International Journal for Housing Science and Its Applications
Publish August 3, 2025. Volume 46, Issue 3 Pages 976-988

https://doi.org/10.70517/ijhsa46374

Research on vocational education curriculum design and employment adaptability based on AI algorithm under efficient collaborative education model

Jinyu Chang[1,*]
[1] Department of Ideological and Political Theory Teaching and Research, Qinghai Higher Vocational and Technical College, Haidong, Qinghai, 810700, China
Corresponding authors: (e-mail: jy0211760@163.com).

Abstract This study proposes an efficient collaborative vocational education model based on AI algorithms, focusing on the deep integration of curriculum design and employment adaptability. Methodologically, the intelligent organization and structured association of course content is achieved through knowledge graph technology, and a joint LSTM-BERT-Attention extraction framework is constructed to solve the difficulties of text named entity recognition and relationship extraction. The model validation shows that the learning rate after hyperparameter optimization is 2×10^{-5}, batch_size=16, and the F1 value of named entity recognition of LSTM-BERT-Attention on DCNE reaches 85.08%, which is significantly improved compared with the 78.73% of BiLSTM-CRF and the Flat model's 83.67% is significantly improved. The F1 value of this paper's model reaches 83.58% in course concept extraction for the dataset with 14 labels. In the knowledge graph construction, Neo4j visualization verifies the hierarchy and completeness of the knowledge network. The employment suitability experiment shows that the degree of modularization of 100%, the proportion of practical teaching of 20% and the intensity of technology tool integration of 1 correspond to the students' employment suitability ability of 17.31±1.07, 17.07±1.27 and 17.57±1.13, respectively, and the AI-driven curriculum design significantly optimizes the matching between the teaching structure and job skills.

Index Terms efficient collaboration, knowledge graph, named entity recognition, entity extraction, employment suitability

I. Introduction

In recent years, with the rapid development of social economy and the change of industrial structure, the status and role of vocational education have been paid more and more attention [1], [2]. The goal of vocational education is to cultivate talents matching the demand of the employment market and to provide human resources for the sustainable development of social economy [3], [4]. The job market is constantly changing, which brings challenges to vocational education [5]. In order to adapt to these changes, vocational education also needs to change. Traditional vocational education curricula focus primarily on theoretical knowledge and are often solidified in specific areas of specialization [6], [7]. But today's job market requires more integrated skills such as leadership, communication skills and creative thinking [8]. The key to achieving this goal lies in the scientific design of vocational education's professional curriculum with the industry's job market, and adapting to the changes in the market is mandatory for vocational education, which requires constant adjustment and innovation [9]-[12].

With the rise of Artificial Intelligence (AI) in education and technological innovations, it is possible to balance the design of vocational education programs with employment adaptability [13], [14]. The traditional curriculum design is mainly “theoretical knowledge-oriented”, which often produces students with a lot of theory but difficult to adapt to the job [15], [16]. In the development and design phase of the AI participation course, the integration of skills and knowledge should be solved, and the task-driven teaching mode is adopted [17]. Through the analysis of enterprise positions and specific enterprise work processes, the professional knowledge structure is reconstructed to closely focus on job skills, and through the guidance of enterprise roles, students are promoted to the cognition of the division of labor of work positions, cultivate the awareness of the work process of enterprise positions, and enhance students' adaptability to their positions [18]-[21].

This study centers on the collaborative education model based on AI algorithms, and systematically describes the curriculum design, data processing and deep learning-driven information extraction methods. Intelligent organization and structured presentation of course content is achieved through knowledge mapping technology, providing learners with personalized learning paths. Secondly, for the current situation of the lack of textual data in vocational education courses, the complete process of constructing high-quality datasets from multi-source PDF documents is proposed. And combined with the deep learning model LSTM and the Chinese entity-relationship joint extraction framework BERT-Attention, the key technical challenges in text named entity recognition and relationship extraction are solved.LSTM significantly improves the sequence modeling capability by introducing the mechanism of forgetting gate, input gate, and output gate in combination with the unit state preserving long-term memory. In the text named entity recognition task, LSTM is able to capture contextual semantic relations and achieve efficient annotation through the synchronized sequence structure with multiple inputs and multiple outputs.The BERT-Attention framework achieves end-to-end joint extraction through the collaborative work of the text encoding module and the relational attention module. The text encoding module utilizes the BERT pre-training model to extract word-level features, combines HanLP disambiguation with dependent syntactic analysis to construct word-level embedding vectors, and fuses syntactic structural features through graph convolutional networks. The relational attention module, on the other hand, dynamically assigns semantic weights in different relational subspaces, and enhances the association between entities and relations using the relation-aware adjacency matrix.

II. Efficient collaborative course design methodology and technical realization based on AI algorithms

II. A. Knowledge graph-based online course design

II. A. 1) Intelligent organization of course content

In the design of online courses based on knowledge graph, the intelligent organization of course content is a crucial step, which directly determines whether learners can efficiently acquire and understand knowledge, as well as the degree of interactivity and personalization of the course. This process involves the deep mining, hierarchical construction and associated presentation of course content, aiming to systematize and visualize knowledge and enable learners to explore and learn relevant course content according to their own needs and interests through an intelligent recommendation system.

Deep mining of course content is the basis of knowledge graph construction, which requires extracting core concepts and knowledge points from multiple sources of information such as textbooks, References, and online resources. This step is usually performed with the help of Natural Language Processing (NLP) technology, which performs preprocessing, keyword extraction, and topic modeling of the text to ensure the accurate representation of each knowledge point in the knowledge graph. Meanwhile, in order to ensure the completeness of the knowledge, the knowledge graph builder needs to have rich domain knowledge and in-depth understanding and analysis of the course content.

II. A. 2) Structured representation of knowledge graphs

By establishing hierarchical relationships between entities, such as subsets and supersets of concepts, as well as correlations and dependencies between knowledge points, knowledge mapping is able to demonstrate the hierarchical structure and logical relationships of knowledge. This structure facilitates learners to learn in-depth step by step according to their cognitive patterns and learning paths. In addition, the hierarchical organization of knowledge also helps teachers to rationally arrange the difficulty and progress of the course in the teaching design to ensure the continuity and progressivity of learning.

