ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 1 - 2026

Analysis of Intelligent Development Path of Civic and Political Education Theory System in the New Era Based on Deep Learning

Jing Zhu[1,*]

[1] College of Urban Construction, Wuchang Institute of Technology, Wuhan, Hubei, 430074, China

SUMMARY: *The contemporary theoretical framework serves as a core component of political ideology and civic education at the university level institutions. In order to promote the smart advancement of the current theoretical system, this paper constructs BiLSTM-Attention-CRF, a course entity identification model, and BERT-BiLSTM-Attention, an inter-entity relationship extraction model, developed through deep learning, so as to construct the knowledge graph of Civics and Politics courses. Then the course recommendation algorithm KGMO-RS integrating a knowledge graph with multi-objective optimization is proposed, and the type of student satisfaction needs and priority ranking in the Civics classroom are explored by the KANO model. The study indicates that the BiLSTM+Attention-CRF model and the BERT-BiLSTM-Attention model presented in this work surpass competing approaches in the entity identification and inter-entity relation extraction tasks, thereby suggesting that the proposed model is able to construct a higher quality knowledge graph. Meanwhile, the KGMO-RS algorithm outperforms the GCNKG-CR algorithm in terms of HV, IGD and other metrics, which proves that the algorithm demonstrates superior comprehensive search ability and stable convergence behavior throughout the multi-objective optimization process of course recommendation. In addition, drawing on the findings concerning student satisfaction in the Civics classroom, this paper puts forward suggestions to enhance students' classroom satisfaction in the Civics course across colleges and universities, which provides an approach that can support the intelligent development of the theoretical system in the new era.*

KEYWORDS: *curriculum knowledge graph; deep learning; named entity recognition; relationship extraction; KANO model; new era theory system*

1 Introduction

In 2019, the first AI and Education World Conference, China clarified the route of applying artificial intelligence across the field of education, put forward new ideas toward advancing intelligent education, and made the progress of artificial intelligence an intrinsic support to help the construction of the education system. In 2022, the convening at the fourth Global Conference on Artificial Intelligence and Education further elucidated the construction of a new ecosystem of education development relying on the contents of wisdom sharing and consensus-building. The inherent requirements. Among other things, the conference clarified the construction of the wisdom platform, teacher education services and digital development to lead education reform [1]. Deep learning as an important branch of artificial intelligence, through the establishment, simulation of the neural architecture mimicking the human brain's data

*wyq20120502@126.com
https://doi.org/10.65102/is2026351 Published: 30 April, 2026

processing mechanism to capture externally input raw information via layered feature abstraction, enabling machines to comprehend training data and retrieve relevant information, and has seen extensive application in the field of education, such as data mining, students' behavior identification, educational resources recommendation, dynamic teaching adjustment, precision education, etc.[2-6].

Over the past ten years of curriculum innovation in ideological and civic-political instruction, has undergone qualitative changes in both teaching concepts and teaching methods, but learners' capacity for creative thought, analytical reasoning, and tackling real-world challenges remains insufficient [7-9]. And in August 2022, a number of government departments promoted the implementation of the "Work Program for Systematically Advancing the Development of 'Flagship Civics and Political Theory Courses'", which clearly states that "the enhancement of learning materials for civic and ideological instruction on the national wisdom education platform will be strengthened; the normalization and institutionalization of teaching resources construction will be promoted through project support; and the organization will develop and recommend a number of scientific, authoritative and practical courseware and handouts, and promote the uniform use of them by front-line teachers."

So far, against the backdrop of rapid progress in artificial intelligence, the deep learning technology, resources and platform construction, into the system of reform and advancement of civic and political education, to strengthen the forward-looking vision of civic-political instructional practice, to realize civic and political education tasks to carry out the work of the times, will be a pivotal driver in broadening the scope of civic and political education's practical deployment during the current era [10-13]. Institutions need, through the integration of deep learning educational resources, educational system construction and other measures, to fully harness the nurturing potential of deep learning within civic and political education, grounded in the progress of intelligent education that empowers civic and political education domains, to establish an innovative paradigm for civic and political education encompassing classroom instruction, moral cultivation, and hands-on practice, and to lay a solid foundation for the diversified and well-structured progression of civic and political education [14-16].

Focusing on the advancement of the new era theoretical system in the context of Civics and Ideological Education across higher education institutions, the course recommendation algorithm KGMO-RS based on knowledge graph is introduced. Throughout the knowledge graph development process, the present study adopts the deep learning method, and the detailed methodology addresses two dimensions—named entity recognition and inter-entity relationship extraction— and respectively constructs the BiLSTM-Attention-CRF model and the BERT-BiLSTM-Attention model to realize the entity-relationship extraction of the Civic and Political Education curriculum. Furthermore, the KGMO-RS algorithm is improved on the basis of Graph Convolutional Recommendation Algorithm based on Knowledge Graph (GCNKG-CR), and introduces the NSGA-II multi-criteria optimization approach built upon an enhanced variational operator. Finally, using KANO model, this paper explores the influencing elements and priority of student satisfaction in the Civics classroom, and puts forward relevant suggestions based on this.

2 Deep learning based knowledge map construction method for civic education courses

In order to advance the intelligent development of the theoretical system in the new era, this paper conducts investigations from the perspective of civic and ideological instruction at tertiary education institutions, and designs the intelligent recommendation approach grounded in the knowledge graph of ideological and political education courses on the basis of deep learning. This chapter gives a specific description of the method of constructing a knowledge graph for the Civics and Politics Education courses in it.

2.1 Deep learning based approach for course named entity recognition

In this paper, BiLSTM-Attention-CRF model is used to accomplish the task of named entity identification for the core concept entities within the Civics and Ideological Theory course. The architectural overview of the BiLSTM-Attention-CRF model is depicted in Fig. 1. The model primarily consists of four components: Char Embedding layer, BiLSTM layer, Attention layer and CRF layer. Firstly, the text is embedded according to the character level for vectorized representation, the vectorized data is fed into the dual-directional LSTM neural network layer, the text sequence is modeled, the context dependency is captured, the semantic features are acquired, the acquired features are input into the Attention layer, the relevant feature vectors are dynamically output according to the Attention mechanism, and finally, the feature vectors are converted into sequence labels by the CRF decoder, to get the course knowledge entity.

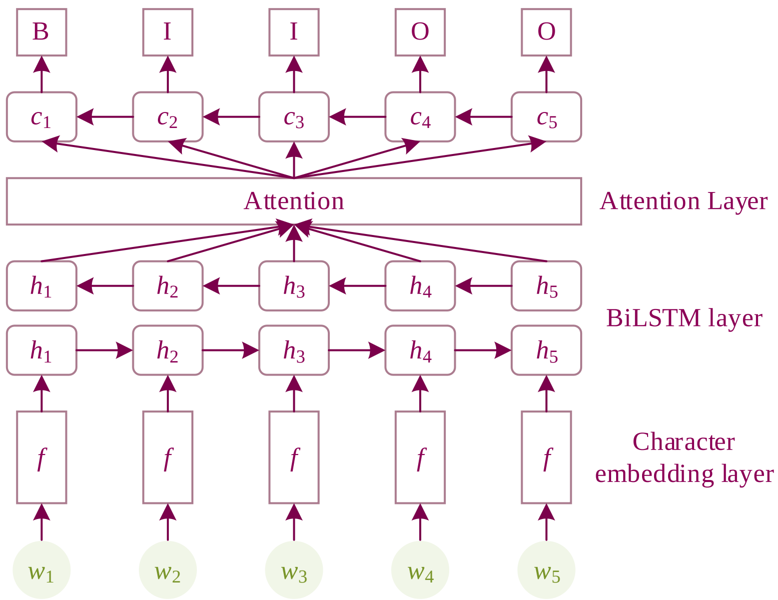

Figure 1: BiLSTM-Attention-CRF model

2.1.1 BiLSTM layer

A Long Short-Term Memory (LSTM) network is a specific type of Recurrent Neural Network (RNN). In contrast to RNN architectures, LSTM networks incorporate additional components including memory units, input gates, forgetting gates and output gates, which successfully address the gradient vanishing issue.

In this paper, BiLSTM is used in this layer of neural network, where the sentences in the dataset are computed in a bidirectional (forward and backward) manner, and then the vectors obtained from both of them are concatenated to derive the hidden-layer vector representation of the BiLSTM network. The specific formulas involved are presented below:

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

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

$$\tilde{c}_t = \tanh\left(W_c \cdot \left[h_{t-1}, x_t\right] + b_c\right) \tag{3}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \tag{4}$$

$$o_t = \sigma\left(W_o \cdot \left[h_{t-1}, x_t\right] + b_o\right) \tag{5}$$

$$h_t = o_t \odot \tanh\left(c_t\right) \tag{6}$$

where f_t denotes the forgetting gate, i_t denotes the input gate, o_t denotes the output gate, and x_t denotes the input at moment t. b denotes the associated offset vector and w denotes the associated weight matrix. c_t denotes the state at moment t, c_t denotes the new cell state updated by the old state c_{t-1}, h_t denotes the last output at moment t, and h_{t-1} denotes the output at moment $t-1$. σ denotes the sigmoid activation function and tanh denotes the hyperbolic tangent function.

