International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 17 (2025) pp. 1-18

https://doi.org/10.70917/ijcisim-2025-0285
Article

Intelligent Construction of the Curriculum System of Master of Arts in Music in the Framework of Great Civic and Political Education Based on Knowledge Mapping

Yu Xia [1] and Fumei Xu [2,*]

[1] Nanchang Hangkong University, Nanchang, Jiangxi, 330063, China
[2] Jiangxi Normal University, Nanchang, Jiangxi, 330063, China; xy_tulip@163.com

Abstract: At present, the use of music teaching on college students' nurturing in the context of big ideology and politics in colleges and universities is still deficient, and most professional teachers ignore the nurturing function in music courses. Therefore, this paper adopts the graphical semantic network in the knowledge graph to represent the concepts and the interrelationships between the concepts, to be an in-depth study and research on the extraction of entity relations. Through the neural network model as well as the pre-trained language model, the text is vectorized and represented. The input layer, feature extraction layer and feature reuse layer of the model are constructed sequentially, and the loss function is introduced to improve the relationship extraction model based on entity information and feature reuse, and the model is applied to the education of music professional courses under the framework of the Big Ideology and Politics. The performance of the model is evaluated, and the average values of P value, R, and F1 of the model are 0.932, 0.945, and 0.939, respectively, and the performance of the model is good. After a period of time after the music classroom integrating the elements of ideology and politics, the students' cultural self-confidence increased by as much as 9.38%, and the degree of understanding of core values increased by 6.14%, which indicates that the construction of music professional curriculum under the framework of the Great Ideology and Politics Education has achieved good results.

Keywords: knowledge graph; entity-relationship extraction; neural network; pre-trained language model; music professional curriculum

1. Introduction

In the new era, master of arts students are the hope of the country, the future of the nation, in the case of economic globalization, the world multipolarization, especially the informationization of the society, in the face of the values of all sides of the intentional or unintentional influence, to buckle up life, if you can't use the right way to stick to the ideological position, the master of arts students will lose the national “root” and “soul” [1-4]. In this rapidly changing Internet era, although the theory of the master of arts students of the ideological education is extremely important, but in practice the results are not particularly significant, the ideological and political class is not high enough, how to play the master of arts students in the adherence to the road of socialism with Chinese characteristics, theories, systems and culture of self-confidence, the integration of the music professional curriculum system and the ideological and political education is particularly important [5-8].

Civic and political education courses are designed to cultivate college students' correct worldview, outlook on life and values, as well as to learn the basic Marxist literacy to set up lofty ideals and beliefs, and then become qualified socialist builders and laborers. And the music professional curriculum system for master of arts students includes professional courses to cultivate specialized musical talents and courses to cultivate basic music general knowledge [9-10]. Although music education and ideological education belong to different disciplines, music education contains rich materials and resources for ideological education, and the combination of the two can form a synergy to play a common effect of educating people [11-13]; music education is centered on aesthetics, which belongs to the category of aesthetic education, and it can enhance the attractiveness and infectious force of ideological education [14-15].

At the same time, the music professional curriculum system of master of arts can play the role of guarding the canal and planting the field on the basis of implementing the fundamental task of moral education in the Civic and Political Classes; therefore, it is imperative to play the efficacy of the music professional curriculum system in the aspect of Civic and Political Education [16]. The application of knowledge graph mapping can realize the intelligent construction of the music professional curriculum system of master of arts in the civic and political education, which not only reduces the pressure of teachers, but also improves the learning efficiency of students [17-19]. Knowledge graph mapping can structurally associate resources such as music professional curriculum system, civic education, cultivation program and syllabus, and form a unified knowledge representation of fragmented material disciplinary knowledge, based on its knowledge structure and intrinsic characteristics, to provide students with rich personalized services, which helps students to understand and apply knowledge at a deeper level [20-24].

In this paper, we focus on the in-depth study of the most critical part of knowledge graph construction, i.e., the extraction of entity relationships. Cultural vectors are represented by word vectors and two mainstream pre-trained models, BERT and RoBERTa algorithms. The entity information and feature reuse relationship extraction model is divided into three-layer structure of input layer, feature extraction layer, feature extraction layer as well as feature reuse layer, and a loss function is added to build out the basic architecture of the relationship extraction model, which is applied to the music professional courses to complete the internal knowledge fusion of the music professional knowledge mapping. The realization path of the music curriculum's ideology under the framework of big ideology education is proposed. Analyze the results of entity relationship extraction dataset, while the knowledge graph represents the relationship between entities and entities, and realize the relationship visualization display of entities. Evaluate the learning effect of music course Civics under the framework of Big Civics Education through comparative experiments.

2. Technology related to the construction of knowledge map for music professional courses

2.1. Knowledge mapping

The construction and updating process of knowledge graph is shown in Figure 1 [25]. Knowledge map adopts graphical semantic network to represent concepts and interrelationships between concepts, and its basic unit is “entity-relationship-entity” triad, as well as entities and their related attribute-value pairs, and the entities are linked to each other through the relationship, constituting a net-like knowledge structure. The basic unit is the “entity-relationship-entity” triad, and entities and their related attribute-value pairs.

Entity-relationship extraction is the most basic and critical part of knowledge graph construction, because the accuracy of its extraction has a great impact on the later tasks, this paper mainly focuses on entity-relationship extraction for in-depth study and research, and the following will introduce various methods and techniques mainly used in the field of entity-relationship extraction.

