Implications of Immersive Online Game Expert System for Ideological and Political Teaching Based on Set Theory and Artificial Intelligence

Hong Chen[1*]

[1]Institute of civil engineering, Putian college, Putian ,Fujian ,351100, China
13850299816@163.com

Corresponding author: Hong Chen,13850299816@163.com

Abstract. This paper analyzes ideological and political teaching by combining covering set theory and artificial intelligence technology, constructs a corresponding expert system, and presents a design method of interval type-2 fuzzy model based on rule reduction and differential evolution.Moreover, the rule reduction method is used in this paper to reduce the number of rules of the interval type-2 fuzzy model to a certain extent, and reduce the parameter scale of the model.On this basis, this paper uses the differential evolution algorithm to optimize the parameters of the reduced interval type-2 fuzzy model to further improve the performance of the model.The research shows that the expert system of ideological and political teaching based on covering set theory and artificial intelligence proposed in this paper has a good effect and can effectively process various information of ideological and political teaching.

Key words: covering set; artificial intelligence; ideology and politics; teaching; Online Game expert system
DOI: https://doi.org/10.14733/cadaps.2024.S5.138-156

1 INTRODUCTION

Knowledge point is the basic unit of mutual information transfer in the course, and the correlation of knowledge point refers to the connection between the knowledge points in the course. It is ubiquitous and affects each other, so it has mutuality and universality. Moreover, the correlation between knowledge points often needs to be discovered through association analysis mining, which is not obvious result data, so it has certain implicitness [4].When the learner's understanding of a knowledge point is insufficient, the possibility of deviation in the learner's understanding of the knowledge point itself is excluded, which may also be affected by its associated knowledge point. The antecedent knowledge points represent forward knowledge points that may lead to biases in understanding a certain knowledge point.The backward knowledge point represents the backward knowledge point that a certain knowledge point may affect [8].Compared with the knowledge points with better understanding and sufficient mastery, teachers should pay more attention to mining and analyzing the knowledge points that student have less mastery. Therefore, dividing the content of ideological and political teaching into multiple independent knowledge points according to the syllabus, and mining the relevance of knowledge points for them has a guiding role in clarifying the relationship between the key content of the students' curriculum and knowledge points [7].

Association analysis has been widely used in the field of education, and three commonly used association analysis algorithms are: Apriori algorithm, FP-Growth algorithm and MS-Apriori algorithm. The Apriori algorithm is a recursive algorithm for mining frequent itemsets based on two stages, in which any itemsets satisfy the downward closed property. If one of the items is a frequent itemset, any subset of the item can be regarded as frequent In the process of mining, the data set needs to be scanned many times. When the amount of data is large, the problem of low mining efficiency will occur. The FP-Growth algorithm finds frequent itemsets in the dataset by creating an FP-Tree. The dataset is stored in a tree structure before mining, and only two scans of the dataset are required during the entire mining process [16]. The advantage of this algorithm is that it can adapt to association rules of different lengths, and the mining speed is greatly improved due to the different structures. MS-Apriori algorithm is a multi-support Apriori algorithm. A min_sup is set for each item, and the minimum value of min_sup of all items in the item set is the support threshold of the item set [6]. The advantage of this algorithm is that it can mine the association rules of rare items, so that the mining results are comprehensive. Due to the large data set in knowledge point correlation mining, the Apriori algorithm and the MSApriori algorithm will generate multiple candidate item sets in the process of repeatedly scanning the database. , Apriori needs to rescan the entire dataset, resulting in slow mining speed and exponential increase in time complexity. The FP-Growth algorithm does not need to generate a candidate item set during the mining process, and only needs to scan the databases on both sides. Compared with the previous two algorithms, the mining speed will be greatly improved. The FP-Growth algorithm is selected as the knowledge point association mining. algorithm [10].

Because the original data set is affected by human factors or other reasons in the actual database entry, there is a possibility that the entry results are inconsistent. Therefore, before the knowledge point correlation mining is carried out, in order to ensure the rationality of the data, avoid mining results. To have an impact, a preprocessing operation on the original dataset is required [13]. Data preprocessing mainly includes data cleaning, data integration, data transformation and data discretization. The data set used in the association mining of knowledge points does not need to change the size of the data volume through data integration [5]. (1) Data cleaning There is usually a lot of redundant information in the original data set, and it is necessary to consider removing useless data and interfering data from it. Data cleaning follows the following two principles: 1) For the processing of missing values, due to the large number of samples in the data set, if a certain knowledge point score of a student is missing, you can choose to remove the entire data record of this student. 2) Redundant column deletion. In order to ensure the privacy of the data, the real student number is not used to number it. Therefore, to mine the correlation between knowledge points, only the four attribute columns of number, knowledge point, knowledge point score, and knowledge point score are needed. Remove other redundant column information. (2) Data transformation Before mining the relevance of knowledge points, data transformation needs to be carried out to prepare for the next mining process. Combined with the specification of the data structure in the mining process, it is necessary to transform the data into a format with students as rows, knowledge points as columns, and data in tables as knowledge point scoring rates [1]. the focus shifts to the potential of online gaming as an educational tool. It examines how gamification elements, such as quests, challenges, rewards, and leaderboards, can be incorporated into the expert system to foster active learning and motivate students. The section also discusses the advantages of collaborative gaming experiences and the potential for social interaction and knowledge sharing

In order to make the algorithm more reliable and accurate to mine the correlation of knowledge points, literature [14] constructed a knowledge point correlation mining method based on association analysis. Classification mining, hierarchical mining according to the proportion of knowledge point samples in the total data volume, and the use of FPGrowth algorithm to mine the correlation between knowledge points. Constructing sample data according to the score rate for classification mining. In the research of knowledge point correlation mining, usually only from the perspective of knowledge point score or loss point to mine the knowledge point correlation, you can get a number of association rules, composed of associations. The rule table is not comprehensive enough [12]. Because in the actual ideological and political teaching situation, while exploring the correlation between knowledge points through the high score rate of students, teachers need to pay more attention to the correlation between knowledge points with a low score rate (loss of points). The correlation of knowledge points is excavated from two aspects: rate and low score rate, and the obtained association rules are more comprehensive, which can provide certain guidance for improving students' learning effect and teachers' ideological and political teaching [11]. Data Layering An exam contains multiple knowledge points. The data sets of all knowledge points come from multiple exams. The knowledge points examined in each exam overlap but are not identical, that is, each knowledge point appears in different exams. The frequency is inconsistent, which will lead to differences in the number of times each knowledge point appears in the data set. For different knowledge points, the number of data sets related to the student's score rate for each knowledge point is different. When the difference between the data sets is too large, the possibility of frequent itemsets in mining the overall data set will drop sharply [15]], to mine the correlation between knowledge points, it is necessary to consider dividing the sample data according to the proportion of knowledge point samples to the total data volume, stratifying the data, and further mining the correlation between knowledge points, so as to avoid the occurrence of factors in each layer during mining. Insufficient number of dataset samples affects the phenomenon that the number of frequent itemsets generated is insufficient. When the data set is large, MySQL can be used to import the students' knowledge point score rate table, design a query statement to count the number of knowledge point samples, and get the number of samples corresponding to each knowledge point and the proportion of the total number of people [9] . The knowledge point correlation mining is performed on each layer separately, and the mining results of different layers can be obtained. In the mining results, only the association rules related to the knowledge points of this layer need to be retained, and any association rules that do not contain the knowledge points of this layer are not retained [2]. The result of knowledge point association mining is determined based on the dataset samples of the corresponding layer, and the knowledge point dataset samples involved in the discarded association rules come from the dataset samples of other layers [3].