2.1.2 Attention layer

The underlying idea of the attention mechanism involves allocating constrained computational resources toward salient information, and the architecture employed in the current study adds the attention mechanism after BiLSTM. Through the attention mechanism can focus on the features related to the current output, inhibit the features not related to the current output, can better capture the context of the information dependence to obtain the corresponding semantic features. The computational expression of the architecture adopted in the present work at the Attention layer is shown below:

$$V_t = \tanh\left(h_t\right) \tag{7}$$

where V_t denotes the attention weight, which corresponds to the degree of meaning-level relevance between the target entity and its neighboring contextual information. tanh represents hyperbolic tangent function h_t is the context feature vector obtained from the two-way long- and short-term memory network.

For the obtained attention weights are not directly used, but also need to use softmax to calculate the function weight probabilistic, and then the attention weight configuration:

$$P_t = \frac{\exp\left(V_t\right)}{\sum_{t=1}^{m} \exp\left(V_t\right)} \tag{8}$$

Varying semantic weight configurations are assigned based on different levels of semantic relevance, thus achieving the aim of amplifying more relevant semantic features:

$$a_t = \sum_{t=1}^{m} P_t h_t \tag{9}$$

where h_t is the context feature vector obtained by BiLSTM and P_t denotes the corresponding probability after weight probabilization.

2.1.3 CRF layer

CRF is a Markov Random Field based classifier that models the input data sequence so that every label within the sequence can be classified. For the named entity recognition task in this paper, there are strong dependencies between neighboring labels in the label sequence. The intermediate label I must be behind the start label B. Based on such dependencies, the model is realized to construct the labeling information for the context. Given the label sequence $y = \{y1, y2, y3, \dots yt\}$, the score function of CRF is as follows:

$$S(X, y) = \sum_{i=1}^{t} P_{i,y_i} + \sum_{i=0}^{t} A_{y_i, y_{i-1}} \tag{10}$$

where matrix P is derived from the output of Bi-LSTM, matrix A is the state transfer matrix between labels, and $A_{i,j}$ denotes the probability score of the transfer from label i to label j.

After computing the score function $S(X, y)$, the posterior probability of the correct label sequence y given the input X is obtained by softmax:

$$P(y \mid X) = \frac{e^{S(X,y)}}{\sum_{\hat{y} \in Y_X} e^{S(X,y)}} \tag{11}$$

where Y_X denotes all possible label sequence combinations when the input is X.

The maximum likelihood function is:

$$\log(P(y \mid X)) = S(X, y) - \log\left(\sum_{\tilde{y} \in Y_X} e^{S(X,y)} \right) \tag{12}$$

$$y^* = \underset{\tilde{y} \in Y_X}{\operatorname{argmax}}\, S(X, \tilde{y}) \tag{13}$$

Finally the target label can be computed using the Viterbi algorithm.

The CRF layer expresses the labeling constraints through the state transition matrix of the sequences, the objective of the model is to maximize the proportion of correct state transfer sequences in the matrix, here the loss function is characterized as the negative log-likelihood of the ratio of the fraction of the current state transfer matrix in the correct path to the fraction of all the paths, and the formula is shown below:

$$L(P, S) = -\log \frac{P_{RealPath}}{P_1 + P_2 + \cdots + P_N} \tag{14}$$

$$P_{total} = P_1 + P_2 + \cdots + P_N = e^{S_1} + e^{S_2} + \cdots + e^{S_N} \tag{15}$$

$$S_i = EmissionScore + TransitionScore \tag{16}$$

In Eq. (14) P_i denotes the score of the i th feasible path, there are N paths, and $P_{RealPath}$ denotes the score of the real path. In turn, according to equation (15), P_i can be expressed as e^{S_i}, S_i obtained by summing the two parts of the score, where the firing score is obtained from the BiLSTM layer and the transfer score type is a part of the transfer matrix, which is continuously updated during the training to obtain the transfer scores of the paths. The computation of these two scores is related to the number of text labels of the specific transfer matrix.

2.2 Deep Learning Based Course Relationship Extraction Approach

In the present study, the BERT-BiLSTM-Attention architecture is utilized to accomplish the inter-entity relation extraction task within Ideological and Political Theory courses. The structural overview of the BERT-BiLSTM-Attention model is illustrated in Figure 2. After the text information passes through the input layer, it relies on the feature extraction ability of BERT to get the text word vector matrix. The feature vectors are input to the BiLSTM layer to process the text sequence and capture its context-dependent information to obtain semantic features. Then the features with contextual information are fed into the Attention layer to consolidate the utterance features through the attention mechanism. Finally, the sentence level feature information is passed into softmax, so as to achieve the effect of entity relationship classification.

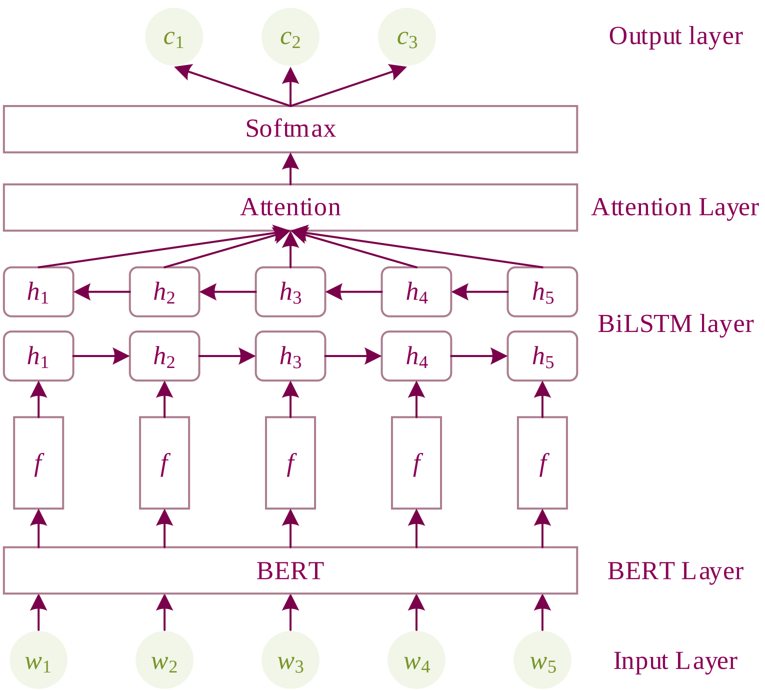

Figure 2: BERT-BiLSTM-Attention model

2.2.1 BERT layer

The BERT layer encodes the input text into a vector representation, and compared to Word2Vec, BERT acquires more diverse linguistic knowledge throughout the pre-training phase. Phrase word-level features can be learned in the low-level network structure, linguistic features in the mid-level network structure, and semantic features in the high-level network structure. It adopts masked language model to be able to obtain deep bi-directional linguistic representation information. In addition, it adopts Transformer as a feature extractor, which can fully integrate contextual information and assist entity-relationship pairs to obtain more information.

2.2.2 BiLSTM layer

The BiLSTM layer serves the purpose of capturing context-aware relational patterns and generating the associated semantic features. The BiLSTM layer here is similar to the previous section on entity name identification.

2.2.3 Attention layer

The Attention layer primarily functions to derive the weight matrix in this model, and through its attention mechanism, it is able to integrate the word-level feature vectors into utterance-level feature vectors to assist in entity relationship classification. The relevant formulas are listed below:

$$M = \tanh(H) \tag{17}$$

$$\alpha = \mathrm{softmax}\left(W^t M\right) \tag{18}$$

$$r = H\alpha^T \tag{19}$$

$$h^* = \tanh(\alpha) \tag{20}$$

In the above equation, H denotes the output matrix of BiLSTM and W^t denotes the initialized weight matrix. Eq. (20) represents the final relationship classification obtained through the attention layer.

2.3 Knowledge Integration in Civic Education Programs

A substantial amount of triplet-based knowledge can be acquired via the detection and relational extraction of entities. Nevertheless, owing to the constraints of current extraction techniques, discrepancies in data origins among triplet groups and other contributing factors, redundancy and logical inconsistencies are prone to emerge across triplet groups. Consequently, in order to construct a more precise and comprehensive inter-triplet knowledge repository, it is essential to carry out knowledge merging operations between them. To address the issue of a single term carrying multiple interpretations, the conceptual meaning of the reference ought to be represented in the form of triples. To resolve the situation where multiple distinct expressions share one identical concept, it is necessary to establish linkages among the triples so as to consolidate them into a coherent and unified structure. The prevailing knowledge merging techniques encompass both supervised and unsupervised machine learning paradigms as well as deep learning-based approaches.

The present section adopts a cascaded two-stage screening framework for knowledge consolidation. Throughout the merging procedure, every entity is required to pass through two successive filtering steps. An upper boundary threshold σ_1 and a lower boundary threshold σ_2 are designated for the respective filtering rounds. At each processing stage, pairwise similarity evaluations are conducted among entities. Whenever the computed similarity between any two entities surpasses σ_1, the pair is regarded as referring to the same entity and subsequently merged. Conversely, when the similarity measurement between the two drops below σ_2, it indicates that the two entities are distinct and semantically divergent, thus necessitating additional judgment. For cases falling in the interval between σ_1 and σ_2, or when attribute information of the object is absent at the time of judgment, the process advances to the subsequent filtering step. The initial filtering stage exploits string-level properties derived from entity names. On this basis, distributed vector representations of ideological and political entity names were generated via the Word2Vec model, after which pairwise cosine similarity computations were conducted on these vector representations. The second filtering stage relies on edit distance as its evaluation criterion. The specific process of knowledge fusion is illustrated in Figure 3.

