增加工期的现象 … In the process of accounting for the stone construction on the ground three floors of the 1# and 2# buildings of the first phase of Green City Jade Garden, it was found that the stone lines were complex and numerous, which wasted resources and increased the construction period...	8.73

Table 2. Data source eg project content

No.	Project content	Scores
1	主持编制了颐和园小区一期项目工程，制定施工工序，明确合同约定要求。主要负责颐和园小区二期项目工程的图纸设计审验，及施工 … Presided over the preparation of the project engineering of the first phase of the Summer Palace Community, the development of the construction sequence, and the clarification of the contractual requirements. Mainly responsible for the project engineering of the second phase of the Summer Palace community, the design of the drawings for examination, and construction...	7.62

Table 3. Data source eg practical skills

No.	Project results	Scores
1	目前项目已竣工验收，水库完工后可以将安全运行中的隐患消除，同时绿化景观的保灌面积恢复至原设计的 1,045.9 亩 … The project has now been completed and accepted, and the reservoir will be completed to eliminate hidden dangers in safe operation, while the green landscape is restored to the original design of 1,045.9 mu of irrigation preservation…	5.78

Table 4. Data source eg project results

No.	Practical skills	Scores
1	在公司承建的《和静县国道 314 线岔口至克尔古提乡公路改建工程（第二合同段）》，四级公路沥青混凝土路面，长 13.47 公里，合同价款 1,984.5 万元 … In the company's construction of the Hejing County National Highway 314 line fork to Kerguti township road reconstruction project (the second contract section), grade four highway asphalt concrete pavement, 13.47 km long, the contract price of 19,845,000 yuan...	9.31

Table 5. Data set description

Data	No. Work summary	No. Project content	No. Practical skills	No. Project results
Total	3,000	3,000	3,000	3,000
Train sets	2,400	2,400	2,400	2,400
Test sets	600	600	600	600

To substantiate the efficacy of the proposed model, we purposefully eschew intricate feature engineering, sidestepping potential error magnification issues intrinsic to text similarity computation methods. Instead, we recast the problem as a question-answer matching task. The dataset is bifurcated into a training subset and a testing subset, at a proportion of 8:2. To manifest the potency of the model, experimental analyses are executed on this partitioned dataset.

4.3 Contrast Model and Evaluation Metrics

This manuscript aims to offer an intelligent assessment based on pre-existing talent data, conceptualized as a question-and-answer matching task. To substantiate the effectiveness of the TE-RCB model, we benchmark it against several high-performing baseline models in the domain of question-and-answer matching and related areas:

Chinese BERT [29] (2018): A model developed by Google, pre-trained and underpinned by the Transformer architecture, which is capable of producing dynamic word embeddings for text matching tasks.

SBERT [31] (2019): An enhancement of BERT that utilizes siamese and triplet networks to derive semantically rich sentence embeddings, which are compared via cosine similarity.

Chinese RoBERTa-WWM [29] (2019): Jointly proposed by the Harbin Institute of Technology and Microsoft, this model uses the Transformer architecture for unsupervised, bidirectional training.

Chinese RoBERTa-WWM-Large (RWL) [29] (2019): This is an augmented variant of RoBERTa-WWM that has been trained on 2.5TB of CommonCrawl data across 100 languages.

ERNIE 3.0 [32] (2021): Developed by Baidu, this Transformer-based pre-training language model is known for its multi-granularity and multi-task learning capabilities.

MACBERT [33] (2022): This is an enhanced version of the RoBERTa model that incorporates MLM as a corrective measure in its masking strategy.

TCBERT [34] (2022): Inspired by BERT, this model features custom adaptations for language and task-specific improvements.

GPT-RoBERTa model [35] (2023): This model combines a pre-trained language encoder, RoBERTa, with a decoder component, which excels at high-quality text generation tasks.

To evaluate the impact of the C module introduced in the TE-RCB model on the semantic feature learning capacity, and the ability of the B module to comprehend contextual information from talent data accurately, we have formulated two models, namely RWL-TextCNN, RWL-BiLSTM and RWL-BiLSTM-TextCNN, for comparative scrutiny.