This paper analyzes ideological and political teaching by combining the covering set theory and artificial intelligence technology, and constructs a corresponding expert system to improve the effect of ideological and political teaching.

2 COVERING SET THEORY

Differential evolution (DE) algorithm is a heuristic random search algorithm, first given by Storn and Price in 1995. The specific process is shown in Figure 1.

The key aspects of the DE algorithm are as follows:

1) Initialize the population

If the population size is assumed to be N and the dimension is D, the initialized population can be obtained as $\left(X_1, X_2, \ldots, X_N\right)$.

2) Variation

Mutation refers to randomly selecting two different individuals in the population to scale their vector differences and then perform vector synthesis with the individuals to be mutated to obtain a mutation vector $V_1(g+1)$.

$$V_1(g+1)=X_{r1}(g)+F\left(X_{r2}(g)-X_{r3}(g)\right) \tag{1}$$

3) Crossover

Crossover is the operation between the individuals of the original population and the generated variants, and the purpose is to improve the diversity of the population, and the operation method is as follows.

$$U_{i,j}(g+1)=\begin{cases} V_{i+1}(g=1) & if \;\; rand(0,1)\le C \\ x_{i,j}(g) & otherwise \end{cases} \tag{2}$$

4) Selection

Selection refers to selecting a better individual as a new individual, and the operation is as follows:

$$X_i(g+1)=\begin{cases} U_{i,j}(g+1) & if \;\; f\left(U_{i,j}(g+1)\right)\le f\left(X_i(g)\right) \\ x_{i,j}(g) & otherwise \end{cases} \tag{3}$$

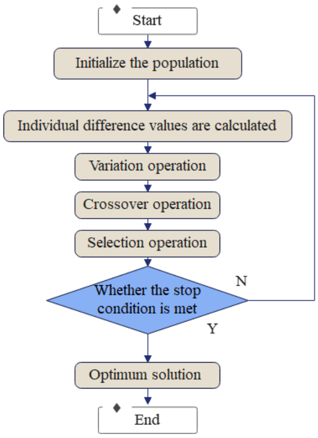


Figure 1: Flow chart of differential evolution algorithm.

The main optimization process of the DE method is as follows:

1) The algorithm initializes the relevant parameters, sets the current algebra $g=0$ and gives the maximum number of iterations $g_{\max}$, as well as the population size N, the crossover probability C, and the scaling factor F;

2) The algorithm randomly generates the initial population $X_1, X_2, \ldots, X_N$, and calculates the fitness value of the individual;

3) The algorithm updates the algebra $g = g + 1$;

4) The algorithm sets the target vector $i = 1$;

5) As described in the mutation operation, the algorithm randomly selects three different individuals X_{r1}, X_{r2} and X_{r3} except the target vector i to generate the mutation vector V_I;

6) The algorithm performs a crossover operation to obtain U_i;

7) The algorithm calculates the U_i fitness value, and then performs the selection operation;

8) The algorithm sets $i = i + 1$, then returns to (5) until $i = N$, otherwise go to the next step;

9) The algorithm judges whether the maximum number of iterations $g_{\max}$ is reached. If it is satisfied, the algorithm outputs the result, otherwise it returns (3).

When the number of interval type-2 covering set rules is large, the optimization difficulty and complexity will increase sharply. In order to reduce the difficulty of model construction, one of the most intuitive ideas is to reduce the number of fuzzy rules, and then optimize the learning of parameters on the basis of rule simplification.

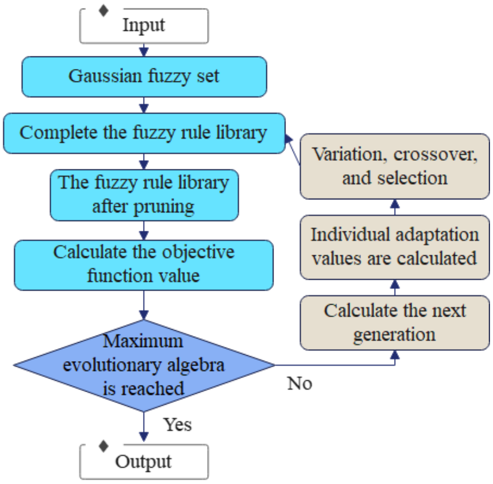


Figure 2: Overall flow chart of interval type-2 covering set based on rule reduction and differential evolution algorithm.

The overall process of the interval type-2 covering set based on rule reduction and differential evolution given in this chapter is shown in Figure 2. The specific process is as follows:

1.According to the training data set, the interval type-2fuzzy division of each input variable is generated, and a complete interval type-2 fuzzy rule base is generated;

2.The algorithm obtains the activation strength matrix of each rule through the training data, and realizes the pruning of the interval type-2 fuzzy rules according to the matrix;

3.On the basis of the reduced interval type-2 fuzzy rule base, the differential evolution method is used to optimize the parameters of the interval-type-2 fuzzy rules;

4.The resulting final interval type-2 fuzzy prediction model is output.

The implementation of the key links in the design method will be given below.

In this chapter, a Gaussian fuzzy set $\bar{A}$ with uncertain standard deviation value is selected, as shown in Figure 3(a).We assume that the constructed model has n input variables $x_1 \in X_1, x_2 \in X_2, \cdots, x_n \in X_n$ and one output variable $y \in Y$ .For the input universe X_i , it is divided by Gaussian fuzzy set.Moreover, we assume that there are N_i interval-type two fuzzy sets used to divide the input variable x_i , which are $\tilde{A}_i^1, \tilde{A}_i^2, \cdots, \tilde{A}_i^{N_i}$, respectively, as shown in Figure 3.