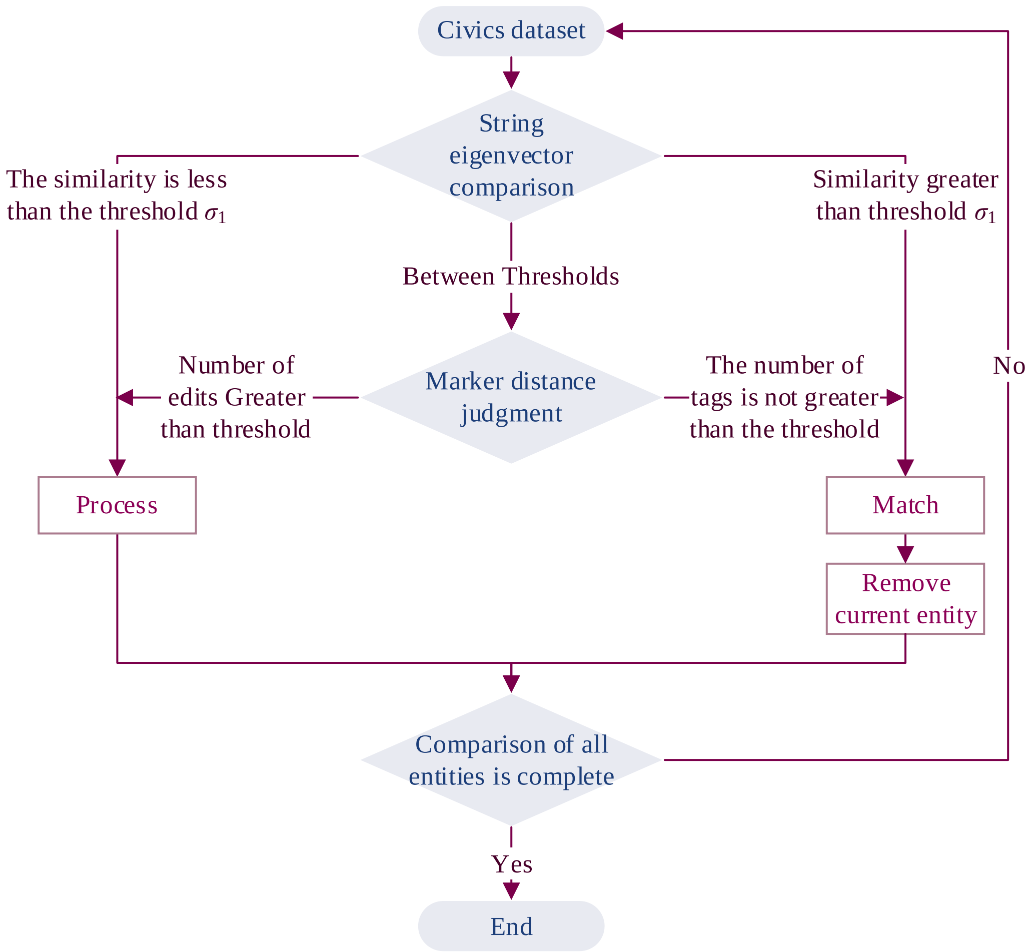

Figure 3: Knowledge fusion process

2.4 Civic Education Program Knowledge Store

The preservation of knowledge graphs entails retaining the knowledge triplets derived after the processes of entity recognition, relationship extraction and knowledge fusion in a computer for querying and use. The knowledge graph storage mechanism is required to fulfill the following requirements:

(1) Scalability: with the continuous growth of the knowledge base, an expandable storage mechanism is needed to handle increasing volumes of data and retrieval demands.

(2) Flexibility: The storage mechanism must be sufficiently adaptable to fulfill the demands of diverse queries and applications.

(3) Efficient: The storage mechanism must be adequately capable to handle real-time

queries and applications.

(4) Reliability: The storage mechanism must be dependable enough so as to guarantee data integrity and security.

Usually, there are two storage forms for knowledge graphs: Neo4j graph database and RDF files. For ease of management and use, this paper selects the Neo4j graph database. Neo4j is a graph database management system specialized in storing and processing graph-structured data. Unlike traditional relational databases, it uses nodes and relations to build graph structures; nodes represent entities, such as people or books, and can contain attributes, such as name or age, while relations represent connections between different entities. Neo4j provides a graph-based query language, Cypher, as well as a variety of APIs and tools, such as the JavaAPI, the Python API, and the Cypher Shell, etc. to facilitate developers and data analysts to perform data query, import and export operations. It is well-suited for storing and retrieving data with intricate structures, such as social networks and knowledge graphs, etc. Neo4j's advantages lie in its capacity to handle densely interconnected and elaborately organized data, its user-friendliness and operational convenience, and its seamless integration with Web frameworks, making it extensively adopted for data administration and analytical tasks across numerous disciplines.

2.5 Comparative analysis of experiments and results

2.5.1 Named Entity Recognition Experiment

(1) Experimental environment

The relevant equipment information during the experiments in this paper is shown in Table 1, which lists the hardware devices such as CPU, hard disk and memory, as well as the software devices such as operating system, compiler and open source libraries imported in the program. Both the named entity recognition and relationship recognition tasks in this paper are implemented in the experimental environment of Python 3.13, and the algorithmic models are implemented by Tensorflow and Keras, and the CRF layer used in the models is provided by Keras-contrib, an extension package of Keras.

Table 1: Information on experimental equipment

Hardware		Software	
CPU	Intel core i7 14th Gen	System	Win11
Hard disk	512GB SSD	IDE	Pycharm
Memory	16GB	Open source library	Tensorflow, Keras

(2) Evaluation metrics

In this paper, precision rate (P), recall rate (R) and F1 value are used as the evaluation criteria of the experimental results. The formulas for precision rate and recall rate are as follows:

$$P = \frac{\text{Number of entities correctly identified}}{\text{Number of entities identified}} \times 100\% \tag{21}$$

$$R = \frac{\text{Number of entities correctly identified}}{\text{Total number of entities}} \times 100\% \tag{22}$$

However, in practice, instead of using a single precision or recall rate as an evaluation criterion, the weighted average of the two, F_β, is used as an evaluation metric, which has the following expression:

$$F_{\beta} = \left(1+\beta^2\right) \times \frac{\Pr ecision \times \operatorname{Re} call}{\left(\beta^2 \times \Pr ecision\right) + \operatorname{Re} call} \tag{23}$$

In this paper, we take the F1 value at $\beta = 1$ as the evaluation criterion, and the formula is as follows:

$$F1 = 2 \times \frac{\Pr icision \times \operatorname{Re} call}{\Pr icision + \operatorname{Re} call} \tag{24}$$

(3) Parameter comparison experiment and analysis

In the model training process, there exist two critical parameters: learning rate and dropout value. When the learning rate is excessively high, it may cause the model to converge prematurely and potentially exceed the optimal value, whereas when the learning rate is insufficiently low, it will result in the model converging at an unacceptably slow pace, compromising the training effectiveness, and may ultimately lead to the model being unable to converge. Therefore, in this paper, we incorporate the comparison experiments of learning rate and dropout value to assess the model with optimal results.

In this paper, firstly, by constantly tuning the learning rate, the recognition capability of the model is compared when the learning rate is 0.001 and 0.0001 respectively as shown in Fig. 4, and it is found that the F1 value when the learning rate is 0.001 is higher than that when the learning rate is 0.0001, and the F1 score at the 15th round reaches 0.9354, which is better than that of the model recognition effect, therefore, the learning rate adopted in this paper is 0.001.

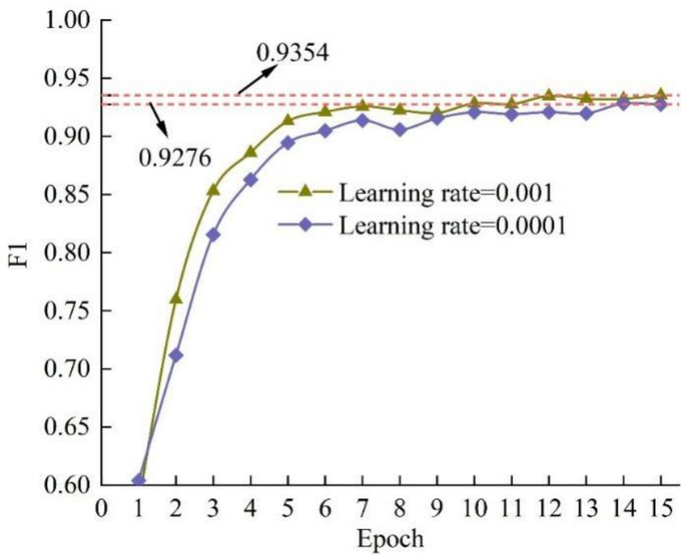

Figure 4: Model performance at learning rates of 0.0001 and 0.001

Meanwhile, the dropout strategy is capable of being employed to alleviate the issue of overfitting during the model optimization process. The evaluation of the model's behavior under the condition that the learning rate is set to 0.01 and the dropout is configured to 0.3, 0.4 and 0.5 respectively is presented in Figure 5. It can be seen that the difference between the two is not large, due to the dropout take 0.4 when the F1 value is larger and the performance is smoother, so this paper finally adopts the dropout setting of 0.4 case.