As this part of the paper focuses on scoring talent characteristics, the highest correlation coefficient is needed to evaluate talent, using mean accuracy (MAP), normalized discounted cumulative gain (NDCG), and Accuracy as evaluation metrics for retrieval-based automated quizzes [36].

$$MAP@N = \frac{1}{Q}\sum_{q=1}^{Q}\sum_{k=1}^{N}(P(k) * rel(k)) \quad \textbf{(14)}$$

MAP@N is the average accuracy on the test set after model learning, and the formula is shown in 14.

$$\text{Accuracy@N} = (\text{TP@N} + \text{TN@N}) / \text{N} \quad \textbf{(15)}$$

Accuracy, as described by Equation X, quantifies the proportion of accurate model predictions within the first N results. Herein, True Positives (TP) represent instances correctly identified as positive, and True Negatives (TN) embody instances correctly recognized as negative, within these N predictions. N signifies the total count of predicted samples.

$$DCG_N = \sum_{i=1}^{N}\frac{2^{rel_i}-1}{\log_2(i+1)} \quad \textbf{(16)}$$

$$CG_N = \sum_{i=1}^{N} rel_i \quad \textbf{(17)}$$

$$NDCG@N = DCG@N / IDCG \quad \textbf{(18)}$$

Normalized Discounted Cumulative Gain (NDCG) is a measure where N represents the top N results assessed, DCG denotes the Discounted Cumulative Gain, and IDCG symbolizes the Ideal Discounted Cumulative Gain. $P(k)$ defines the precision of the top k responses for the present query. In this investigation, q corresponds to the quantity of query texts, and N assumes values of 1, 3, 5, and 10 et to affirm the robustness of the results. k designates the rank of a retrieved result in the outcome list, while reli ascertains whether the text at position k pertains to the query-answer relevance, with relevance denoted by 1, and irrelevance by 0.

4.4 Results and Analysis

4.4.1 Comparison and Analysis of Baseline Models versus Model Ablation

Table 6 indicates a significant enhancement in the performance of the proposed TE-RCB model. Compared to models 1-7, it surpasses them by 3% to 35% in MAP@10, NDCG@10, and Accuracy@10 metrics. These results highlight the TE-RCB model's ability to extract semantic information from talent assessment data efficiently, bolstering its utility in intelligent talent evaluation. Its improvement in the accuracy metric suggests that TE-RCB ensures high precision in talent datasets, making it a suitable choice for professional title evaluations. The increase in NDCG signifies that TE-RCB successfully heightens the correlation between talent assessment data and ratings, even in scenarios with many candidate ratings.

Table 6. Talent dataset model comparing MAP@10 and other evaluation results

id	Model	MAP@10 (%)	Accuracy@5 (%)	NDCG@10 (%)
1	BERT	81.51	91.02	83.44
2	Ernie3.0	77.13	88.00	81.10
3	GPT-RoBERTa	54.13	66.91	57.10
4	TCBERT	81.37	89.85	84,74
5	MACBERT	77.80	87.00	81.45
6	SBERT	79.91	90.35	83.06
7	RoBERTa-WWM	80.82	88.90	84.02
8	RWL	83.61	92.25	86.72
9	TE-RCB	85.05	93.05	87.91

Table 7 showcases how different models impact the learning of talent data features. The RWL-TextCNN and RWL-BiLSTM models exhibit similar performances on MAP@10, Accuracy@5, and NDCG@10 metrics, suggesting these two models have comparable effects on talent data feature learning.

The performance changes when BiLSTM and TextCNN are combined, regardless of whether BiLSTM is used before TextCNN or vice versa. Particularly, the RWL-BiLSTM-TextCNN model shows a decline in all three metrics, implying that the combination of BiLSTM and TextCNN might not be suitable for this talent data feature learning task.

Table 7. Talent dataset model modular ablation experiments

id	Model	MAP@10 (%)	Accuracy@5 (%)	NDCG@10 (%)
1	RWL- TextCNN	84.49	92.65	87.44
2	RWL- BiLSTM	84.56	92.65	87.43
3	RWL-BiLSTM-TextCNN	81.81	91.40	85.26
4	TE-RCB	85.05	93.05	87.91

However, the proposed TE-RCB model outperforms the other three models on all metrics, indicating that using TextCNN before BiLSTM is a more fitting approach for this task. This might be due to TextCNN superior capability in extracting text features and BiLSTM ability to better capture long-term dependencies in the features extracted by TextCNN.