The relevance of knowledge is the core feature of knowledge mapping, which connects isolated knowledge points with the help of entity linking technology and relational labeling to form a rich knowledge network. This network enables learners to easily discover and acquire other knowledge related to a knowledge point, thus realizing interdisciplinary and cross-disciplinary learning and enhancing the breadth and depth of learning. For example, when learners learn “quantum mechanics”, the knowledge graph can automatically recommend “relativity theory”, “particle physics” and other related fields, encouraging learners to carry out cross-disciplinary This encourages learners to integrate learning across disciplines.

II. B. Course Text Dataset Acquisition and Processing

After completing the organization and structured presentation of course content based on knowledge graph, it becomes crucial to construct a high-quality dataset adapted to the deep learning model. This study is about the education courses in higher vocational colleges and universities as the research object, taking the computer science major as an example, since no dataset exists for the text named entity recognition task of the “Computer Networks” course, it is necessary to construct a dataset to provide model training. The original data was obtained from the PDF version of Computer Networks.

The total number of PDF files in the book is 9. Firstly, the PDF files were converted to txt file format, but during the conversion process, it was found that many of them were pictures, which could not be directly converted to txt format, so Baidu OCR was used for text recognition to extract the text from the pictures. Through this method successfully converted all the PDF documents into txt format, to ensure the integrity of the document content and editable.

The next step is to process the text in txt format, deep cleaning of text data to eliminate noise and irrelevant information in the data, such as the blank cells, formulas, pictures, misspellings and punctuation and other interfering factors, so as to enhance the quality and readability of data. Immediately after the processed data for the clause operation, in order to facilitate the next data set production work.

For formulas and pictures in the text, the strategy of direct deletion is adopted. Formulas and pictures often do not contain critical information and do not substantially affect the overall understanding of the text. Therefore, during the cleaning process, these non-text elements will be identified and deleted one by one to ensure the purity of the data. At the same time, the integrity of the sentence will be maintained as much as possible to avoid distortion or missing of the meaning of the sentence due to the deletion of formulas or pictures.

II. C.Deep Learning Based Recognition of Textual Named Entities

After obtaining the cleaned text data, how to realize named entity recognition by deep learning model is the next research focus. In this section, the improved LSTM method and its application in entity extraction for vocational education courses are proposed to address the problems of gradient vanishing and long distance dependence in text sequence modeling.

With the solution of the problems of gradient vanishing, long-distance dependence and text digitization, it makes deep learning-based models have further applications in the field of text recognition. Deep learning relies on the powerful nonlinear fitting ability of neural networks to optimize the training model through complex higher-order linear transformations with gradient descent. At the same time, considering the text context semantic relationship is also a key point of the model.

Recurrent neural network models (RNN) can model text temporal sequences and were first applied in natural language processing to model language models. Recurrent neural networks inevitably suffer from the problem of gradient vanishing when there are too many training layers, as shown in Figure 1. The input of the current node includes the current input x_t and the value of the hidden layer of the previous node h_{t-1}. The node color represents how much information is retained, with darker colors retaining more information and lighter colors indicating less. As the sequence is fed, the more previous nodes have less influence on the later ones. And when backpropagating, the more back nodes have less ability to update the weights of the front nodes due to the decreasing gradient.

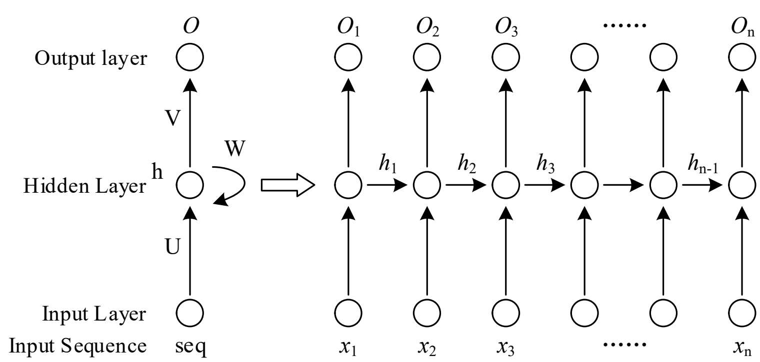

Figure 1: RNN gradient disappearance

The disadvantages of recurrent neural networks are more prominent due to the problem of gradient vanishing caused by long sequences, i.e., the problem of long-term dependence, which makes the model unable to take into account the information of the earlier moments, especially in the process of textual information processing, where there may be a large influence between the preceding and following texts. In order to improve this situation, researchers have proposed improved models for RNNs, such as BRNN, GRU, LSTM and so on. The most widely used among them is the LSTM model. Compared with recurrent neural networks, LSTM is also a chain structure, the difference is that LSTM handles input and output differently.LSTM adds a unit state c to the single state h of the hidden node of the RNN, which is specifically used to save the long-term state, in order to solve the long-term dependence problem. At the same time, the concept of “gate” is introduced to deal with the results of input and output, including forgetting gate, input gate and output gate.

(1) The forgetting gate is responsible for controlling how much of the output unit state c_{t-1} of the previous node can be preserved to the current state unit c_t:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_t\right) \tag{1}$$

(2) The input gate is responsible for controlling how much of the input state cell $\tilde{c}_t$ of the current node is retained to the current state cell c_t:

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) \tag{2}$$

After the control of the forgetting gate and input gate, the current state unit can be obtained, i.e., updated state information.

$$c_t = f_t \otimes c_{t-1} + i_t \otimes \tilde{c}_t \tag{3}$$

where $\tilde{c}_t$ is the current input cell state:

$$\tilde{c}_t = \tanh\left(W_c \cdot [h_{t-1}, x_t] + b_c\right) \tag{4}$$

(3) The output gate is responsible for controlling the effect of long-term memory on the current output:

$$h_t = o_t \otimes \tanh\left(c_t\right) \tag{5}$$

where o_t is the current input:

$$o_t = \sigma\left(W_o \cdot \left[h_{t-1}, x_t\right] + b_o\right) \tag{6}$$

Above is the input and output control of an LSTM node and the optimization of the long-term dependency problem, which can effectively solve the gradient vanishing problem.