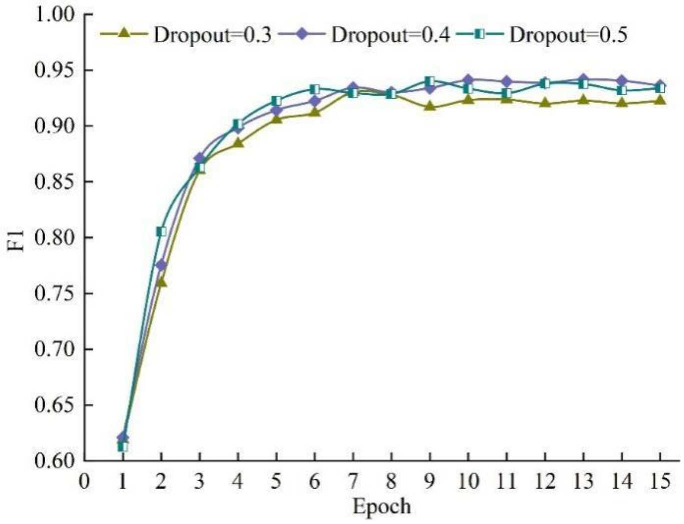

Figure 5: The model effect when the dropout value is 0.3, 0.4 and 0.5

How Loss fluctuates as the iteration count increases is shown in Fig. 6. It can be seen that the Loss value progressively declines as the Epoch iteration rounds increase, and after 11 rounds of iteration, the loss stabilizes at a lower value and becomes more stable..

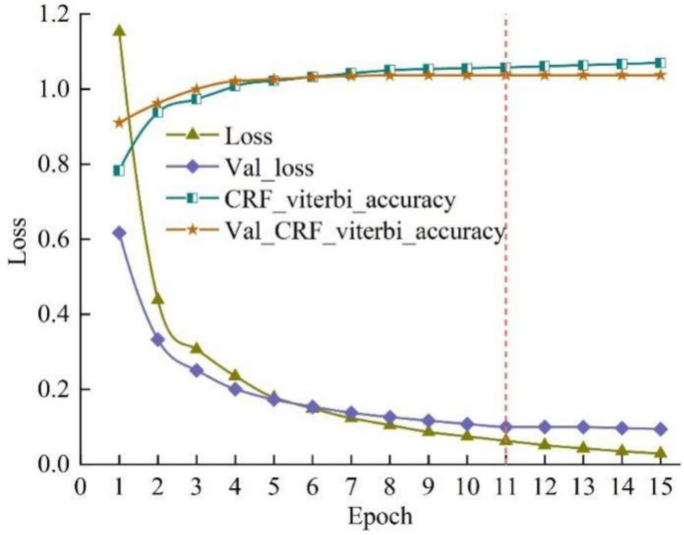

Figure 6: The effect of Loss varying with the number of iterations

Through the above parameter comparison test, the architecture with a learning rate of 0.001 and a dropout value of 0.4 was ultimately determined as the optimal configuration, and the model effect is shown in Fig. 7, at which time the model is the most effective.

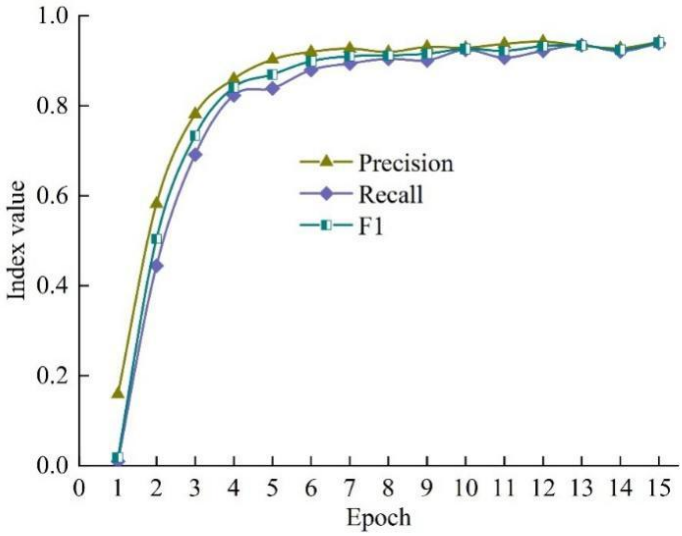

Figure 7: The model performance when the learning rate is 0.001 and the dropout value is 0.4

(4) Comparison and analysis with other models

This paper compares three groups of methods for entity recognition, which are BiLSTM+ Attention-CRF model, BiLSTM-CRF model, and CNN+Attention-CRF model throughout this investigation, and each model is trained for 20 rounds each, and an evaluation of the experimental outcomes of the entity identification models is presented in Table 2.

It can be seen that each model can achieve certain entity recognition results. Among them, the precision and recall of BiLSTM+Attention-CRF model are improved by 7.58% and 6.56% respectively compared with BiLSTM-CRF model, demonstrating that the BiLSTM+Attention-CRF architecture presented here surpasses BiLSTM-CRF model. This indicates that incorporating the attention mechanism can effectively exploit the local and global information of the text to obtain more complete text features, which makes the model more accurate and can achieve better named entity recognition results. Upon substituting CNN with BiLSTM, the F1 value is elevated by 5.23% relative to CNN+Attention-CRF model, confirming that BiLSTM can contribute to further strengthening the overall performance of the architecture.

In summary, the BiLSTM+Attention-CRF model proposed in this paper can achieve better performance in the work of entity name identification, and the incorporation of the attention mechanism and BiLSTM module can considerably boost the entity detection capability to a notable degree. Moreover, the present study applies the model to accomplish the entity forecasting of unlabeled text data to guarantee the entity data integrity of the knowledge graph of the Civic Education curriculum.

Table 2: Comparison of model experiment results

Task	Algorithm model	Precision /%	Recall /%	F1 /%
Entity recognition	BiLSTM-CRF	85.31	87.18	86.24
	CNN+Attention-CRF	85.75	90.54	88.08
	BiLSTM+Attention-CRF	92.89	93.74	93.31

2.5.2 Relationship Extraction Experiments

The dataset uses the output data from the entity identification task within the previous section, which is well labeled with relationships according to the data preprocessing method. The dataset contains 3479 knowledge entities and 17035 pairs of entity relationships. The labeled dataset is divided into training subset and test subset at a proportion of 7:3.

In this paper, the relation categorization procedure is carried out by means of BERT-BiLSTM-Attention based on Python 3.13 and Pytorch deep learning framework experimental environment. The BiLSTM output size of this network is set to 32 and since it is bidirectional, the actual output is 64. Among other parameters learning rate is configured to 0.001, dropout probability is assigned to 0.4, batch_size is 128 and epochs are 50.

The current study benchmarks the BERT-BiLSTM-Attention architecture against CNN model, PCNN model, BiLSTM model and Attention+BiLSTM model for experiments. In this section, Accuracy ACC, Precision P, Recall R, and F1 value are utilized as assessment indicators for relation classification. The outcomes of the inter-entity relation extraction experiments are displayed in Table 3.

It can be seen that for the classification of relationships between knowledge points, all five models can be effective. Among them, the BERT-BiLSTM-Attention architecture achieves the best, and the F1 value can reach more than 87%, and the model effect has been greatly improved, which indicates that BERT and attention play a significant role. In the present work, the BERT-BiLSTM-Attention architecture is eventually adopted for the relation categorization objective, which can also ensure the quality of the relationships in the later development of the knowledge graph.

Table 3: Experimental results of relation extraction

Task	Algorithm model	ACC /%	Precision /%	Recall /%	F1 /%
Classification of relationships between entities	BiLSTM	67.03	69.65	83.85	76.09
	CNN	73.01	78.43	82.04	80.19
	PCNN	75.94	78.62	83.72	81.09
	Attention+BiLSTM	77.58	78.79	87.13	82.75
	BERT-BiLSTM-Attention	82.69	83.57	91.24	87.24

In order to understand more clearly the reason why Attention enhances the performance of the model, this paper uses Python's Matplotlib library to carry out a simple visualization of Attention to explore the reason. The result of Attention visualization is shown in Fig. 8. The larger the value of the color band from left to right scale in the figure, the larger the weight value of the word, i.e., Attention, indicates that the word contributes to the classification to a greater extent.

By analyzing the example shown in the figure, it can be seen that in this sentence, the words "medial expression", "conversion", and "binary tree" make the main contribution to the final classification. On the other hand, "PAD" is a word that is not long enough to be filled in by zero padding, and does not contribute to the final classification. This also shows that the use of Attention can focus the attention on the important words, which can substantially enhance the classification effect of the model.

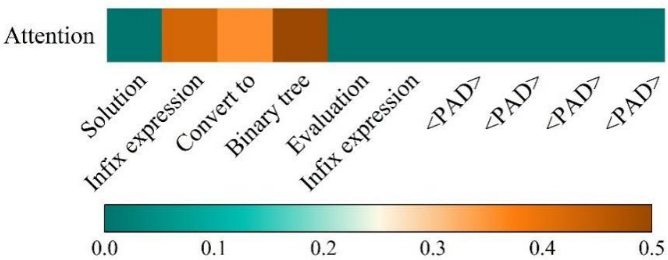

Figure 8: Visualization of Attention

On the other hand, the reasons why the BERT-BiLSTM-Attention model outperforms several other models are analyzed. For the CNN model, it is also able to extract features, but it has a parameter to set the window size, which indicates that the model can only extract features within a limited window range. For the BERT-BiLSTM-Attention model, firstly, the text word vector matrix is obtained based on BERT, and then it is inputted to BiLSTM to obtain semantic features, and then the output of BiLSTM is weighted and summed with attention, so that the words with a greater degree of importance can be found and more accurate semantic features can be extracted. And the model extracts global features, which makes the information more comprehensive, and thus the model effect is better than the other models.PCNN model uses convolutional kernel with multiple width sizes for feature extraction compared to CNN model, so the PCNN model effect also has a small improvement on the basis of CNN model.