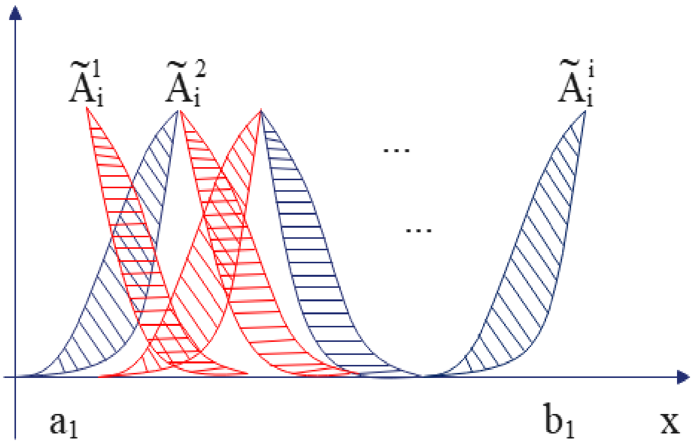


Figure 3: Type-2 fuzzy partition of input variable x_i **.**

In Figure 3, the upper and lower membership functions of fuzzy set $\bar{A}$ are determined by the following formula.

$$\bar{\mu}_{\tilde{A}_i^j}(x) = \exp\left[-\frac{1}{2}\left(\frac{x-m_i^j}{\underline{\sigma}_i^j}\right)^2\right] \tag{3}$$

$$\underline{\mu}_{\tilde{A}_i^j}(x) = \mathrm{xp}\left[-\frac{1}{2}\left(\frac{x-m_i^j}{\underline{\sigma}_i^j}\right)^2\right] \tag{4}$$

Among them, $m_i^j = a_i + \frac{(j-1)(b_i - a_i)}{N_i - 1}, \underline{\sigma}_i^j = \frac{b_i - a_i}{6(N_i - 1)}, \overline{\sigma}_i^j = \frac{b_i - a_i}{3(N_i - 1)}$.

Therefore, the rules of the complete interval type-2 fuzzy rule base of the designed interval type-2 covering set are as follows:

$$\left\{R(i_1,\cdots i_n): if\ x_1\ is\ \tilde{A}_1^{i1}, x_2\ is\ \tilde{A}_2^{i2}, \cdots, x_n\ is\ \tilde{A}_n^{in}\ then\ y\ is \left[\underline{w}^{i_1,\cdots,i_n}, \overline{w}^{i_1,\cdots,i_n}\right]\right\} \tag{5}$$

In the case of input $x = (x_1, x_2, \cdots, x_n)^T$, according to the inference process of the interval type-2 covering set, the expression of the activation strength $F^{i_1,\cdots,i_n}(x)$ corresponding to the rule $R(i_1,\cdots,i_n)$ can be obtained.

$$F^{i_1,\cdots,i_n}(x) = \left[\underline{f}^{i_1,\cdots,i_n}(x), \overline{f}^{i_1,\cdots,i_n}(x)\right] \tag{6}$$

Among them,

$$\underline{f}^{i_1,\cdots,i_n}(x) = \underline{\mu}_{\tilde{A}_1^{i_1}}(x_1) * \underline{\mu}_{\tilde{A}_2^{i_2}}(x_2) * \cdots * \underline{\mu}_{\tilde{A}_n^{i_n}}(x_n) = \prod_{j=1}^{n} \underline{\mu}_{\tilde{A}_j^{i_j}}(x_j) \tag{7}$$

$$\overline{f}^{i_1,\cdots,i_n}(x) = \overline{\mu}_{\tilde{A}_1^{i_1}}(x_1) * \overline{\mu}_{\tilde{A}_2^{i_2}}(x_2) * \cdots * \overline{\mu}_{\tilde{A}_n^{i_n}}(x_n) = \prod_{j=1}^{n} \overline{\mu}_{\tilde{A}_j^{i_j}}(x_j) \tag{8}$$

On this basis, the specific output results obtained by the BMM method are:

$$y(x) = \alpha y_1(x) + (1-\alpha) y_r(x) \tag{9}$$

Among them, y_1 and y_r are the left and right endpoints of the reduced output, respectively, and its expression is:

$$y_1(x) = \frac{\sum_{i_1=1}^{N_1} \cdots \sum_{i_n=1}^{N_n} \underline{f}^{i_1,\cdots i_n}(x) \underline{w}^{i_1,\cdots i_n}}{\sum_{i_1=1}^{N_1} \cdots \sum_{i_n=1}^{N_n} \underline{f}^{i_1,\cdots i_n}(x)} \tag{10}$$

$$y_r(x) = \frac{\sum_{i_1=1}^{N_1} \cdots \sum_{i_n=1}^{N_n} \overline{f}^{i_1,\cdots i_n}(x) \overline{w}^{i_1,\cdots i_n}}{\sum_{i_1=1}^{N_1} \cdots \sum_{i_n=1}^{N_n} \overline{f}^{i_1,\cdots i_n}(x)} \tag{11}$$

For the convenience of calculation, formula (9) is rewritten as follows:

$$y(x)=\alpha\frac{\sum_{i_1=1}^{N_1}\cdots\sum_{i_n=1}^{N_n}\underline{f}^{i_1,\cdots i_n}(x)\underline{w}^{i_1,\cdots i_n}}{\sum_{i_1=1}^{N_1}\cdots\sum_{i_n=1}^{N_n}\underline{f}^{i_1,\cdots i_n}(x)}+(1-\alpha)\frac{\sum_{i_1=1}^{N_1}\cdots\sum_{i_n=1}^{N_n}\overline{f}^{i_1,\cdots i_n}(x)\overline{w}^{i_1,\cdots i_n}}{\sum_{i_1=1}^{N_1}\cdots\sum_{i_n=1}^{N_n}\overline{f}^{i_1,\cdots i_n}(x)}$$

$$=\alpha\frac{\sum_{k_{i_1i_2\cdots i_n}=1}^{M}\underline{f}^{k_{i_1i_2\cdots i_n}}(x)\underline{w}^{k_{i_1i_2\cdots i_n}}}{\sum_{k_{i_1i_2\cdots i_n}=1}^{M}\underline{f}^{k_{i_1i_2\cdots i_n}}(x)}+(1-\alpha)\frac{\sum_{k_{i_1i_2\cdots i_n}=1}^{M}\overline{f}^{k_{i_1i_2\cdots i_n}}(x)\overline{w}^{k_{i_1i_2\cdots i_n}}}{\sum_{k_{i_1i_2\cdots i_n}=1}^{M}\overline{f}^{k_{i_1i_2\cdots i_n}}(x)} \tag{12}$$