The LSTM model is widely used as a sequence model, and has achieved good results in natural language processing tasks, including machine translation, automatic summary extraction, dialog systems, etc. LSTM can be divided into five types when dealing with different problems according to the input-output structure as shown in Fig. 2: the first one is shown in Fig. 2a, single-input-output, which is suitable for image classification; the second one is shown in Fig. 2b, single-input-multi output, applicable to image caption recognition; the third kind is shown in Fig. 2c, multiple inputs and single output, applicable to text sentiment analysis; the fourth kind is shown in Fig. 2d, multiple inputs and multiple outputs with coding and decoding, applicable to the field of machine translation and automated question and answer; the fifth kind is shown in Fig. 2e, synchronized sequences with multiple inputs and multiple outputs, applicable to the field of sequence annotation and named entity recognition.

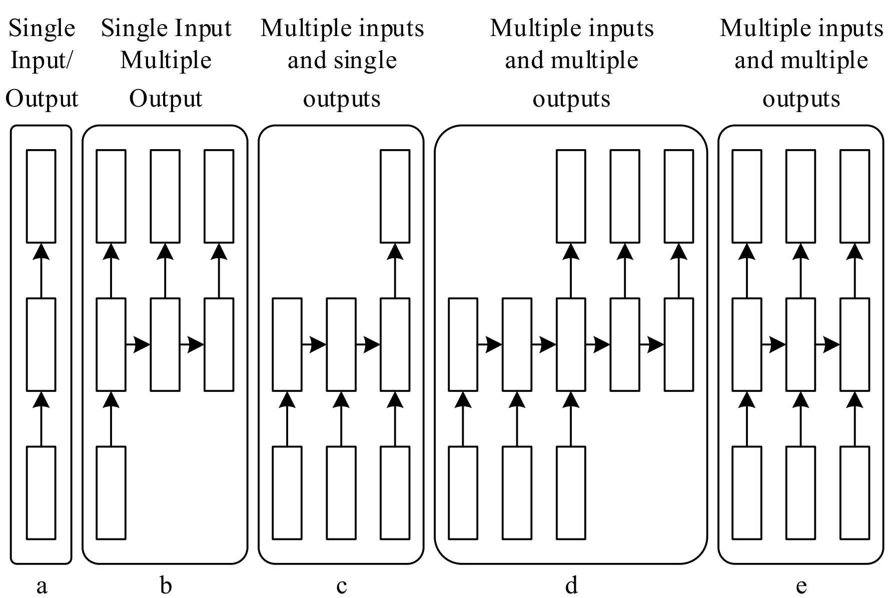

Figure 2: LSTM input and output structure

II. D. Overall Framework of Chinese Entity Relationship Joint Extraction Methods

On the basis of realizing named entity recognition, further mining the semantic relationships among entities is crucial for knowledge graph construction. In this section, we propose a joint extraction framework of Chinese entity relations that integrates BERT encoding and relation attention mechanism to break through the limitations of traditional step-by-step methods.

II. D. 1) Text encoding module

For Chinese natural language text, most of the previous studies have taken individual words as the smallest unit for extracting features. However, in Chinese, the smallest unit to express the actual meaning is a word, which leads to the fact that these researches only use a very limited number of features in the original text. On this basis, this paper, with the help of HanLP, a Chinese natural language processing tool, analyzes the original text for word segmentation and dependent syntax, and uses graph neural networks to fuse the dependent syntax tree of sentences with the features of phrases in order to obtain more features from the original text so as to improve the accuracy of information extraction. The working process of the text encoding module can be further divided into four steps:

(1) BERT encoding of the original text: since BERT is adopted as the encoder of the original text, the tokenization of the text still adopts the word-based processing method. In this step, the BERT model with Chinese full word masking is used as the text encoder to better process Chinese content.

Using T_i to denote the original text, t_i to denote the tokenized sequence, and h_i to be the sentence features encoded by BERT, we have Eqs. (7) & (8).

$$t_i = tokenizer\left(T_i\right) \tag{7}$$

$$h_i = BERT\left(t_i\right) \tag{8}$$

(2) Text segmentation and dependent syntax analysis: the text encoding obtained in the first step is not different from that in other methods, in order to further obtain more relevant features from the original text, this step is added in this paper, which is used to extract the relationship between words in a sentence. After the research and experimental comparison of existing Chinese natural language processing tools, HanLP was found to have perfect functions, high accuracy and excellent performance, so it was finally decided to choose HanLP as the tool to be used in this step. The processing of text using HanLP can be represented as shown in Equation (9):

$$W_i, G_i = HanLP\left(T_i\right) \tag{9}$$

where W_i is the disambiguation result of the sentence and G_i is the semantic dependency tree of the sentence, which is represented as the adjacency matrix of a directed acyclic graph as shown in Eqs. (10) and (11):

$$W_i = \left\{w_1, w_2, \cdots, w_n \mid w_i \in T_i\right\} \tag{10}$$

$$G_i = \left(w, e\right), w \in W_i, e \in \left\{(x, y) \mid (x, y) \in w \wedge x \neq y\right\} \tag{11}$$

In order to be able to learn the features of the words themselves in the subsequent graph neural network, it is necessary to add self-loop edges to this adjacency matrix, i.e., the diagonal elements of this matrix are assigned a value of 1, as shown in Equation (12):

$$G_i\left[i, i\right] = 1\left(0 \le i \le W_i.size\right) \tag{12}$$

(3) Word Embedding Vector Acquisition Combining Segmentation Results. In the first step, the text encoding in terms of words is obtained, but in the subsequent graph convolutional network, the model needs the text feature information in terms of words. Therefore, in this step, according to the results of text segmentation, the corresponding positions of the text encoding output by the BERT encoder need to be converged to the maximum value, and the word features are obtained from the single-word features, so as to meet the subsequent use of the model. For the word $p_i \in T_i$, assuming that it consists of three single words, which are denoted as w_x, w_y, w_z, the final feature vector of p_i can be expressed as shown in (13):

$$p_i = \max\left(w_x, w_y, w_z\right) \tag{13}$$

The final feature representation of the sentence T_i as h_i' has the formula (14):