Given the dearth of intelligent models in talent evaluation, conventional methods primarily rely on manual assessment.

However, after using the proposed model for evaluation, a simulation of 50 expert assessments on the talent data yielded an average error of around 0.5 points. This signifies that the model can provide satisfactory results in practical applications, marking a significant leap forward in the field of intelligent talent evaluation.

4.4.2 Experiment on Multiscale Convolution Kernels

During the construction of this model, the convolution window in the multi-scale convolutional neural network has a certain impact on the model. Therefore, this paper uses the real dataset of talent job title reviews to conduct experiments on the selection process of the convolution kernel. The model proposed in this chapter employs a multiscale convolution module, which utilizes convolution kernels of three distinct sizes for feature extraction from talent evaluation data vectors. Convolution layers of different sizes have varying feature extraction capabilities. To ascertain the optimal sizes for the three convolution layers, we experimented with this section. Referring to parameter settings in existing research, we selected candidate convolutions as [1,3,5], [2,3,4], [3,4,5], [4,5,6], and [2,4,6]. As illustrated in Figure 6, The model training loss values are as shown in the figure. The experimental results reveal that when the convolution combination is [1,3,5], the model exhibits the fastest decrease in loss and the best learning capability. This combination of multiscale convolution modules has a sufficiently large receptive field for feature extraction from attribute feature vectors relevant to talent evaluation.

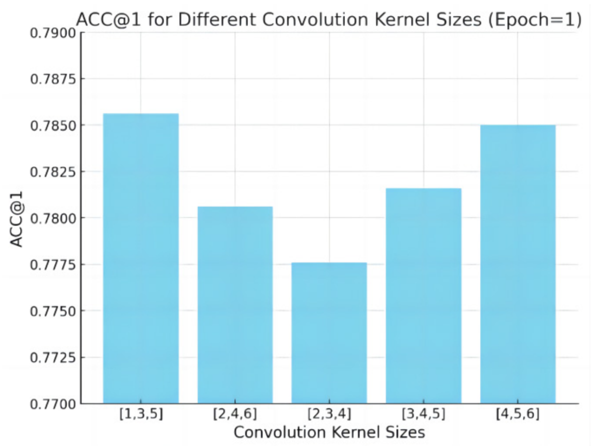

Figure 6. The impact of convolution kernel size on the efficiency of Accuracy@1

4.4.3 Hyperparametric and Time Analysis

Due to the complexity of the model, which exceeds the time of about half an hour for basic pre-training models like BERT when epoch=1, this paper conducts experiments on the main hyperparameters. For both the TE-RCB model and the RWL baseline model, each epoch takes approximately fifteen minutes with an improvement in accuracy of about 0.05%. Consequently, an epoch value of 1.0 was chosen. Suggesting that the hyperparameter configurations in the experimental setup were optimized, balancing time and model accuracy, as demonstrated in Figure 7.

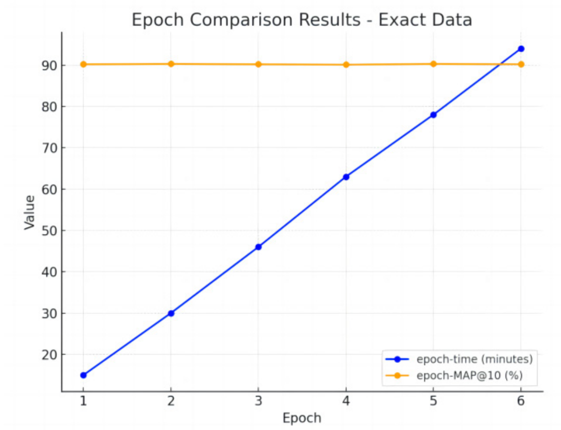

Figure 7. TE-RCB epoch impact time (minutes) or MAP@10(%)

4.4.4 Loss Function Analysis

During the experimental phase, the selection of an appropriate loss function emerged as a pivotal factor influencing the model's learning and training process. To assess the effectiveness of the loss functions chosen in this study, L1 and L2 were designated as baseline loss functions for experimental comparison. Both L1 and L2 represent variants of the contrast similarity loss function, typically harnessed in tasks necessitating the learning of ranking or the differentiation of samples based on their similarity. They exhibit proficiency in ensuring representational divergence of dissimilar samples. As illustrated in Figure 8, the chosen loss function in this study demonstrated superior performance in the task of learning features from talent data, as per the MAP evaluation metrics.