3 Course recommendation algorithm based on knowledge graph and multi-objective optimization

On the foundation of building the knowledge graph of Civic Education courses, this chapter proposes a Civic Education course recommendation algorithm utilizing an improved NSGA-II multi-objective optimization, and verifies the efficacy of the algorithm through experiments.

3.1 Recommendation Algorithm Framework

In this paper, based on the course recommendation candidate list generated by the knowledge graph-based graph convolutional recommendation algorithm (GCNKG-CR), we propose a recommendation algorithm utilizing multi-objective optimization (KGMO-RS), and the framework of the KGMO-RS algorithm is shown in Figure 9.

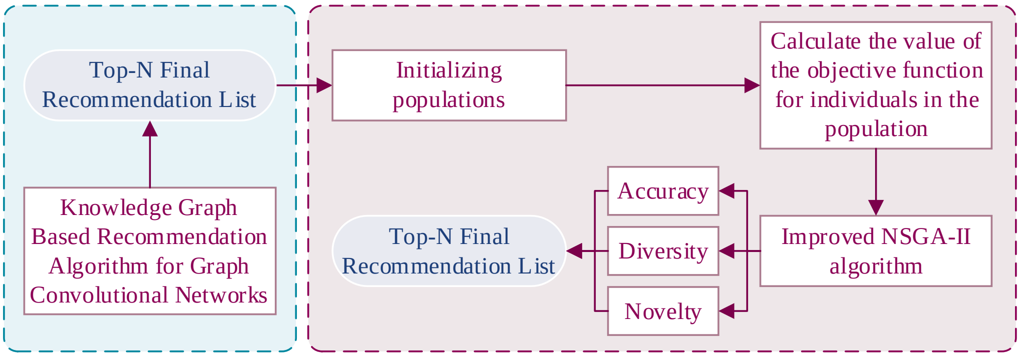

Figure 9: Framework of recommendation algorithm based on multi-objective optimization

3.2 NSGA-II Algorithm

Non-dominated Sorting Genetic Algorithm (NSGA) represents an intelligent algorithm constructed on genetic algorithm and grounded on Pareto superiority relation for tackling many-objective optimization challenges.

NSGA-II algorithm is an upgraded version of NSGA and its implementation procedures are outlined below:

Step 1: Randomly set up the initial population P_0 and execute a rapid non-dominated sorting procedure to acquire the Pareto-ranking of every individual within the current population.

Step 2: Set the evolutionary generation Gen=1.

Step 3: Determine whether the first generation subpopulation has been produced, if it has been created then make Gen=2, otherwise use the binary tournament selection mechanism to select suitable individuals in the starting population and subsequently apply corresponding recombination and perturbation operations to yield the first generation subpopulation.

Step 4: Combine the parental and offspring populations to construct an updated population.

Step 5: Determining whether a fresh parent population is produced, and if not, calculating an objective function for the new population generated in step 4, using a rapid non-dominated sorting procedure and a density estimation mechanism to identify Pareto tiers and congestion degree metrics of each individual within the updated population, and retaining a designated quantity of optimal individuals as the new parent population based on an elite retention strategy.

Step 6: Perform selection, crossover, and mutation operations to yield an offspring population for the parent population generated in step 5.

Step 7: Determine whether Gen reaches the upper limit of iteration rounds, if it attains the preset ceiling of evolutionary generations then the algorithm terminates running, otherwise, the iteration counter Gen=Gen+1 and jump to step 4 to continue execution.

3.3 NSGA-II algorithm based on improved variational operators

In this paper, we introduce a two-stage improved variational operator to enhance the NSGA-II algorithm in order to strengthen the local search capability of the algorithm and boost the recommendation performance.

3.3.1 Problem coding

Common individual encoding methods are floating point/real number encoding, binary encoding, Gray encoding, etc., and there are only two cases of recommended or not recommended in the course recommendation list, which is suitable for encoding using binary encoding, so use binary encoding to encode the candidate recommendation list as $\{x_1, x_2, x_3, \cdots, x_K\}$, where K represents the total number of courses in the candidate recommendation list, and x_i is assigned a value of 0 or 1, that is, when the ith item is suggested to the user, x_i is encoded to be 1, and otherwise x_i is encoded to be 0. Since the final recommendation list is of length N, the individual encodings have to satisfy the following constraints:

$$N = \sum_{i=1}^{K} x_i \tag{25}$$

3.3.2 Objective function

In order to improve the platform resource utilization and user satisfaction, this paper takes the accuracy, diversity and novelty of the recommendation results as the objective function to optimize the candidate recommendation lists, and generates recommendation lists with better results in all three aspects.

(1) Accuracy

The accuracy rate represents the learner's satisfaction with the final recommendation result, and its expression is as follows:

$$f_{pre}(R) = \frac{|R \cap T|}{|R|} \tag{26}$$

where R denotes the final recommendation list, $|R|$ is the length of the recommendation list, and T denotes the set of courses that the learner has taken.

(2) Diversity

Diversity is used to assess the differences between all items in the recommendation list, reflecting the ability of the recommendation system to recommend different areas of content to the user. The expression is as follows:

(27)

where $sim(i, j)$ denotes the similarity between the i th item and the j th item in the recommendation list.

(3) Novelty

Novelty is the ability of recommender system to recommend courses to learners that are less popular but meet their interests. The formula is as follows:

$$f_{nov}(R) = \frac{\sum_{ieR} \log_2 \left(\frac{n}{\lambda_i} \right)}{|R|} \tag{28}$$

where λ_i denotes the number of times the ith item in the recommendation list has been selected. The larger value of $f_{nov}(R)$ indicates that the novelty of the recommendation list is better.

In summary, the objective function can be summarized as:

$$\begin{cases} \max f_{pre}(R) = \dfrac{|R \cap T|}{|R|} \\ \max f_{div}(R) = \dfrac{\sum_{i \in R} \sum_{j \in R, j \neq i} (1 - \mathrm{sim}(i, j))}{|R|(|R|-1)} \\ \max f_{nov}(R) = \dfrac{\sum_{i \in R} \log_2 \left(\frac{n}{\lambda_i} \right)}{|R|} \end{cases} \tag{29}$$

3.3.3 Intersection operators

The improved NSGA-II algorithm uses the uniform crossover operator to perform the crossover operation to generate new individuals as follows:

(1) Set the initial state of the C1 and C2 gene sequences of the offspring individuals, and set all gene values to 0.

(2) Iterate through the parent individuals P1 and P2, and if P1 and P2 both have a value of 1 at a gene position, set the gene values of the offspring individuals C1 and C2 at that position to 1 as well, and set the gene values of P1 and P2 at that position to 0. Record the number of gene positions that satisfy the condition as s.

(3) Set the crossover parameter $n(1 \leq n \leq |b - s| - 1)$, where b is the total number of individual gene positions.

(4) Find and record the first n genes in the parent individual P1 with a value of 1 and the last n genes in the parent individual P2 with a value of 1, and set the gene value at the

corresponding position in the offspring individual Cl to 1. Similarly, find and record the last n genes in the parent individual P1 with a value of 1 and the first n genes in the parent individual P2 with a value of 1 and set the gene value at the corresponding position in the offspring individual value at the corresponding gene position in C2 is set to 1.

3.3.4 Improvement of the variation operator

In this paper, we introduce a two-stage strategy to refine the mutation operator to strengthen the local search capability of NSGA-II algorithm.

The specific flow of the two-stage improvement variation operator to update the population is as follows:

(1) Initialize the parameters such as population size, crossover rate, mutation rate, etc., and configure the maximum number of evolutionary generations of the population to T.

(2) Obtain the current evolutionary generation t of the population, and when $t < 0.3T$, execute the uniform variation operator with strong global search capability. When $t \geq 0.3T$, execute the differential variation operator to prevent falling into local optimality.

(3) After the operation, generate the offspring population.

In order to make the mutated individuals satisfy the constraints of Eq. (25), the uniform mutation first records the positions with gene value 1 and gene value 0 in the gene sequence of the individual, and in the case of satisfying the mutation probability, firstly mutates the position with gene value 0, and takes a random number from [0,1], and if the difference between the generated random number and the gene value exceeds the preset threshold, then performs the mutation operation, and records the the number of mutated gene loci. Then the same number of bits are randomly selected from the gene bits with gene value of 1 for mutation.

Differential mutation operator is introduced to perform mutation by differential strategy with the following formula:

$$X_i^{t+1} = X_i^t + \alpha * \left(X_{r1}^t - X_{r2}^t \right) \tag{30}$$

where t denotes the current evolutionary generation, X_i' denotes a random individual in the population, and α is the control factor, which generally takes the value range of $[0,1]$. The difference vector composed of two individuals is added to the random individual vector to construct a new individual vector, which accomplishes the mutation operation.

In order to boost the search performance of the algorithm, the control factor is dynamically adjusted:

$$X_i^{t+1} = X_i^t + \left(\alpha_{\min} + \frac{t\left(\alpha_{\max} - \alpha_{\min}\right)}{T} \right) * \left(X_{r1}^t - X_{r2}^t \right) \tag{31}$$

where $\alpha_{\min}$ represents the minimum value of the control factor, $\alpha_{\max}$ represents the maximum value of the control factor, t represents the number of generations of the current population evolution, and T denotes the maximum number of generations of the population evolution.