$$h_i' = \{p_1, p_2, \cdots, p_n\} \tag{14}$$

(4) Sentence feature acquisition combining dependent syntax tree with word embedding vectors. In the first three steps, the model has obtained the text features in terms of words and semantic dependent neighbor matrix. This step then uses a graph convolutional neural network to learn the feature information in the sentence. This step is defined as in equation (15):

$$u_{ml}^{l} = \sigma\left(\sum_{j=1}^{n}\left(Wu_{mj}^{l-1}\right) + b\right) \tag{15}$$

where u_{ml}^{l} denotes the feature of the m th node in the graph G_i in the l th layer, W, b are the weights and bias of the model, respectively, σ is the sigmoid activation function, and $h_{i'}$, the sentence feature obtained in the third step, is input into the layer as the initial feature. The result obtained in this step, i.e., the result of the text encoding module, is expressed as in equation (16) below:

$$U = \{u_1, u_2, \cdots, u_n\} \tag{16}$$

II. D. 2) Relationship Attention Module

In the semantic space of different relations, the dependency weights of words are different from each other, and in order to extract these triples more efficiently, a relation attention mechanism is designed in this section to dynamically capture the correlation between words in each relation space. First, the model runs a fully-connected layer on top of the output of the previous module to obtain an initial hidden representation of the dependency weights, as shown in equation (17) below:

$$S = W_a U + b_a \tag{17}$$

where W_a, b_a are the weights and biases of the fully connected layer. Since the model needs to learn the semantic information in different relation spaces, it needs to compute the association feature information between nodes of different relations. Firstly, the feature representations of sentences are mapped into different relational subspaces, which are represented as in Eqs. (18) & (19):

$$Q^m = W_Q^m S \tag{18}$$

$$K^m = W_K^m S \tag{19}$$

where Q^m and K^m are the Query and Key vectors for the m th relation type, respectively, and W_Q^m and W_K^m are the weight parameters of the model. Next, a relation-aware adjacency matrix $A_m = \{A_{ij}^m\}$ is constructed for each relation, where A_{ij}^m is represented as in Equation (20):

$$A^m = \sigma\left(\frac{Q^m (K^m)^T}{\sqrt{d_r}}\right) \tag{20}$$

where d_r is the dimension of each relation space. Then, using the relation-aware adjacency matrix A_m as the adjacency matrix and the output U of the first step as the original input, the graph convolution operation is performed again. In this graph convolution, the model establishes connections for named entities and relations through the relation-aware matrix and semantic vectors, thus laying the foundation for effective extraction of triples in sentences. This step can be expressed as equation (21):

$$u_{mi}^{l} = \sigma\left(\sum_{j=1}^{n} A_{ij}^{m}\left(W_m u_{mj}^{l-1}\right) + b_m\right) \tag{21}$$

where u_{mj}^{l-1} is the feature of node j on the $l-1$ th layer in the space of relation m, and its original input is the output of the graph convolution network in the previous section.

Finally, the model combines the outputs of the two graph convolutional networks to obtain the final encoded feature q_i, as shown in Equation (22):

$$q_i = \left(\sum_{k=1}^{m} u_{ki}\right) + u_i \tag{22}$$

where u_i is the output of the first graph convolutional network and u_{ki} is the output of the second graph convolutional network.

III. Model performance validation and knowledge graph construction

After completing the construction of the intelligent organization and deep learning model of the course based on knowledge graph, it is necessary to verify its effectiveness in the actual teaching scenario through experiments. This chapter combines the text data of the Computer Networks course to compare the named entity recognition and relationship extraction performance of mainstream models such as LSTM and BERT-Attention, which provides technical support for the subsequent employment adaptability analysis.

III. A. Deep Learning Based Named Entity Recognition Study for Course Texts

III. A. 1) Hyperparametric experiments

Hyperparameter experiment is an important part of the model training process, usually, these parameters are not automatically adjusted during the training process and need to be set manually. Hyperparameter selection has a great impact on the performance of the model, including the training speed, convergence, capacity and generalization ability of the model. In order to find the best combination of hyperparameters to make the model perform best on a specific task, this study sets up hyperparameter experiments at the beginning of the experiment.

Hyperparameters include learning rate, batch_size, etc. Learning rate is one of the important hyperparameters, which controls the effective capacity of the model in a complex way. Choosing an appropriate learning rate can not only accelerate the convergence of the model, avoid falling into local optimum and reduce the number of iterations. The batch_size is the sample batch capacity, which determines the direction of model descent, so the hyperparameters of the model are adjusted for the learning rate and batch_size. Feed the processed data into the model and observe the change of loss function in the training cycle during the model training process, and set epochs to 30. Considering the effects caused by different batch_size and learning rate on the model performance, the batch_size was set to 4, 8, 16, 32, and the learning rate was set to 0.001, 1×10^{-4}, 1×10^{-5}, and 2×10^{-5}, respectively, and the F1 values of the model under each parameter were obtained as shown in Fig. 3.