Among them,

$M=\prod_{j=1}^{n}N_j$, $i_j=1,2,\cdots,N_j$, $k_{i_1i_2\cdots i_n}=(i_n-1)\cdot N_1\cdots N_{n-1}+(i_{n-1}-1)\cdot N_1\cdots N_{n-2}+\cdots+(i_2-1)\cdot N_1+i_1$ For simplicity, $y(x)$ is further rewritten as:

$$y(x)=\alpha\frac{\sum_{k=1}^{M}\underline{f}^k(x)\underline{w}^k}{\sum_{k=1}^{M}\underline{f}^k(x)}+(1-\alpha)\frac{\sum_{k=1}^{M}\overline{f}^k(x)\overline{w}^k}{\sum_{k=1}^{M}\overline{f}^k(x)}$$

$$=\alpha\sum_{k=1}^{M}\underline{\gamma}^k(x)\underline{w}^k+(1-\alpha)\sum_{k=1}^{M}\overline{\gamma}^k(x)\overline{w}^k \tag{13}$$

Among them,

$$\underline{\gamma}^k(x)=\frac{\underline{f}^k(x)}{\sum_{k=1}^{M}\underline{f}^k(x)} \tag{14}$$

$$\overline{\gamma}^k(x)=\frac{\overline{f}^k(x)}{\sum_{k=1}^{M}\overline{f}^k(x)} \tag{15}$$

In the above formula, $\left[\underline{\gamma}^k(x)\ \ \overline{\gamma}^k(x)\right]$ represents the activation strength of the fuzzy rule $R(i_1,\cdots,i_n)$ corresponding to the fuzzy rule $k_{i_1i_2\cdots in}$ in the complete fuzzy rule base, which reflects the importance of the fuzzy rule to the input data.

Therefore, the specific pruning steps of the complete fuzzy rule base are as follows:

1.First, the algorithm calculates the activation strength of the input vector x^t in the training data (X^t,y^t) and its corresponding M fuzzy rules, and constructs an initial activation matrix H_0.

$$H_0 = \begin{bmatrix} \gamma^1(x^1) & \cdots & \gamma^M(x^1) \\ \vdots & \vdots & \vdots \\ \gamma^1(x^N) & \cdots & \gamma^M(x^M) \end{bmatrix}_{N\times M} \tag{16}$$

2.Then, the algorithm calculates the maximum value of the matrix H_0 in the k-th column.

$$\gamma^k = \max_{t=1}^{N} \gamma^k(x^t) \tag{17}$$

Among them, $k = 1, 2, \cdots, \mathrm{M}$.

3.Finally, the algorithm sets the threshold ε. When $\gamma^k / \max_{k=1}^{M} \gamma^k < \varepsilon$, the algorithm deletes the corresponding k-th fuzzy rule and the k-th column in the matrix H_0.

We assume that the fuzzy rule base has M ' interval type-2 fuzzy rules remaining after rule reduction, it is expressed as:

$$\left\{R^k : if\ x_1\ is\ \tilde{A}_1^k, x_2\ is\ \tilde{A}_2^k, \cdots, x_n\ is\ \tilde{A}_n^k\ then\ y\ is \left[\underline{w}^k, \overline{w}^k\right]\right\}_{k=1}^{M} \tag{18}$$

In the coverage set of interval type-2, the interval type-2 fuzzy set in the antecedent of the interval type-2 fuzzy rule can be obtained by intuitive division, as shown in Figure 3. However, in order to obtain good performance, it is still necessary to optimize the rule consequent of the interval type-2 covering set.

In this chapter, the DE algorithm will be used to optimize the parameters. The specific optimization learning steps are as follows:

1.The algorithm initializes the consequent parameters of the interval type-2 coverage set and defines the objective function;

2.The algorithm encodes M ' fuzzy rules after rule reduction;

3.The algorithm calculates the corresponding objective function value according to the objective function;

4.The algorithm performs crossover, mutation and selection operations, and determines whether the maximum number of iterations is reached. If the maximum number of iterations is reached, the algorithm stops iterating, otherwise returns (3);

5.The algorithm outputs the optimal parameter value that satisfies the condition.

3 EXPERIMENTAL VERIFICATION

In order to verify the effectiveness of the proposed method, this section will apply it to the ideological and political teaching prediction and ideological and political teaching information prediction experiments, and will compare and analyze with the ANFIS and BPNN models.

The original data is shown in Figure 4 (in order to visually show the characteristics of the dataset, some data are selected for drawing).

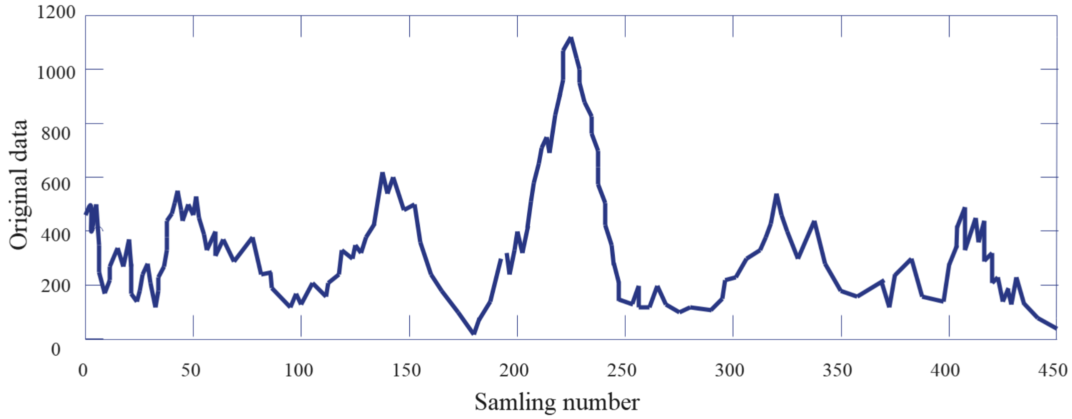


Figure 4: Experimental data of ideological and political teaching.

It can be seen from Figure 5 that the data set has relatively large daily fluctuations under the premise of a certain periodicity, and the uncertainty is obvious.

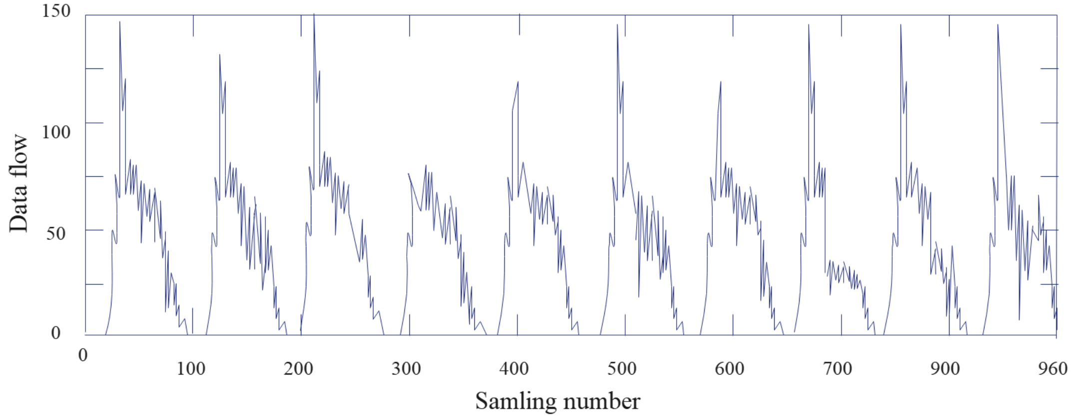


Figure 5: Experimental data on the amount of information in ideological and political teaching.