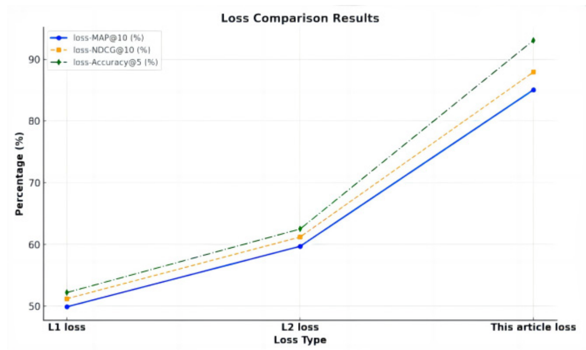

Figure 8. TE-RCB loss impact MAP@10 et.(%)

The contrastive loss values are represented by L1 and L2, defined as follows:

$$L1 = (1-y) * D^2 + y * \max(0, m-D)^2 \quad \textbf{(19)}$$

$$L2 = \max(0, D(A,P) - D(A,N) + \text{margin}) \quad \textbf{(20)}$$

L1, L2 is the contrastive loss value. y signifies a binary label denoting whether the pairs are similar (y=0) or dissimilar (y=1). D represents a chosen similarity metric, typically the Euclidean distance, between the representations of the anchor and positive (similar) samples. m denotes a margin hyperparameter that defines the minimum distance between the anchor and negative (dissimilar) samples. D(A,P) indicates the distance between the anchor (A) and the positive (P) samples, whereas D(A,N) signifies the distance between the anchor (A) and the negative (N) samples. The margin is a positive hyperparameter that establishes the minimum required difference between the distances of anchor-positive and anchor-negative pairs.

4.5 Results of Experiments on Public Data Sets and Experiment Analysis

In this investigation, the talent data was sourced from a specific subset, and a random sample of 100,000 question-answer pairs was extracted from the public DouBan movie dataset for experimental purposes. To minimize label discrepancies, the data format was conformed to match the format of the chosen talent data source. Given the uniformity of the labels used in this study, despite their varying lengths, and the predominance of short texts in the DouBan movie dataset, the cMedQA2 public dataset was employed to adjust labels and supplement the experimental evaluation, thereby corroborating the model's adaptability and scalability within its operational domain.

Throughout the experimental process of public datasets, we opted for several high-performing, standard methods, previously implemented in talent intelligence evaluation, to serve as the baseline models for comparison.

Table 8. Model comparison of MAP@100 and other rating results on the Douban Q&A collection

Model	NDCG@100 (%)	MAP@100 (%)	Accuracy@100 (%)
BERT	59.05	53.26	83.03
Ernie3.0	49.95	43.35	77.31
SBERT	37.83	30.68	68.49
GPT-RoBERTa	30.08	24.30	55.46
Macbert	59.05	53.26	83.05
RoBERTa-WWM	63.61	58.22	85.86
RWL	59.72	55.83	78.32
RWL- TextCNN	64.02	58.42	86.85
TE-RCB	66.86	58.84	92.30

Table 9. cMedQA2 dataset model comparing MAP@1000 and other evaluation results

Model	NDCG@1,000 (%)	MAP@1,000 (%)	Accuracy@1 (%)
BERT	78.98	72.90	62.40
MacBERT	79.44	74.72	63.28
Ernie3.0	77.27	71.01	61.17
RWL	80.57	74.79	64.32
RWL-TextCNN	81.63	76.09	65.70
TE-RCB	81.06	75.44	65.12

Given the multifarious nature of the DouBan movie dataset, and the potential for multiple responses to a singular query, the evaluation was methodically crafted around NDCG@100, MAP@100, and Accuracy@100 metrics. Concurrently, the medical Q&A dataset exhibits a similar trait, albeit with a distinct focus on model precision due to the constraints of the applied scenario. Consequently, NDCG@1,000, MAP@1,000, and Accuracy@1 were selected as evaluative indices to gauge the model's proficiency.