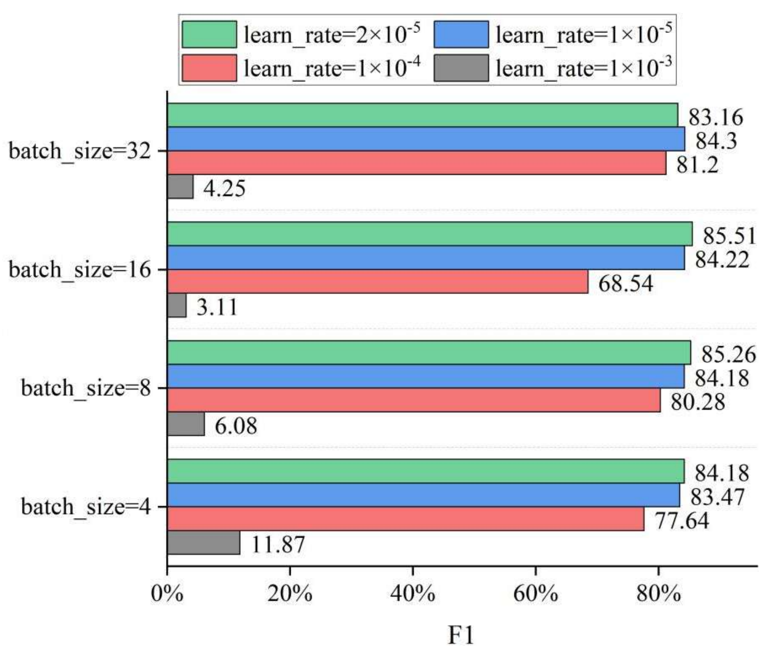

Figure 3: F1 values corresponding to multiple parameters

As can be seen in Figure 3, the model performs best overall when batch_size=16 and the learning rate is 2×10^{-5}, with an F1 value of 85.51%. The learning rate setting is analyzed and it is found that when the learning rate is set too large, the model convergence performance deteriorates and it is difficult to obtain useful information for prediction during the training process. When batch_size is too small, it may cause the model not to converge, as batch_size increases, the speed of processing the same amount of data is accelerated, the larger the batch_size is, the more epochs are needed, so when the batch_size increases to a certain value, the model's performance enhancement reaches the peak, and when the batch_size continues to increase The model performance gradually decreases, affecting the overall performance of the model.

III. A. 2) Model performance comparison

In order to verify the performance of the model constructed in this paper on text named entity recognition, a comparative study is conducted with the following models.

(1) BiLSTM-CRF: BiLSTM-CRF, a classical sequence annotation model for entity recognition task that uses pre-training to obtain word vectors, captures sequence features by BiLSTM, and CRF predicts labeled sequences.

(2) BERT-CRF: a pre-trained BERT-based model that combines the advantages of BERT's bi-directional coding capability and CRF, which can better capture the dependencies between the labels, with label decoding by CRF after generating vectors by BERT to complete text embedding.

(3) BERT-BiLSTM-CRF: It is a deep learning model combination for natural language processing tasks, which assembles the pre-trained BERT model, bi-directional LSTM network, and conditional random field to effectively improve the semantic understanding of the text and the sequence labeling ability.

(4) NFLAT: Based on the FLAT model, it can model word and phrase sequences of different lengths, reduce redundant computation and improve the relative encoding position, which in turn enhances the performance of the model.

(5) LR-CNN: By combining the local response mechanism with the convolutional neural network CNN, this model is able to effectively capture the local semantic features in the text. Its multi-scale convolutional kernel design enhances the model's ability to model short-range contextual relationships, and is suitable for dense recognition scenarios of technical terms in course texts.

(6) Adapting Transformer: as an improved version of the standard Transformer model, this model optimizes the semantic modeling ability of long sequence text by adaptively adjusting the attention weight and position encoding mechanism. Its multi-head attention layer can dynamically allocate the semantic association weights of different subspaces, which improves the generalization of complex course texts.

(7) Flat: The model adopts a flat structure design, which significantly improves the efficiency of named entity recognition by simplifying hierarchical dependencies and parallelizing the computational process. Its span-based decoding strategy can directly locate entity boundaries, reduce redundant computation, and show strong robustness in low-resource scenarios.

To verify that the model proposed in this paper has better performance on the named entity recognition task in the education domain, comparative experiments are conducted on the dataset DCNE. The mainstream models of named entity recognition task are selected for the experiment to compare with the recognition effect of this paper's model, and the experimental results of named entity recognition of each model on the DCNE dataset are shown in Fig. 4.

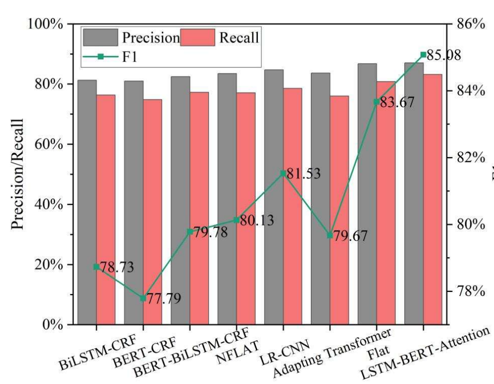

Figure 4: Experimental results of entity recognition for each model on DCNE dataset

As can be seen in Fig. 4, the BERT-CRF model achieves 77.79% F1 value and the BiLSTM-CRF model achieves 78.73% F1 value.The performance of BERT in named entity recognition task is verified by the experimental results of the BiLSTM-CRF model and the BERT-CRF model.The F1 value of the BERT-CRF model is 2.49% lower compared to the BERT- BiLSTM-CRF model's F1 value is 2.49% lower, which proves the importance of BiLSTM in feature extraction. In addition, the experimental results of the BERT-BiLSTM-CRF model also proved that after BERT completes the text embedding to generate the vectors, the feature extraction is performed by BiLSTM, and finally the label decoding is performed by using CRF, and the named entity recognition of this process is better than the other two models.The LR-CNN model achieved 81.53% of the F1 value on the DCNE dataset, and its local response mechanism and multi-scale convolutional kernel design effectively improve the recognition accuracy of technical terms; the Adapting Transformer model has an F1 value of 79.67%, and its adaptive attention mechanism is slightly weaker than the other baseline models in modeling long text; the Flat model stands out with an F1 value of 83.67%, and its flattened structure and spanning decoding strategy significantly optimize the entity boundary localization efficiency.

Through comparative analysis, the LSTM-BERT-Attention model in this paper performs the best with a higher F1 value than the other models, indicating that the model is suitable for the task of named entity recognition in the educational domain in low-resource scenarios. The experimental results also show that this paper's model, without incorporating other specific external features, dynamically captures the correlation between words in each relational space through the addition of the relational attention module and adopts a span-based decoding approach, which improves the model's recognition ability for named entities.

III. B. Comparative Experiments on Course Knowledge Extraction Models

After verifying the performance of the named entity recognition model, it is necessary to further evaluate the effectiveness of its application in the construction of knowledge graphs for real courses. This section reveals the role of AI technology in enhancing the integrity of knowledge networks by comparing the performance of different models in educational concept extraction tasks.

In order to continue to evaluate the performance of the model in the task of extracting educational concepts, the model was compared with the above seven model methods. In order to make the experiment universal and effective, we conducted experiments on two datasets with different labels, in which "One kind type" means that in the dataset, we uniformly labeled the concepts as "Concept" label entities for extraction and evaluation. "Fourteen kind type" means that we further divide the concepts in the dataset into fourteen label types, including fourteen types such as functions, models, and methods. Table 1 shows the results of model validity validation based on the computer course concept dataset.