In this paper, two models, adaptive fuzzy inference system (ANFIS) and back propagation neural network (BPNN), are selected as comparison models. A brief introduction to the two models is given below.

a. Adaptive fuzzy inference system

ANFIS organically combines fuzzy logic and neural network, automatically extracts $if - then$ rules, and uses back propagation algorithm (BP) and least squares to adjust the parameters of the front and rear components to obtain the optimal solution. Its structure is shown in Figure 6.

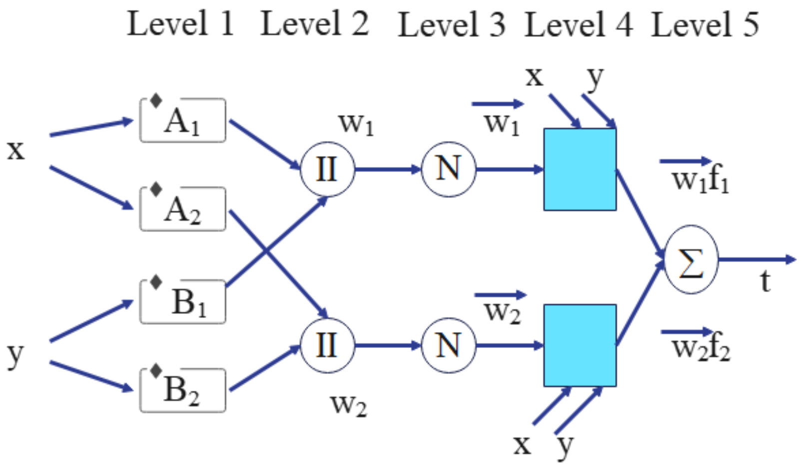


Figure 6: ANFIS structure diagram.

It is assumed that the rule base adopted by ANFIS is as follows:

$$\left\{ R^l : x_1 = A_1^l, x_2 = A_2^l, \cdots, x_n = A_n^l \rightarrow y_l\left(x\right) = c_0^l + \sum_{i-1}^{n} c_i^l x_i \right\}_{i=1}^{M} \tag{19}$$

The structural framework of ANFIS is shown in Figure 6. There are 5 layers in total.

Layer 1 is the blurring layer. This layer uses the membership function to fuzzify the input data $x_1, x_2, \cdots x_n$, and the membership of the output corresponding fuzzy set is as follows:

$$\mu_{A_j^l}\left(x_j\right) = \exp\left[-\frac{1}{2}\left(\frac{x_j - m_i}{\sigma_i}\right)^2\right] \tag{20}$$

Among them, if a Gaussian fuzzy set as shown in Figure 2.1 (a) is assumed, its membership function is a Gaussian function, m_i is the center point of the fuzzy set, and σ_i is its standard deviation.

The second layer is the rule layer. This layer multiplies the membership degree output by the fuzzification layer to obtain the activation strength of the corresponding rule, which is calculated by the following formula.

$$f_l\left(x\right) = \prod_{j=1}^{n} \mu_{A_j^l}\left(x_j\right) \tag{21}$$

The third layer is the normalization layer. This layer is used to normalize the activation strength of all fuzzy rules. The normalized fuzzy rule activation strength is calculated as follows:

$$\overline{f}_l(x) = \frac{f_l(x)}{\sum_{l=1}^{M} f_l(x)} \tag{22}$$

The fourth layer is the output layer. This layer is used to compute the output $\overline{f}_l(x)\, y_l(x)$ of each fuzzy rule. The parameters in this layer are consequent parameters.

The fifth layer is the total output layer. This layer is used to compute the final output of ANFIS, which can be expressed as:

$$\hat{y}(x) = \sum_{l=1}^{M} \overline{f}_l(x)\, y_l(x) = \frac{\sum_{l=1}^{M} f_l(x) y_l(x)}{\sum_{l=1}^{M} f_l(x)} \tag{23}$$

b. Backpropagation neural network

As the most widely used neural network today, BPNN has good approximation performance. Its structure is shown in Figure 7. It can be seen from the figure that the composition of BPNN includes: input layer, hidden layer and output layer. The learning process is forward propagation of information and back propagation of error. When the error between the actual output and the expected output is too large, the error back propagation process is entered. In addition, in the process of backpropagation, it is necessary to update the weights of the entire network through the error rate. After the weights are updated, they enter the next iterative learning. The whole process ends when the requirement to stop the iteration is reached. The requirement to stop an iteration is generally a given error rate or a given specific number of iterations.

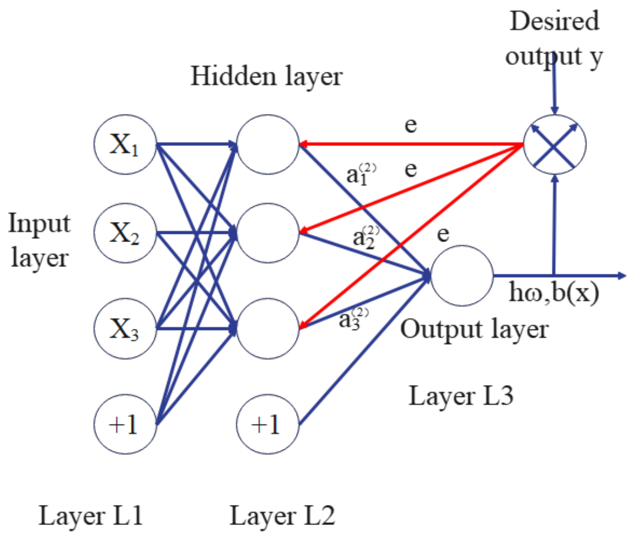


Figure 7: BPNN structure diagram.

In order to measure the performance of each prediction model, this paper will use symmetric mean absolute percentage error (SMAPE) and the mean absolute error percentage (MSPE) as measurement standards. The specific calculation formulas of SMAPE and MSPE are as follows:

$$SMAPE = \frac{1}{K}\sum_{k=1}^{K}\left|\frac{\hat{y}^k - y^k}{\hat{y}^k + y^k}\right| \tag{24}$$

$$MSPE = \frac{1}{K}\sqrt{\sum_{k=1}^{K}\left(\frac{\hat{y}^k - y^k}{y^k}\right)^2} \tag{25}$$

Among them, y^k represents the k-th actual data, $\hat{y}$ represents the predicted value of the k-th input, and K represents the number of data. For SMAPE and MSPE, the smaller the index, the better the performance of the model.