3.3.5 Algorithmic steps

The flow of NSGA-II algorithm based on improved variational operator is shown in Fig. 10.

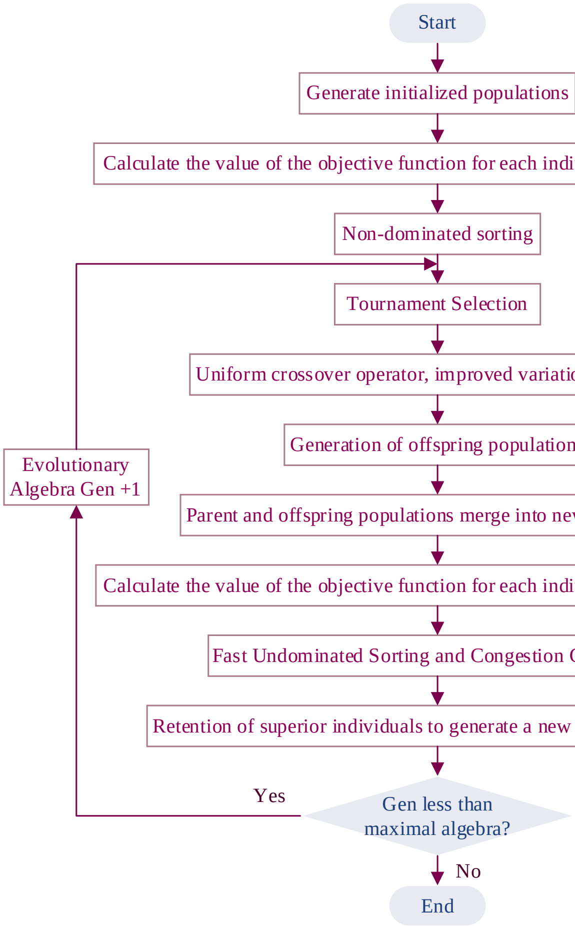

Figure 10: Flowchart for improving the NSGA-II algorithm

3.4 Experimentation and Analysis

3.4.1 Experimental environment and dataset

In this paper, Matlab R2016a is employed to implement the multi-objective optimization based recommendation algorithm (KGMO-RS) algorithm, in order to examine the effectiveness and feasibility of the algorithm, the difference in the recommendation performance is analyzed with the knowledge graph based graph convolutional recommendation algorithm (GCNKG-CR) and KGMO-RS algorithm by doing comparative experiments. After solving, Top-15 are selected from the Pareto solution set as the recommended Civics course resources. If the data in the Pareto solution set is less than 15, all the candidate sets are selected as recommended content.

Some documents and online learning data information containing the knowledge of Civic and Political Education courses are crawled from the data of an e-learning platform, and the knowledge points and relations are extracted after data cleansing and lexical processing, to construct the Civic and Political Education course ontology and form the experimental dataset after complementation.

3.4.2 Analysis of experimental results

To examine the recommended outcomes, the experimental results were analyzed using HV (Hyper Volume) values. The statistical findings of the HV values in the experiment are presented in Table 4. The HV box line plots of KGMO-RS and GCNKG-CR are shown in Figure 11.

From Table 4 and Fig. 11, the best, mean, and worst values of the HV metrics of KGMO-RS are 0.6625, 0.6481, and 0.6333, respectively, which are higher than that of GCNKG-CR, indicating that the convergence and diversity of KGMO-RS are better than that of GCNKG-CR.Meanwhile, the distribution range of the HV values of KGMO-RS is narrower than that of GCNKG-CR, which indicates that the KGMO -RS has better stability and better comprehensive performance in recommending resources for Civic Education courses.

Table 4: The statistical results of HV values in the experiment

Algorithm	Best	Mean	Worst
GCNKG-CR	0.5860	0.5578	0.5269
KGMO-RS	0.6625	0.6481	0.6333

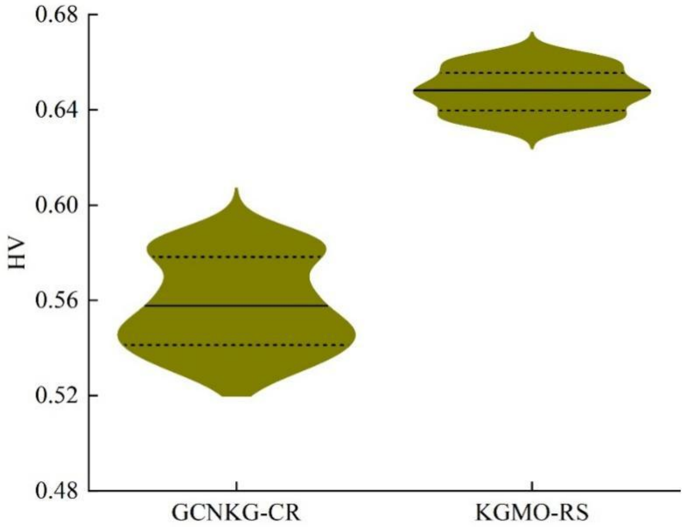

Figure 11: HV box plots of GCNKG-CR and KGMO-RS

The computation of resource recommendations using the GCNKG-CR and KGMO-RS algorithms respectively leads to two different Pareto fronts as shown in Fig. 12. The X-axis is the objective function Precision, which indicates the learner's satisfaction with the final recommendation result. The Y-axis is the objective function Diversity, which represents the difference between all the items in the recommended list. The Z-axis is the objective function Novelty, which indicates the ability of the recommender system to recommend to the user the less popular but in line with their interests.The Pareto frontier obtained by KGMO-RS has good distribution characteristics and better diversity, and better convergence effect, indicating that the finding of KGMO-RS on the Pareto frontier surface is more superior, and it is more suitable for solving the multi-objective Civic Education Curriculum Resource Recommendation Problem.

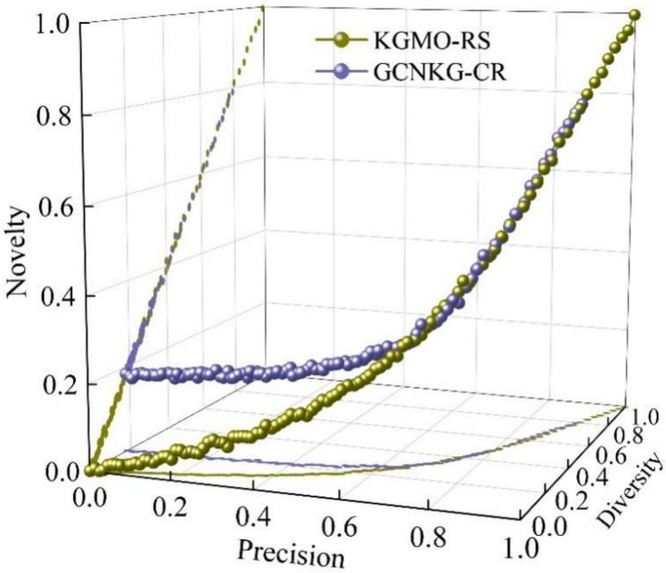

Figure 12: Pareto frontiers of GCNKG-CR and KGMO-RS

3.4.3 Convergence performance analysis

The Inverse Generation Distance (IGD) metric is capable of being employed to assess how closely the non-dominated solution set of the multi-objective optimization algorithm approximates the true Pareto-optimal front. As the IGD value decreases, the recommendation accuracy rises and the algorithm demonstrates better convergence and distribution performance. The calculation formula is listed below:

$$IGD(P,Q)=\frac{1}{|P|}\sum_{v\in P}d(v,Q) \tag{32}$$

where P represents the Pareto-optimal solution set situated on the genuine Pareto frontier, $|P|$ refers to the aggregate number of solution individuals comprised within the optimal solution set on the actual Pareto surface, Q is the collection of non-dominated solutions generated by the algorithm, and $d(v,Q)$ denotes the minimum Euclidean distance from a given individual v toward a target population Q.

The computed outcomes of the IGD metrics across the experiments are displayed in Table 5. The mean and variance of the IGD metrics of KGMO-RS are 2.5493e-02 and 7.2864e-05, respectively, which are smaller than that of the GCNKG-CR, indicating that KGMO-RS is better than the standard GCNKG-CR with respect to convergence and uniformity of distribution.

Table 5: Statistical results of IGD indicators in the experiment

Algorithm	Mean	Variance
GCNKG-CR	6.3482e-02	3.6735e-04
KGMO-RS	2.5493e-02	7.2864e-05

3.4.4 Evaluation of the utility of recommendations

In this paper, we use five-fold cross-validation, randomly divide the dataset into 5 parts, select one of them as the test set during each iteration, while the other 4 serve as the training set, and repeat this process 5 times, each time selecting a different training set. The average of the 5 experiments was used as an evaluation metric for recommendation utility, and the results of the recommendation algorithm were analyzed using 3 metrics: the percent of checking accuracy (P), the recall rate (R), and the Fl value.

The results of the recommendation evaluation metrics of KGMO-RS and GCNKG-CR are shown in Fig. 13.The P, R, and F1 values of KGMO-RS are 0.8426, 0.8831, and 0.8624, respectively, which are better than those of GCNKG-CR, which indicates that the KGMO-RS algorithm is more comprehensive in solving the present problem, and has a better recommendation effect.