Table 8 illustrates that despite the RWL model employed in this research failing to surpass the performance of the RoBERTa-wwm model with a smaller parameter count on the Douban movie dataset, this is likely due to the inherently short nature of question-and-answer texts within this public dataset. Furthermore, the RoBERTa-wwm-large model, possessing a larger size and higher complexity, tends towards overfitting and has not undergone specific fine-tuning for this relatively small dataset. Contrastingly, our proposed TE-RCB model demonstrates a significant performance enhancement, with improvements of approximately 10% across NDCG, MAP, and Accuracy metrics compared to baseline models. This indicates its ability to surmount the constraints of RWL on small datasets, showcasing its robust generalization capacities.

For Table 9, which comprises longer texts, the semantic association between questions and answers is more focused on local key features, thus leading to a superior performance of the RWL-TextCNN model. However, our model remains competitive, surpassing some of the baseline models.

Upon experimental application to public datasets, the TE-RCB approach exhibits strong performance in both short-text and long-text matching tasks. While there is a marginal decrease in accuracy for long-text to long-text matching, it nonetheless highlights the robust learning capacity of the TE-RCB model when applied to talent evaluation data. This analysis corroborates the broad applicability of our proposed method in the sphere of intelligent talent evaluation, thereby underlining its generalizability.

5 Conclusion

In this research, we introduce an innovative TE-RCB framework designed for Intelligent Talent Assessment. This structure not only effectively harnesses semantic information via character-level embeddings, but it also adeptly extracts pertinent talent correlation data. This proposition is subjected to meticulous testing and validation, employing extant talent evaluation datasets. The experimental outcomes reported herein confirm that our model exhibits formidable performance, surpassing other state-of-the-art models applied to the talent dataset. Moreover, our empirical evaluations, carried out on two separate public datasets, substantiate the model's potent usability. Foreseeably, the incorporation of knowledge graphs in the talent domain may significantly enhance recommendation systems, providing improved services for talent and fostering more equitable evaluative methodologies.

Acknowledgment

This work is sponsored by the Xinjiang Key Research and Development Program (2022B01007-3), the Natural Science Foundation of Xinjiang Uygur Autonomous Region (2022D01A225), and the National Natural Science Foundation of China (62241209).

References

[1] D. Ulrich, N. Smallwood, What is talent?, *Leader to Leader*, Vol. 2012, No. 63, pp. 55-61, Winter, 2012.

[2] S. Beechler, I. C. Woodward, The Global "War for Talent", *Journal of International Management*, Vol. 15, No. 3, pp. 273-285, September, 2009.

[3] E. Gallardo-Gallardo, N. Dries, T. F. González-Cruz, Commentary on "The meaning of 'talent'in the world of work", *Human Resource Management Review*, Vol. 23, No. 4, pp. 290-300, December, 2013.

[4] P. van Arensbergen, I van der Weijden, P van den Besselaar, Different views on scholarly talent: What are the talents we are looking for in science?, *Research Evaluation*, Vol. 23, No. 4, pp. 273-284, October, 2014.

[5] H. Ma, A. Ding, Construction and implementation of a college talent cultivation system under deep learning and data mining algorithms, *The Journal of Supercomputing*, Vol. 78, No. 4, pp. 5681-5696, March, 2022.

[6] W. Wisetsri, C. Vijai, K. Chueinwittaya, P. Jirayus, Artificial Intelligence in Human Resources Management-An Overview, *Journal of Positive School Psychology*, Vol. 6, No. 2, pp. 2688-2693, 2022.

[7] G. Chen, X. Li, X. Zhang, G. Reniers, Developing a talent training model related to chemical process safety based on interdisciplinary education in China, *Education for Chemical Engineers*, Vol. 34, pp. 115-126, January, 2021.

[8] L. Yao, X. Li, R. Zheng, Y. Zhang, The impact of air pollution perception on urban settlement intentions of young talent in China, *International Journal of Environmental Research and Public Health*, Vol. 19, No. 3, Article No. 1080, February, 2022.