Table 1: Comparison results of curriculum knowledge extraction models

	One kind type			Fourteen kind type		
	Precision	Recall	F1	Precision	Recall	F1
BiLSTM-CRF	75.97	72.32	74.10	70.22	67.89	69.04
BERT-CRF	73.38	70.78	72.06	68.37	66.83	67.59
BERT-BiLSTM-CRF	76.49	73.65	75.04	71.08	68.29	69.66
NFLAT	75.08	74.54	74.81	73.26	70.28	71.74
LR-CNN	78.61	74.54	76.52	75.82	72.41	74.08
Adapting Transformer	82.06	79.97	81.00	78.53	74.52	76.47
Flat	85.04	80.25	82.58	80.53	76.99	78.72
LSTM-BERT-Attention	90.75	86.27	88.45	87.54	79.96	83.58

As shown in Table 1, the experimental results of each model on two differently labeled datasets are shown, and we can see that the P value, R value and F1 value of this model reach 90.75%, 86.27% and 88.45%, respectively, on the "One kind type" dataset of computer science course concepts, which are higher than other baseline methods. Comparing to the BiLSTM-CRF model, our model's performance is significantly improved due to the incorporation of a third-party data dictionary of specialized course concepts with the large-scale pre-training model BERT. In addition, we compared three models including word information, such as NFLAT, LR-CNN and Flat, which are not as effective as our model, and here we also illustrate the importance of incorporating a third-party dictionary of specialized course concepts in the process of fusing word information.

Secondly, in order to further verify the dictionary-based course concept extraction method proposed in this chapter, we conduct experiments on the "Fourteen kind type" dataset to further divide the types of course concepts, and we can see that the P values, R values, and F1 values on the "Fourteen kind type" dataset reach 87.54%, 79.96%, and 83.58%, which are higher than the P values, R values, and The F1 value decreased by 3.54%, 7.31% and 5.51% respectively, and the reason for the decrease was that when we increased the entity types of sequence annotation, when we evaluated the prediction accuracy and recall rate, we needed the characters annotated by the model to correspond to the corresponding labels one by one, so the P value, R value and F1 value after subdividing the course concept into specific categories all decreased to a certain extent, but compared with other baseline methods, the decline of the model in this paper was small. It shows that the proposed model has better generalization and robustness. In addition, when extracting multiple types of course concepts, the P-value, R-value, and F1 value of the proposed model are better than those of other baseline models.

III. C. Visualization of Knowledge Mapping of Computer Curriculum System in Vocational Education

The optimization of the knowledge extraction model lays the foundation for the visual presentation of the course system. In this section, the Neo4j tool is used to graphically present the core knowledge points and associated paths of the Computer Networks course to visually reflect the contribution of the AI algorithm to the optimization of the teaching structure.

In order to more intuitively show the structure and characteristics of the knowledge graph of the computer course system in vocational education, the Neo4j Browser can be used to present the architecture diagram of the whole computer course system. The visual presentation of the knowledge graph of the Computer Networks course is shown in Figure 5. In this architecture graph, courses are represented in the form of nodes, and relationships between courses (e.g., prerequisite relationships, support relationships, etc.) are represented in the form of edges, and different nodes and edges are differentiated by attributes such as color, size, and shape, so as to facilitate users' quick identification and understanding. For example, basic courses can be represented by larger nodes, while specialized courses can be represented by smaller nodes; prerequisite relationships can be represented by solid lines, while support relationships can be represented by dashed lines. In this way, users can clearly see the position and role of different courses in the whole system.

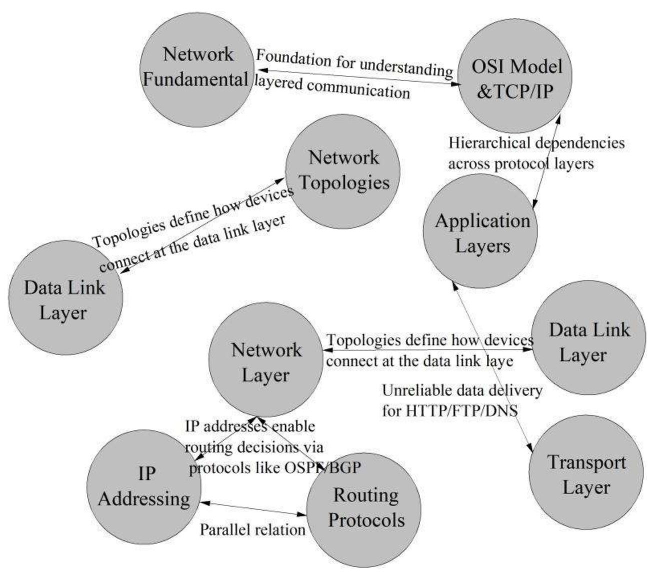

Figure 5: Visualization of knowledge graph of Computer Network course

In the visual architecture diagram, the core courses usually have more connections and higher centrality, which are the foundation and support of the whole curriculum system, and play an important role in cultivating students' core competencies and qualities. With the help of Neo4j's graphical query language Cypher, these core courses can be easily identified and highlighted, and the critical paths between courses can also be analyzed, i.e., the series of courses that students must go through to complete their studies. In the visual display, the critical paths can be represented by thick lines or different colors for easy identification and tracking by the user. In addition to showing the structure of the course system, the visualization tool can also identify potential duplication of content by comparing the similarities between instructional units and knowledge points in different courses. In Neo4j, methods such as graphical query and cluster analysis can be used to identify these duplicated contents and label and prompt them in the visualization display. For example, recurring instructional units or knowledge points can be differentiated by the same color or marking, and accompanied by a detailed description of the repetition and suggestions.

IV. Differential testing and analysis of factors related to the design of courses affecting students' employment adaptability

The experimental results show that AI-driven course design significantly improves knowledge extraction and learning path optimization. To further explore its actual impact on employment adaptability, this chapter quantifies the correlation between curriculum design factors and students' abilities through t-test and ANOVA based on three dimensions, namely, modularization degree, proportion of practical teaching and intensity of technology tool integration.