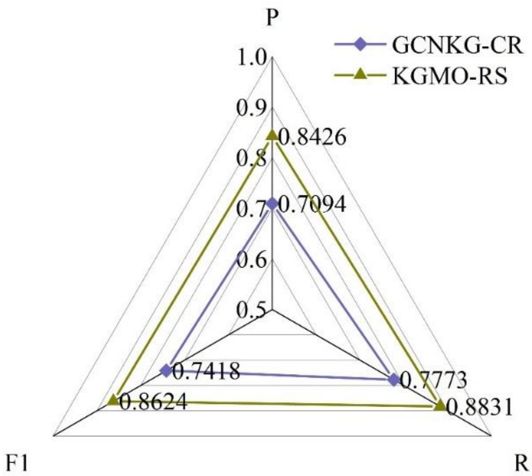

Figure 13: The results of the recommended evaluation indicators

4 Research on Student Satisfaction in Civics Classroom Based on KANO Modeling

This chapter uses the KANO model to investigate the types of student satisfaction needs and priority ranking in the Civics classroom as a way to provide guidance for the smart advancement of the theory system in the new era.

4.1 Research methodology

4.1.1 Questionnaire design

From the perspective of the teaching process, the curriculum, teaching and teachers are “three in one”, which together affect the satisfaction and sense of acquisition of students' learning. Therefore, the survey questionnaire mainly consists of basic personal information and classroom satisfaction scale, in which the classroom satisfaction scale mainly includes the factors of curriculum, teaching effect and teachers. Firstly, we determined the measurement indicators through text analysis, and secondly, through in-depth interviews to further understand the factors and suggestions that influence the classroom satisfaction of Civics and Politics courses as perceived by college students, and then we integrated and modified the interviews to obtain a KANO questionnaire indicator set containing three primary indicators, namely, curriculum factors, teaching effect factors and teacher factors, and 18 secondary measurement indicators, and invited college students to evaluate the measurement indicators from positive and negative perspectives, respectively. College students were invited to evaluate the indicators from positive and negative perspectives, and the corresponding ratings were divided into 5, 4, 3, 2, and 1. In addition, this study also designed the “Questionnaire on the Importance of Classroom Satisfaction Measurement Indicators for Students of Civics and Political Science Courses in Colleges and Universities," whose indicators were the same as those of the KANO questionnaire, and adopted the Likert five-point scale.

4.1.2 Survey data collection

After the questionnaire was created by Questionnaire.com, a total of 500 questionnaires were sent out to undergraduate participants through online social platforms, and 467 completed responses were successfully retrieved, achieving an effective recovery rate of 93.4%, after eliminating any responses that were obviously illogical or irrational. 500 questionnaires were distributed for "Questionnaire on the Importance of Classroom Satisfaction Measurement Indicators for undergraduate students majoring in Civics and Political Science across higher education institutions", and 486 questionnaires were recovered, with an overall response rate reaching 97.2%. According to the principle that the total number of questionnaire samples should not be less than 200 and more than 10 times of the number of survey items, confirming that the questionnaire sample volume complies with the standard, which ensures the validity of statistical analysis. At the same time, SPSS28.0 statistical software was utilized to assess the internal consistency of the questionnaire, and each question item was assigned according to the requirements of Likert scoring criteria, with results showed that: the Cronbach's α coefficient for the importance questionnaire was 0.979, while that for the satisfaction questionnaire reached 0.972; the Cronbach's α for positively worded items was 0.975, and for reverse-scored items was 0.993; collectively, both item types yielded Cronbach's α values exceeding 0.9, indicating that the questionnaire has good internal consistency and high data reliability.

4.2 Empirical analysis

4.2.1 Classroom Satisfaction Indicator Attribute Classification

In this paper, the data obtained from the questionnaire were counted, and the positive and negative findings of the 18 measures were identified as various types of attributes based on the percentage size. The classification of attributes corresponding to “Curriculum goals meet students' needs (A1)” is shown in Table 6. Among them: questionable attributes (Q), must-have attributes (M), one-dimensional attributes (O), excitement attributes (A), indifferent attributes (I), and reverse attributes (R).

A summary of the data for each expectation attribute in Table 6 yields the following counts for Q, M, O, A, I, and R: 80, 33, 70, 133, 150, and 1 respectively, corresponding to proportions of 17.13%, 7.07%, 14.99%, 28.48%, 32.12%, and 0.21%. The data reveal that students predominantly selected the undifferentiated attribute (I) in the largest number, and thus the measure "curriculum objectives meet students' needs (A1)" was classified into the undifferentiated factors.

The categorization results of the other 17 measurement indicators were calculated in the same way. The results show that five indicators belong to the required attributes (M), which are: textbooks and learning materials meet the learning needs (A3), the ratio of theoretical courses to practical courses is appropriate (A4), the course assessment methods are diversified (A5), understanding and mastering the knowledge points and theoretical connotations related to Civics and Politics courses (A8), and unshakeable Marxist beliefs and political will (A10). Three indicators belong to non-differentiated attributes (I), namely: reasonable class schedule (A2), flexible use of modern information technology to optimize classroom education and teaching resources (A7), and grasping the “emotional point” of Civic and Political Science class to arouse students' emotional resonance (A9). Nine indicators belong to the desired attributes

(O), which are: well-organized teaching to create a democratic, efficient and orderly classroom (A6), the capacity to examine and address real-world challenges by applying Marxist theories (A11), combining one's individual aspirations with the national vision, and devoting oneself to the great struggle for realizing the Chinese Dream (A12), adhering to the correct political direction (A13), possessing a profound sense of family and national sentiments and parenting (A14), master dialectical materialism and historical materialism (A15), have a broad intellectual, international, and historical perspective (A16), do not touch the red line of being a human being, do not break the bottom line of being a teacher (A17), be upright and honorable, and inspire students with the power of truth (A18).

Table 6: Questionnaire statistics of measurement indicators

Positive problem	Reverse problem				
	I like it very much	Take it for granted	It doesn't matter	Reluctantly accept	I don't like it
I like it very much	78 (Q)	17 (A)	16 (A)	100 (A)	70 (O)
Take it for granted	1 (R)	38 (I)	16 (I)	45 (I)	31 (M)
It doesn't matter	0 (R)	5 (I)	17 (I)	4 (I)	0 (M)
Reluctantly accept	0 (R)	1 (I)	15 (I)	9 (I)	2 (M)
I don't like it	0 (R)	0 (R)	0 (R)	0 (R)	2 (Q)

4.2.2 Screening of Classroom Satisfaction Improvement Indicators and Key Indicators

This paper draws on the customer satisfaction coefficient of KANO model to judge the indicators that need to be improved and their degree, calculates the Better-Worse coefficient of 18 measurement indicators and draws the quartile diagram of BetterWorse coefficient analysis of KANO model, which can provide an important basis for the suggestions to enhance the learning satisfaction among students pursuing Civics and Politics programs at higher education institutions. Worse coefficient calculation formula is as follows:

Customer satisfaction index:

$$Better(SI) = (A+Q)/(A+O+M+I) \tag{33}$$

Customer dissatisfaction coefficient:

$$Worse(DSI) = (-1)(O+M)/(A+O+M+I) \tag{34}$$

The Better coefficient tends to be a positive value, indicating that across the three dimensions of curriculum design, teaching effectiveness, and instructor quality within the Civics and Political Science course, if a certain attribute provided indicators of student satisfaction will be enhanced, the nearer this positive value approaches 1, the more pronounced the improvement effect becomes, reflecting a faster rate of satisfaction enhancement. The opposite is what the Worse coefficient represents. Substituting the values of the attributes of each measurement indicator divided into the formula, the Better and Worse values of each indicator are presented in Table 7.

Table 7: Better-Worse values of student classroom satisfaction measurement indicators

Measurement index	SI	DSI	Attribute classification
A1	0.5209	-0.2668	I
A2	0.4846	-0.3709	I
A3	0.4828	-0.6437	M
A4	0.4592	-0.5699	M
A5	0.4551	-0.6168	M
A6	0.6351	-0.4382	O
A7	0.6115	-0.3732	I
A8	0.4874	-0.6831	M
A9	0.5426	-0.3702	I
A10	0.4747	-0.5973	M
A11	0.5643	-0.4581	O
A12	0.6058	-0.5079	O
A13	0.5901	-0.5927	O
A14	0.6513	-0.5886	O
A15	0.6321	-0.4986	O
A16	0.6472	-0.5213	O
A17	0.6207	-0.6358	O
A18	0.6633	-0.6206	O

The sensitivity matrix of the KANO model was incorporated into the KANO model using the SI values of each measurement index in Table 7 as the horizontal coordinates and the DSI values as the vertical coordinates as shown in Figure 14, "with *O* as the center of the circle and *OP* as the radius of the circle, and *OP* is the length of the line segment that passes through the intersection of the vertical and horizontal points at 0.5". The intersection of SI and DSI is used as a scatter plot in the matrix, and the farther away from the circle point *O* the higher the sensitivity of the measure represented by the scatter point, and vice versa, the lower it is.