[9] N. Jokuschies, V. Gut, A. Conzelmann, Systematizing coaches' 'eye for talent': Player assessments based on expert coaches' subjective talent criteria in top-level youth soccer, *International Journal of Sports Science & Coaching*, Vol. 12, No. 5, pp. 565-576, October, 2017.

[10] L. Verginer, M. Riccaboni, Talent goes to global cities: The world network of scientists' mobility, *Research Policy*, Vol. 50, No. 1, Article No. 104127, January, 2021.

[11] A. McDonnell, D. G. Collings, R. Carbery, The Identification and Evaluation of Internal Talent in MNEs, in: D. G. Collings, H. Scullion, P. M. Caligiuri (Ed.), *Global Talent Management*, Routledge, 2011, pp. 1-16.

[12] N. Reyaz, G. Ahamad, N. J. Khan, M. Naseem, Machine Learning in Sports Talent Identification: A Systematic Review, *2022 2nd International Conference on Emerging Frontiers in Electrical and Electronic Technologies (ICEFEET)*, Patna, India, 2022, pp. 1-6.

[13] M. N. H. Hamidi, M. R. W. N. Wazir, A Systematic Review on Psychological and Physical Factors in Talent Identification, *Journal of Social Science and Humanities*, Vol. 5, No. 1, pp. 7-21, 2022.

[14] S. Bubeck, V. Chandrasekaran, R. Eldan, J. Gehrke, E. Horvitz, E. Kamar, P. Lee, Y. T. Lee, Y. Li, S. Lundberg, H. Nori, H. Palangi, M. T. Ribeiro, Y. Zhang, *Sparks of artificial general intelligence: Early experiments with gpt-4*, April, 2023, https://arxiv.org/abs/2303.12712v4

[15] S. Singh, A. Sharma, State of the Art Convolutional Neural Networks, *International Journal of Performability Engineering*, Vol. 19, No. 5, pp. 342-349, May, 2023.

[16] G. Dai, J. Zhang, X. Han, A Novel Attention-Based BiLSTM-CNN Model in Valence-Arousal Space, *International Journal of Performability Engineering*, Vol. 18, No. 12, pp. 833-843, December, 2022.

[17] L. Dong, Y. Liu, Design of performance evaluation algorithm for diversified talent training in modern universities considering innovative thinking, *Mobile Information Systems*, Vol. 2022, Article No. 2374468, June, 2022.

[18] J. Devlin, M. W. Chang, K. Lee, K. Toutanova, *Bert: Pre-training of deep bidirectional transformers for language understanding*, October, 2018, https://arxiv.org/abs/1810.04805v1.

[19] C. Wan, B. Li, Financial causal sentence recognition based on BERT-CNN text classification, *The Journal of Supercomputing*, Vol. 78, No. 5, pp. 6503-6527, April, 2022.

[20] L. You, F. Han, J. Peng, H. Jin, C. Claramunt, ASK-RoBERTa: A pretraining model for aspect-based sentiment classification via sentiment knowledge mining, *Knowledge-Based Systems*, Vol. 253, Article No. 109511, October, 2022.

[21] Y. Sun, D. Gao, X. Shen, M. Li, J. Nan, W. Zhang, Multi-Label Classification in Patient-Doctor Dialogues With the RoBERTa-WWM-ext + CNN (Robustly Optimized Bidirectional Encoder Representations From Transformers Pretraining Approach With Whole Word Masking Extended Combining a Convolutional Neural Network) Model: Named Entity Study, *JMIR Medical Informatics*, Vol. 10, No. 4, Article No. e35606, April, 2022.

[22] S. Zhang, X. Zhang, H. Wang, L. Guo, S. Liu, Multi-scale attentive interaction networks for chinese medical question answer selection, *IEEE Access*, Vol. 6, pp. 74061-74071, November, 2018.

[23] X. Zhang, Y. Ma, An ALBERT-based TextCNN-Hatt hybrid model enhanced with topic knowledge for sentiment analysis of sudden-onset disasters, *Engineering Applications of Artificial Intelligence*, Vol. 123, Article No. 106136, August, 2023.