In this group of experiments, the ideological and political teaching volume data set was firstly divided to obtain a training data set and a test data set. The training dataset contains 8000 data points and the test dataset contains 1984 data points. The prediction performance of the three experimental models is verified by using the divided data sets. In this experiment, all of them have 3 inputs, that is, the ideological and political teaching amount of the three moments before the time is used to predict the power generation at the time. The interval type-2 fuzzy set selected in T2FM-DE is a Gaussian fuzzy set with uncertain standard deviation, and the number of iterations is set to 100 when using DE evolution optimization.

The parameters of the other two comparison models are set as follows:

- In the ANFIS model, Gaussian fuzzy sets are used, the number of fuzzy sets for each input variable is set to 3, and the number of iterations is set to 100.
- In the BPNN model, the number of iterations is set to 10,000, and the activation function is selected as the sigmoid function.

2.Experimental results

The experimental results of the ideological and political teaching volume prediction of the T2FM-DE model in this experiment are shown in Figure 8. For the sake of intuition, the prediction results of some data are selected for display. From the prediction experimental results in Figure 8, it can be seen that the T2FM-DE model can better track the dynamic changes in the amount of ideological and political teaching.

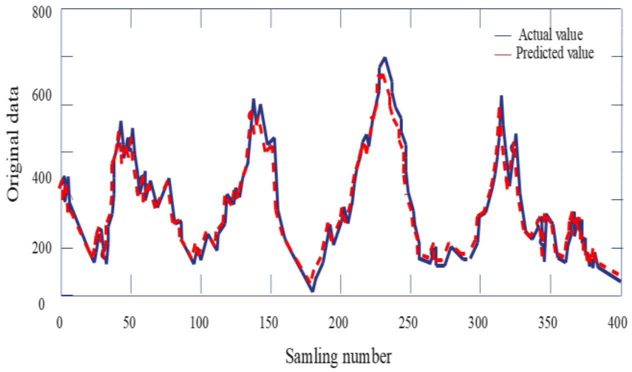


Figure 8: Prediction results of ideological and political teaching based on T2FM-DE model (partial data).

The prediction error value refers to the difference between the predicted value and the actual value. According to the distribution of the error value, it can directly reflect the prediction ability of the model. In order to compare the error distributions of the three models, the error distribution histograms of the three models are given in Figure 9. It can be clearly observed from this figure that the prediction error of the T2FM-DE model is more distributed near "0" than other models, which indicates that the effect of the T2FM-DE model is better than the comparison model ANFIS and BPNN in this experiment.

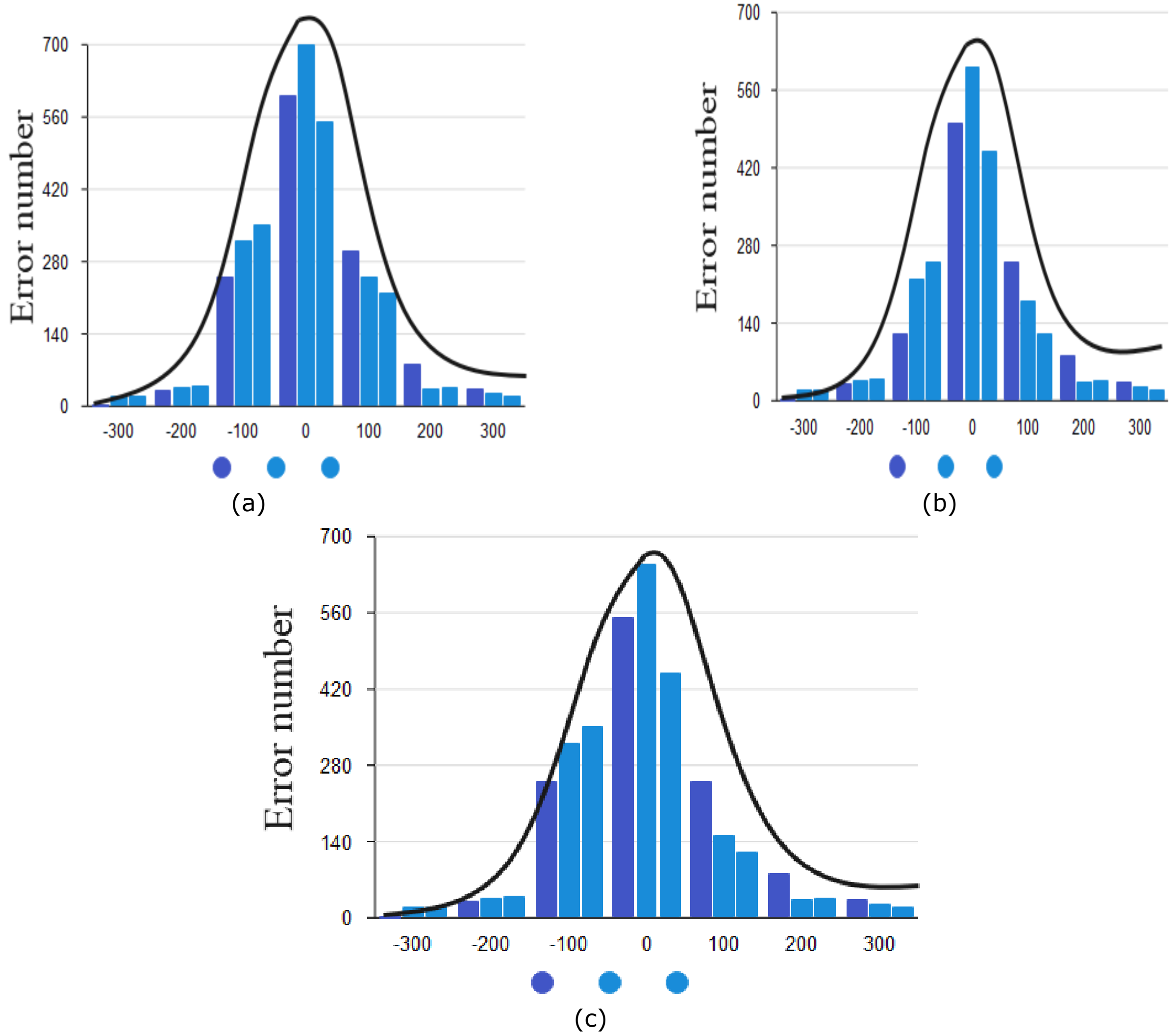
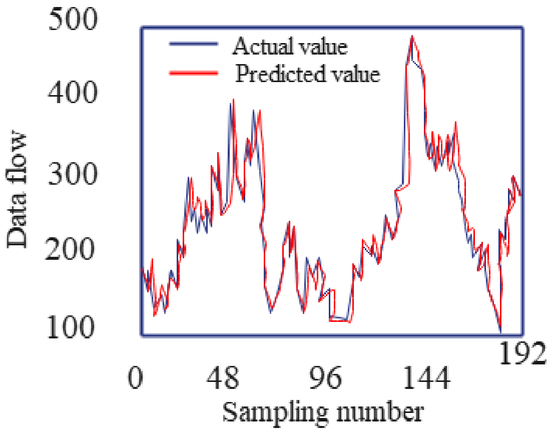
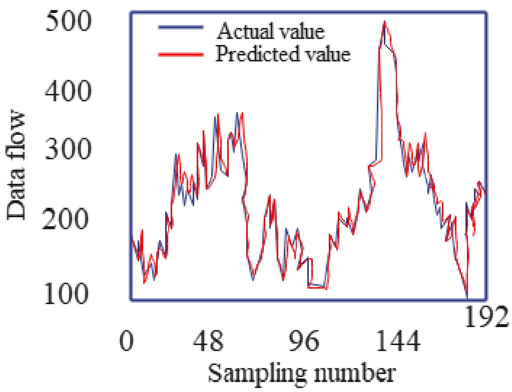