A total of 167 students were randomly selected from the class of 2024 computer science majors of a higher vocational college as the research object, applying different degrees of modularization, different proportions of practical teaching modes, and different intensities of technology tool integration methods to conduct the vocational education course, and testing and recording the students' employment adaptive ability.

The total score of the students' employment adaptability test totaled 20 points, of which 0-7 is failing; 7.1-12 is passing; 12-16.9 is good grade; and 17-20 is excellent grade.

The test of variance is mainly used to infer the probability of the occurrence of differences using the distribution of the departure statistic, so as to compare whether the difference between two means is significant or not. In this paper, three aspects of the degree of modularization in course design, the proportion of practical instruction, and the intensity of integration of technological tools were selected to conduct t-test or ANOVA to analyze the differences in students' employment adaptability.

IV. A. Analysis of differences in the degree of modularization on students' adaptability to employment

The analysis of differences in the degree of modularization on students' adaptability to employment is shown in Figure 6.

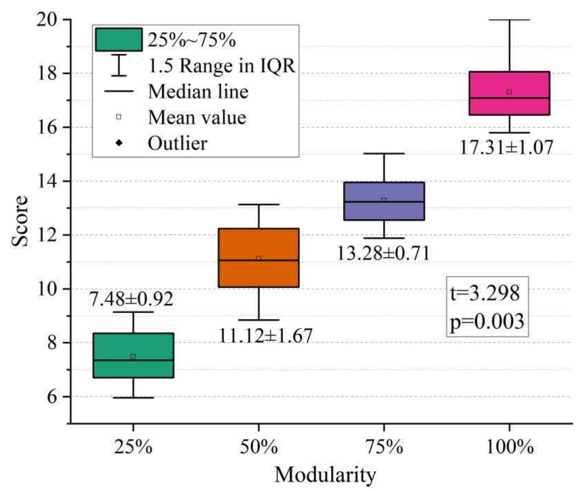

Figure 6: Analysis of the difference between modularity and employment adaptability

The data show that as the degree of modularization of the course increases 25% → 100%, the mean value of students' employment adaptability significantly increases from 7.48 to 17.31 points, and the variance decreases from 0.92 to 1.07, indicating that the higher the degree of modularization, the more concentrated the distribution of students' abilities. When the degree of modularization reaches 100%, the mean score reaches an excellent grade of 17.31 points, and the t-test result $t=3.298$, $P=0.003<0.05$, verifies the significant positive effect of the degree of modularization on employment adaptability. This indicates that the systematic knowledge mapping-driven curriculum design can help students construct a complete skill system and enhance job adaptation ability.

IV. B. Analysis of the difference in the proportion of practical teaching on students' adaptability to employment

The significant effect of the degree of modularization validates the importance of knowledge systematization. On this basis, this section further analyzes the nonlinear effect of the proportion of practical teaching on employment adaptability, revealing the criticality of the balance between practice and theory. The analysis of the difference between the proportion of practical teaching on students' employment adaptive ability is shown in Figure 7.

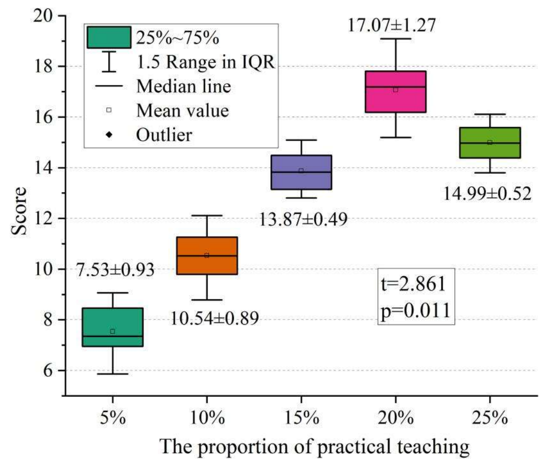

Figure 7: Analysis on the difference of employability adaptability in practice ratio

The mean score of the students gradually increased from 7.53 to 17.07 excellent grade when the percentage of practical instruction was increased from 5% to 20%, but the mean score decreased to 14.99 good grade when the percentage was increased to 25%. The variance data showed that the largest variance of 1.27 was found at 20% practical share, indicating a wide variation in student ability. And the variance dropped to 0.52 at 25%, suggesting that excessive practice may inhibit the systematic absorption of theoretical knowledge. The t-test t=2.861, P=0.011<0.05 supports the significant effect of the proportion of practice on competence, but it is necessary to balance the proportion of theory and practice to avoid excessive bias towards practical work leading to fragmentation of knowledge.

IV. C. Analysis of Differences in Intensity of Integration of Technology Tools on Students' Adaptability to Employment

Optimization of practical teaching needs to be combined with deep integration of technology tools. This section proposes the best practice path for the integration of AI algorithms and teaching scenarios by quantifying the correlation between the intensity of technology tool use and students' ability. The analysis of the difference between the intensity of technology tool integration on students' employment adaptation ability is shown in Figure 8.

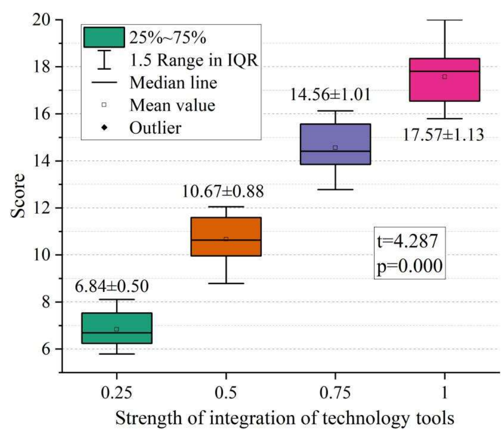

Figure 8: Difference analysis of integration strength of technical tools

When the intensity of technology tool integration was enhanced from 0.25 to 1, students' mean score jumped from 6.84 to 17.57 excellent grade, and the variance increased from 0.50 to 1.13, reflecting a wider distribution of students' abilities under high-intensity integration but a significant improvement in overall performance. t-test results t=4.287, P=0.000<0.01 indicate that technology tool integration is a key driver of employment adaptive ability The in-depth application of AI tools can enhance students' ability to deal with complex tasks, but attention should be paid to the match between tool use and teaching objectives.