[24] Y. Liang, Y. Lin, Q. Lu, Forecasting gold price using a novel hybrid model with ICEEMDAN and LSTM-CNN-CBAM, *Expert Systems with Applications*, Vol. 206, Article No. 117847, November, 2022.

[25] W. Liu, M. Fan, K. Feng, D. Guo, Automatic Chinese knowledge-based question answering by the MGBA-LSTM-CNN model, *AI Communications*, Vol. 36, No. 2, pp. 93-110, May, 2023.

[26] K. W. Church, Word2Vec, *Natural Language Engineering*, Vol. 23, No. 1, pp. 155-162, January, 2017.

[27] J. Pennington, R. Socher, C. D. Manning, Glove: Global vectors for word representation, *Proceedings of the 2014 conference on empirical methods in natural*

language processing (EMNLP), Doha, Qatar, 2014, pp. 1532-1543.

[28] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, V. Stoyanov, *Roberta: A robustly optimized bert pretraining approach*, July, 2019, https://arxiv.org/abs/1907.11692.

[29] Y. Cui, W. Che, T. Liu, B. Qin, Z. Yang, Pre-training with whole word masking for chinese bert, *IEEE/ACM Transactions on Audio*, *Speech, and Language Processing*, Vol. 29, pp. 3504-3514, November, 2021.

[30] S. Siami-Namini, N. Tavakoli, A. S. Namin, The performance of LSTM and BiLSTM in forecasting time series, *2019 IEEE International Conference on Big Data (Big Data)*, Los Angeles, California, USA, 2019, pp. 3285-3292.

[31] N. Reimers, I. Gurevych, *Sentence-bert: Sentence embeddings using siamese bert-networks*, August, 2019, https://arxiv.org/abs/1908.10084.

[32] Y. Sun, S. Wang, S. Feng, S. Ding, C. Pang, J. Shang, J. Liu, X. Chen, Y. Zhao, Y. Lu, W. Liu, Z. Wu, W. Gong, J. Liang, Z. Shang, P. Sun, W. Liu, X. Ouyang, D. Yu, H. Tian, H. Wu, H. Wang, *ERNIE 3.0: Large-scale Knowledge Enhanced Pre-training for Language Understanding and Generation*, July, 2021, https://arxiv.org/abs/2107.02137.

[33] Y. Cui, W. Che, T. Liu, B. Qin, S. Wang, G. Hu, *Revisiting pre-trained models for Chinese natural language processing*, November, 2020, https://arxiv.org/abs/2004.13922v2.

[34] T. Han, K. Pan, X. Chen, D. Song, Y. Fan, X. Gao, R. Gan, J. Zhang, *TCBERT: A Technical Report for Chinese Topic Classification BERT*, November, 2022, https://arxiv.org/abs/2211.11304.

[35] B. Guo, X. Zhang, Z. Wang, M. Jiang, J. Nie, Y. Ding, J. Yue, Y. Wu, *How close is chatgpt to human experts? comparison corpus, evaluation, and detection*, January, 2023, https://arxiv.org/abs/2301.07597.

[36] N. Thakur, N. Reimers, A. Rücklé, A. Srivastava, I Gurevych, *BEIR: A heterogenous benchmark for zero-shot evaluation of information retrieval models*, October, 2021, https://arxiv.org/abs/2104.08663v4.

Biographies

Youren Chen received the B.S. degree in Network Engineering from Xinjiang Normal University. He is currently a graduate student at Xinjiang Normal University. His current research interests include Deep Learning, Intelligent Systems and Knowledge Graph.

Yong Li received the Ph.D. degree in Computer Science from Nanjing University of Aeronautics and Astronautics in 2018. He is currently an Associate Professor of Xinjiang Normal University. His research interests include Machine Learning and Intelligent Software Engineering.

Ming Wen graduated from Xi'an University of Technology in 1988 with a major in automatic control. Now he is the director and researcher of software development and testing center of Xinjiang Electronic Research Institute. His research interests include Software Engineering and Artificial Intelligence.

Xiaohui Peng graduated from shihezi University in 1990 with a major in Software Engineering. Now he is the Deputy Director of the Software Development and Testing Center, and Senior Engineer at the Xinjiang Electronic Research Institute. His research interests include Software Engineering and Artificial Intelligence.