Figure 9: Histograms of error distributions for three models: (a) T2FM-DE (b) BPNN (c) ANFIS.

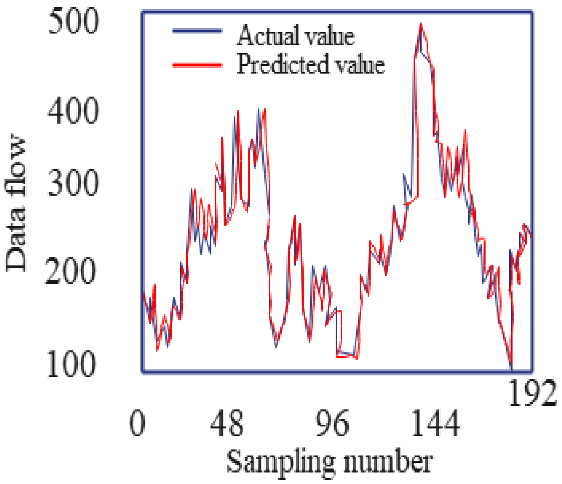
In addition, in Figure 10, the experimental results of the ideological and political teaching volume prediction of the three models on a certain day are given. Through this figure, the three models can clearly compare the situation of tracking the dynamic changes of ideological and political teaching.



(a)



(b)



(c)

Figure 10: Comparison of the results of each model in the prediction experiment of ideological and political teaching (partial data): (a) T2FM-DE (b) BPNN (c) ANFIS.

In this experiment, the experimental results of the T2FM-DE model's prediction of the amount of information in ideological and political teaching are shown in Figure 11. For the sake of intuition, a certain ten-day data forecast result is selected. As can be seen from the figure, except for a few points, the predicted value of the T2FM-DE model is very close to the actual value, and it can also well track the changes in the amount of ideological and political teaching information.

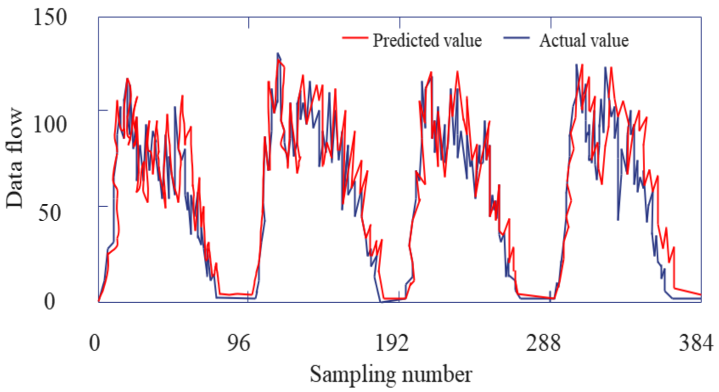
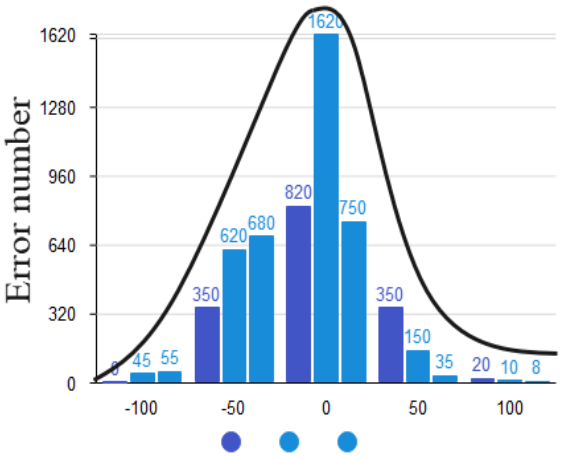


Figure 11: Prediction experiment results of some data of ideological and political teaching information (T2FM-DE).

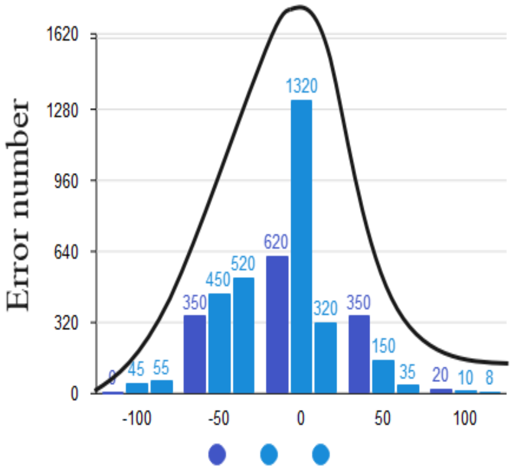
Similarly, in order to better compare the prediction effects of the three experiments, the error histograms of the three models in the ideological and political teaching information prediction experiment are also given in Figure 12. It can be observed from Figure 12 that the T2FM-DE model has a larger number of errors near "0", so that it can perform better in the prediction of information content in ideological and political teaching.

In order to compare the prediction results of the three models more clearly, some data were selected for plot comparison. The comparison of the prediction results of the three models is shown in Figure 13. From the comparison chart, it can also be observed that the T2FM-DE model can also better track the changes in ideological and political teaching information in the prediction of ideological and political teaching information.

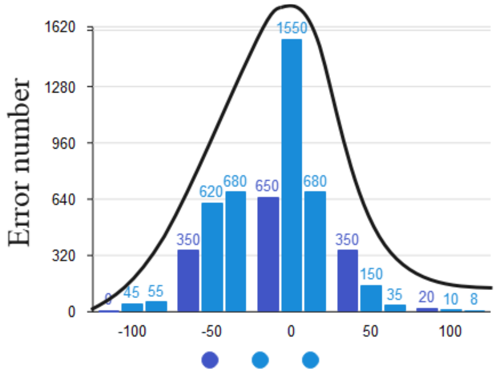
It can be seen from the above research that the expert system of ideological and political teaching based on covering set theory and artificial intelligence proposed in this paper has good effects and can effectively process various information of ideological and political teaching.