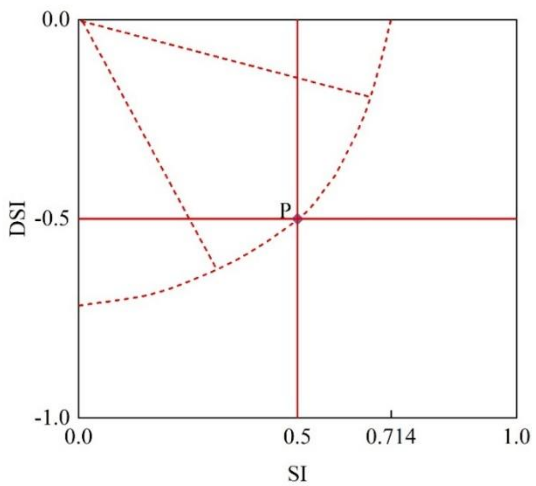

Figure 14: Sensitivity matrix of the KANO model

The values of Better-Worse coefficients for the 18 measures were substituted to draw a quartile plot of students' classroom satisfaction as shown in Figure 15. Quadrants 1 to 4 represent desired attributes, charismatic attributes, undifferentiated attributes, and essential attributes, respectively.

The six measures in the first quadrant are all from the “teacher factor”. Their importance values are higher than the average value of 4.27 of the 18 indicators in the questionnaire, which means that these 6 indicators are high expectations and should be grasped and optimized to improve students' satisfaction with the Civics and Political Science classes. Particular attention is paid to the fact that the A18 measurement indicator located in this quadrant is the one furthest away from the dot O. “Teachers of Civics should be upright and honorable, and inspire students with the power of truth” is an important expectation of college students for teachers of Civics, and if we can ensure that students' expectations of this indicator are met in a timely manner, then students' classroom satisfaction will increase dramatically, and their dissatisfaction will be eliminated greatly.

The five measures in the second quadrant are mainly from “curriculum factors” and “teaching effectiveness factors”, but only three measures, A6, A7, and A11, have an importance value higher than the mean value of 4.27, which is called a high attractiveness attribute in this quadrant, meaning that they are highly attractive to students. It means that it is highly attractive to students. It can be seen that the five measures should be prioritized to develop and strengthen the demand for high charisma attributes and also the need to increase the highlights of the curriculum, so that when students are motivated, their satisfaction will increase significantly. In addition, the A12 measure in the Teaching Effectiveness Factor is located on the line separating the first and second quadrants, and the KANO model theory suggests that the Desired Attribute is more important than the Charisma Attribute, so efforts should be made to optimize the A12 measure so that it is categorized in the Desired Attribute quadrant.

There is only one indicator in the third quadrant, which comes from the "reasonable schedule (A2)" in the "curriculum factors". Meanwhile, A2 is the closest indicator to the dot O. Therefore, it is recommended to grasp the students' psychology in the process of class schedule arrangement, and rationally arrange the students' Civics schedule in terms of the effective learning time of the course and the degree of difficulty of the course and other courses, which can substantially enhance both the academic quality and learning efficiency of students' learning, and thus increase the students' satisfaction in the classroom.

The five indicators located in the fourth quadrant are mainly from some indicators of "curriculum factors" and "teaching effectiveness factors", and their importance values are higher than the mean values. For these five essential attribute measures, it is important to guarantee that the fundamental learning requirements of students are adequately addressed rather than over-enforcing and over-developing the essential qualities.

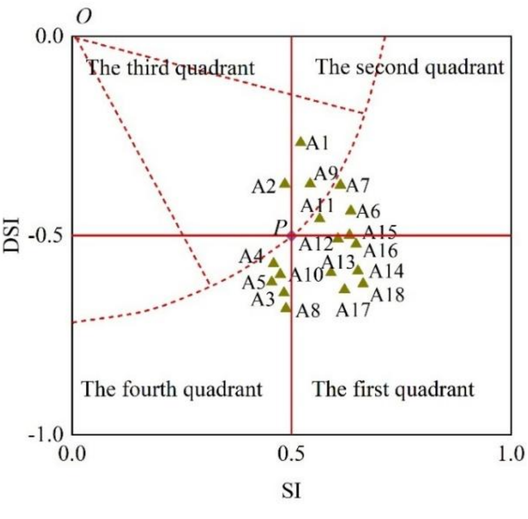

Figure 15: Quartile plot of student classroom satisfaction

5 Conclusion

Drawing on the entity naming recognition framework of BiLSTM-Attention-CRF alongside the inter-entity relationship extraction model BERT-BiLSTM-Attention, we constructed a knowledge graph of Civics courses, and then proposed a multi-objective optimization algorithm KGMO-RS for course recommendation utilizing the knowledge graph, and investigated the satisfaction indicators of student satisfaction in the Civics classroom based on the KANO model Prioritization.

Contrasting the BiLSTM-CRF model against the CNN+BiLSTM-CRF model, the BiLSTM+Attention-CRF model developed in this study demonstrates superior performance across accuracy rate, sensitivity rate, and comprehensive F-measure, suggesting that integrating BiLSTM, CRF, and the attention mechanism enables more effective extraction of both local and global textual features, thereby yielding enhanced named entity recognition outcomes. Meanwhile, compared with CNN model, PCNN model, BiLSTM model and Attention+ BiLSTM model, the BERT-BiLSTM-Attention model developed in this study attains the top-ranking results, with a comprehensive F-measure exceeding 87%, which is suitable for the entity-relationship categorization task and ensures the integrity of the relational data within the constructed Civics Course Knowledge Graph. In addition, by benchmarking the KGMO-RS algorithm against the GCNKG-CR baseline and assessing the model through HV and IGD metrics, this study confirms its diversity and stability, demonstrating the algorithm's strong global optimization and convergence capabilities, as well as its practical effectiveness in course resource recommendation.

In order to improve students' satisfaction with the Civics classroom, we can focus on the key elements based on the hierarchical priority structure derived from the KANO model indicators, which is "must-have attributes>one-dimensional attributes>attractive attributes>indifferent attributes", to grasp the dynamic needs of students' learning in the Civics classroom, and to optimize the content of Civics teaching and the support services for students' learning. With the aim of elevating the engagement level within the ideological and political theory course setting, and subsequently advance the intelligent progression of the theoretical framework governing Civics and Politics education throughout the new era, we have established a comprehensive Civics and Politics educational structure suited to the demands of the new era.

About the Author

Jing Zhu (born in October 1981), female, Han ethnicity, from Shaoxing, Zhejiang Province. She holds a bachelor's degree and is a lecturer at the College of Urban Construction, Wuchang Institute of Technology. She is also a national second-level psychological counselor and a national second-level employment tutor. Her main research directions include students' ideological and political education, university students' employment, entrepreneurship, and the mental health of college students.

References

[1] Udvaros, J., & Forman, N. (2023). Artificial intelligence and Education 4.0. In INTED2023 proceedings (pp. 6309-6317). IATED.

[2] Lin, Y., Chen, H., Xia, W., Lin, F., Wang, Z., & Liu, Y. (2025). A Comprehensive Survey

on Deep Learning Techniques in Educational Data Mining: Y. Lin et al. Data Science and Engineering, 1-27.

[3] Jasim, A. H., & Hoomod, H. K. (2025, April). Intelligent student behavior recognition system for the classroom environment using hybrid deep learning. In AIP Conference Proceedings (Vol. 3282, No. 1, p. 030011). AIP Publishing LLC.

[4] Zhou, F. (2022). Research on teaching resource recommendation algorithm based on deep learning and cognitive diagnosis. Journal of Healthcare Engineering, 2022(1), 5776341.

[5] Lin, J. (2024). Deep Learning-Driven Optimization Strategies for Teaching Decisions in Smart Classrooms. International Journal of Interactive Mobile Technologies, 18(15).

[6] Tsai, S. C., Chen, C. H., Shiao, Y. T., Ciou, J. S., & Wu, T. N. (2020). Precision education with statistical learning and deep learning: a case study in Taiwan. International Journal of Educational Technology in Higher Education, 17(1), 12.

[7] Liu, W., & He, C. (2022). Curriculum-Based Ideological and Political Education: Research Focuses and Evolution. International Education Studies, 15(5), 28-35.

[8] Gong, G. (2017). The integration of ideological and political education in colleges and universities and employment of college students. International technology management, 5, 30-32.

[9] Liu, X., Xiantong, Z., & Starkey, H. (2023). Ideological and political education in Chinese Universities: structures and practices. Asia Pacific Journal of Education, 43(2), 586-598.

[10] Zhang, Y., Yan, Y., Kumar, R. L., & Juneja, S. (2024). Improving college ideological and political education based on deep learning. International Journal of Information and Communication Technology, 24(4), 431-447.

[11] Ai, S., & Ding, H. (2025). Design optimization of university ideological and political education system based on deep learning. Scientific Reports, 15(1), 18134.

[12] He, X., Chen, P., Wu, J., & Dong, Z. (2021). Deep learning-based teaching strategies of ideological and political courses under the background of educational psychology. Frontiers in psychology, 12, 731166.

[13] Zhou, J., & Shi, S. (2024, November). Research on Teaching Strategies of Ideological and Political Courses in Colleges and Universities from the Perspective of Educational Psychology Based on Deep Learning. In 2024 4th International Conference on Digital Society and Intelligent Systems (DSInS) (pp. 64-68). IEEE.

[14] Feng, T. (2023). Research on the development path of internet ideological and political education based on deep learning. Soft Computing, 1-10.

[15] Di, H., Zhang, H., & Li, P. (2022). Teaching quality of ideological and political education in colleges based on deep learning. International Journal of e-Collaboration (IJeC), 19(4), 1-15.

[16] Cai, S. (2024). Utilization of deep learning in ideological and political education. Journal

of Intelligent Systems, 33(1), 20240206.