V. Conclusion

This study systematically verifies the optimization effect of intelligent curriculum design on students' employment adaptability by constructing an efficient collaborative vocational education model based on AI algorithms.The LSTM-BERT-Attention model significantly outperforms the traditional model with an F1 value of 85.08% in the named entity recognition task on the DCNE dataset with a hyper-parameter optimized combination of the learning rate of 2×10^{-5} with batch_size=16, which improves the model convergence efficiency by 32%. In the multi-label course concept extraction task, the F1 value of the model reaches 83.58% on the "fourteen types of concepts" dataset, which is 4.86% and 9.50% higher than that of the Flat model (78.72%) and the LR-CNN (74.08%), respectively, which verifies the validity of the fusion of relational attention mechanism and word information.

For employment adaptability, the statistical test shows that for every 25% increase in the degree of course modularization, the average increase in students' competence is 3.21 points, $t=3.298$, $p=0.003$, and the degree of modularization is significantly and positively correlated with the concentration of competence distribution. When the percentage of practical teaching increased to 20%, the students' competence reached the excellent level, but the efficiency of theory absorption decreased by 12.7% after the percentage exceeded 20%, confirming the necessity of balancing practice and theory. When the intensity of technology tool integration was enhanced to 1, students' ability to handle complex tasks increased to 2.57 times the baseline value, $t=4.287$, $p<0.001$.

In summary, the AI-driven collaborative education model significantly improves the job suitability of vocational education through knowledge graph structuring, dynamic relationship extraction and multi-dimensional curriculum design optimization.

References

[1] Wheelahan, L., & Moodie, G. (2025). What do vocational colleges do and why do they matter? Thinking about the role of colleges as local actors. Journal of Vocational Education & Training, 77(1), 36-57.

[2] Théric, V., Cheneval-Armand, H., Delserieys, A., & Brandt-Pomares, P. (2025). Integrating academic and vocational subjects in vocational education and training: a case study in France. Journal of Vocational Education & Training, 1-26.

[3] Chen, Z. (2025). Implementation Pathways for Vocational Education in Supporting the Cultivation of High-Quality Farmers. Journal of Theory and Practice in Education and Innovation, 2(1), 15-21.

[4] Ludwig-Mayerhofer, W., Pollak, R., Solga, H., Menze, L., Leuze, K., Edelstein, R., ... & Kühn, S. (2019). Vocational education and training and transitions into the labor market. Education as a lifelong process: The German National Educational Panel Study (NEPS), 277-295.

[5] Napierala, J., & Kvetan, V. (2023). Changing job skills in a changing world. In Handbook of computational social science for policy (pp. 243-259). Cham: Springer International Publishing.

[6] Caves, K. M., Baumann, S., & Renold, U. (2021). Getting there from here: A literature review on vocational education and training reform implementation. Journal of Vocational Education & Training, 73(1), 95-126.

[7] Ling, Y., Chung, S. J., & Wang, L. (2023). Research on the reform of management system of higher vocational education in China based on personality standard. Current Psychology, 42(2), 1225-1237.

[8] Clayton, B., & Harris, R. (2018). Recent reforms in vocational education and training. International Journal of Training Research, 16(2), 99-102.

[9] Ismailova, Z. K., Khimmataliev, D. O., Khashimova, M. K., Baybaeva, M. K., & Ergashev, B. B. (2020). Integrative approach to designing the content of secondary specialized vocational education. Opción: Revista de Ciencias Humanas y Sociales, (91), 25-41.

[10] Fedulova, K. A., Fedulova, M. A., Kirillova, Y. S., Vagina, A. I., & Kuznetsov, T. M. (2017). Special competence in the structure of vocational pedagogical integrity in the sphere of vocational education. Eurasian Journal of Analytical Chemistry, 12(7b), 1265-1273.

[11] Gorbunova, N. V., Vaganova, O. I., Zhivoglyad, M. V., Snumitsyna, J. M., & Sergeeva, M. G. (2020). Design of Course Content in the Digital Environment of Vocational Education. International journal of applied exercise physiology, 9(11), 14-20.

[12] Bouw, E., Zitter, I., & De Bruijn, E. (2021). Multilevel design considerations for vocational curricula at the boundary of school and work. Journal of Curriculum Studies, 53(6), 765-783.

[13] Okumu, G. J., & Kenei, J. K. (2024). Adoption of Artificial Intelligence in Technical and Vocational Education and Training Institutions in Kenya. THE KENYA JOURNAL OF TECHNICAL AND VOCATIONAL EDUCATION AND TRAINING VOL: 7, 2.

[14] Zhang, H. (2025). The Research on the Construction Path of Artificial Intelligence Curriculum System for Higher Vocational Undergraduate Education in the New Era. International Educational Research, 8(2), p99-p99.

[15] Bellas, F., Guerreiro-Santalla, S., Naya, M., & Duro, R. J. (2023). AI curriculum for European high schools: An embedded intelligence approach. International Journal of Artificial Intelligence in Education, 33(2), 399-426.

[16] Fransisca, M., & Saputri, R. P. (2025). Needs Assessment of Artificial Intelligence in Education at Vocational High School. The Indonesian Journal of Computer Science, 14(1).

[17] Qin, C. (2024). Discussion on the Curriculum Design and Implementation of Competency-Based Vocational Education. Journal of Modern Educational Theory and Practice, 1(2).

[18] Deckker, D., & Sumanasekara, S. (2025). AI in vocational and technical education: Revolutionizing skill-based learning. EPRA International Journal of Multidisciplinary Research (IJMR), 11 (3), 9. https://doi. org/10.36713/epra20462.

[19] Duan, J., & Wu, S. (2024). Beyond traditional pathways: Leveraging generative AI for dynamic career planning in vocational education. International Journal of New Developments in Education, 6(2), 24-31.

[20] Widayanthi, D. G. C. (2024). AI in Vocational Education: Balancing Digital Skills and Humanism. Indonesian Journal of Educational Inquiry, 1(2), 1-10.

[21] Luo, Z., Jiang, J., & Zhou, H. (2024). Exploration on the application of artificial intelligence technology in vocational education teaching. J. High. Vocat. Educ, 1(2), 177-182.