(a)

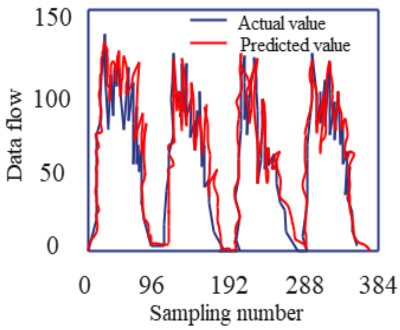


(b)

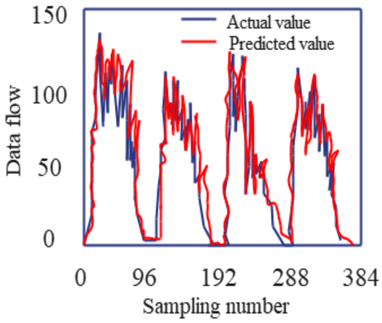


(c)

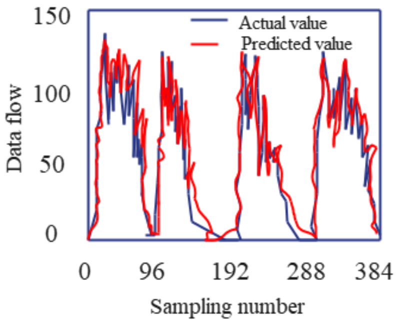
Figure 12: Error distribution histogram of three models in the prediction of ideological and political teaching information: (a) T2FM-DE (b) BPNN (c) ANFIS.



(a)



(b)



(c)

Figure 13: Prediction results of ideological and political teaching information for three models (partial data): (a) T2FM-DE (b) BPNN (c) ANFIS.

4 CONCLUSION

As an interdisciplinary frontier discipline, artificial intelligence is changing people's way of thinking and traditional concepts, improving human knowledge and human education. Moreover, the research hotspot of educational technology is closely linked with the continuous development of technology. Since the birth of artificial intelligence and artificial intelligence science, its research and application fields have been closely related to education. At the same time, the application research of artificial intelligence in education and teaching is to study the science and technology that allows computers to receive education and improve intelligence. In addition, the research results of artificial intelligence are in turn applied to the education and teaching process to promote the teaching efficiency of education, trigger the innovative thinking of educators, and generate new teaching models. This paper uses covering set theory and artificial intelligence technology to analyze ideological and political teaching, and constructs a corresponding expert system. The research shows that the expert system of ideological and political teaching based on covering set theory and artificial intelligence proposed in this paper has a good effect and can effectively process various information of ideological and political teaching.

Hong Chen, https://orcid.org/0009-0006-7550-6892

REFERENCES

[1] Alonso, J. M.: Teaching Explainable Artificial Intelligence to High School Students, International Journal of Computational Intelligence Systems, 13(1), 2020, 974-987. https://doi.org/10.2991/ijcis.d.200715.003
[2] Dilmurod, R.; Fazliddin, A.: Prospects for the Introduction of Artificial Intelligence Technologies in Higher Education, Academicia: an International Multidisciplinary Research Journal, 11(2), 2021, 929-934. https://doi.org/10.5958/2249-7137.2021.00468.7
[3] Huang, S.: Analysis of Psychological Teaching Assisted by Artificial Intelligence in Sports Training Courses, Journal of Applied Science and Engineering, 24(5), 2021, 743-748.
[4] Keating, J.; Nourbakhsh, I.: Teaching Artificial Intelligence and Humanity, Communications of the ACM, 61(2), 2018, 29-32. https://doi.org/10.1145/3104986
[5] Kim, K.; Park, Y.: A Development and Application of the Teaching And Learning Model Of Artificial Intelligence Education for Elementary Students, Journal of The Korean Association of Information Education, 21(1), 2017, 139-149. https://doi.org/10.15833/KAFEIAM.23.1.139
[6] Kong, F.: Application of Artificial Intelligence in Modern Art Teaching, International Journal of Emerging Technologies in Learning, 15(13), 2020, 238-251. https://doi.org/10.3991/ijet.v15i13.15351
[7] Li, X.: The Construction of Intelligent English Teaching Model Based on Artificial Intelligence, International Journal of Emerging Technologies in Learning, 12(12), 2017, 35-44. https://doi.org/10.3991/ijet.v12i12.7963
[8] Mianroodi, J. R.; Siboni, N. H.; Raabe, D.: Teaching Solid Mechanics to Artificial Intelligence—a Fast Solver for Heterogeneous Materials, Npj Computational Materials, 7(1), 2021, 1-10. https://doi.org/10.1038/s41524-021-00571-z
[9] Ni, D.: The Application Of Artificial Intelligence Technology in International Chinese Language Teaching, Chinese Journal of Social Science and Management, 5(1), 2021, 101-112.
[10] Pantic, M.; Zwitserloot, R.; Grootjans, R. J.: Teaching Introductory Artificial Intelligence Using A Simple Agent Framework, IEEE Transactions on Education, 48(3), 2005, 382-390. https://doi.org/10.1109/TE.2004.842906
[11] Rampton, V.; Mittelman, M.; Goldhahn, J.: Implications of Artificial Intelligence for Medical Education, The Lancet Digital Health, 2(3), 2020, e111-e112. https://doi.org/10.1016/S2589-7500(20)30023-6

[12] Ribeiro, P.; Simões, H.; Ferreira, M.: Teaching Artificial Intelligence and Logic Programming in a Competitive Environment, Informatics in Education, 8(1), 2009, 85-100. https://doi.org/10.15388/infedu.2009.06

[13] Yang, C.; Huan, S.; Yang, Y.: A Practical Teaching Mode for Colleges Supported By Artificial Intelligence, International Journal Of Emerging Technologies in Learning,15(17), 2020, 195-206. https://doi.org/10.3991/ijet.v15i17.16737

[14] Yang, R.: Artificial Intelligence-Based Strategies for Improving the Teaching Effect of Art Major Courses in Colleges, International Journal of Emerging Technologies in Learning , 15(22), 2020, 146-160. https://doi.org/10.3991/ijet.v15i22.18199

[15] Zhou, Q.: Reforms in Teaching the Course of Introduction to Artificial Intelligence, Solid State Technology, 64(2), 2021, 6331-6335.

[16] Zou, S.: Designing and Practice of a College English Teaching Platform Based on Artificial Intelligence, Journal of Computational and Theoretical Nanoscience, 14(1), 2017, 104-108. https://doi.org/10.1166/jctn.2017.6133