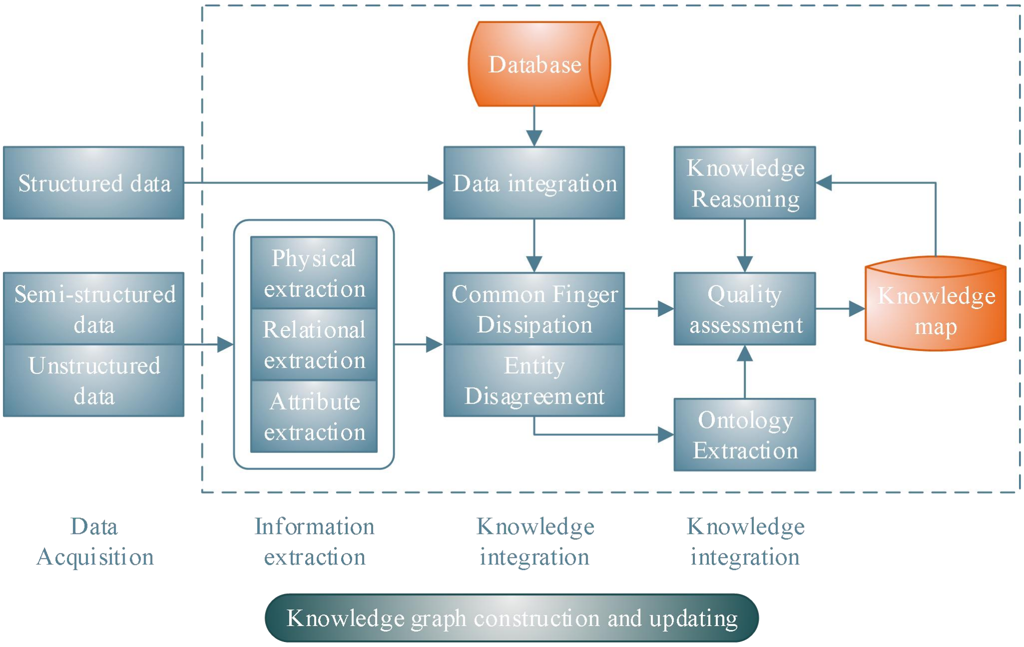

Figure 1. Knowledge graph construction and update process.

2.2. Text vectorization

2.2.1. Word Vector

In NLP, word vectors are an important step in realizing various natural language processing tasks by transforming input natural language information into digital information that can be understood by computers. There are two main types of representations for word vectors, discretized representation and distributed representation.

Among the discretized representation models, one-hot is the most commonly used method, also called one-hot coding, which is the earliest coding method. The method refers to the conversion of words in a sentence into the form of a numerical matrix. One-hot coding treats each word in a sentence as a unit, and each unit corresponds to a state value of 1, while the others are 0, which results in an N × N numerical matrix. For example, in the sentence "cat eat fish", cat is coded as [1,0,0], eat is coded as [0,1,0], and fish is coded as [0,0,1]. The advantages of one-hot coding are that it is simple, easy to implement and fast, but there are also many disadvantages, such as: (1) When the number of words in the word list is very large, the dimension of the word vector will be very large, which may eventually lead to dimensional explosion. (2) There are a large number of values 0 in the word vectors obtained by this encoding method, and it is more difficult to compute the correlation between words, which can lead to the sparsity problem. (3) Since the cosine similarity calculation of the unique heat encoding of any two non-repeated words must be 0, the words are independent of each other and cannot reflect the correlation between words.

The text word sequence is represented as $w=\left(w_1, w_2, \ldots, w_n\right)$, with a word list size of V and a sliding window size of m, W_t being the one-hot vectors corresponding to the word w_t, assuming that the context words are independent.

Skip-gram method is to use the center word to predict the word in context, the core idea of the method is that for a center word, the model has to learn the word vectors of its context, each word will be influenced by its surrounding words when it becomes a center word, and as the context changes, the current word will change, the advantage of this method is that it can obtain a more accurate representation of the word vectors, and the disadvantage is that the training time is relatively long. The objective function is shown in the following equation:

$$L=\sum\nolimits_{w \in text} \log p(Context(w) \mid w) \tag{1}$$

The CBOW method is based on the opposite idea, which predicts the probability of the current word based on the context in the given sentence, which reduces the training time compared to the Skip-gram method, with the objective function as shown in the following equation [26]:

$$L=\sum_{w\in text}\log p(w\mid Context(w))\tag{2}$$

Compared with the one-hot coding method, the Word2Vec coding method obtains word vectors with lower dimensions, which can better express the semantic features of words, is computationally simple and occupies less space. However, the Word2Vec representation only utilizes the information of the word in a certain regional context, and does not connect the word with other distant words, ignoring the global statistical information of the corpus.

Some scholars proposed the GloVe model in 2014, which is an unsupervised word vector training model based on matrix decomposition. The model builds a word co-occurrence matrix based on the corpus, by analyzing the approximate relationship between the co-occurrence matrix and the word vectors, and then constructs the loss function. The value of i row j column of the matrix is the logarithm of the number of co-occurrences t_{ij} of the words w_i and w_j, and t_i denotes the total number of times of all the word occurrences in the context of w_i. p_{ij} denotes the probability that w_j occurs in the context of w_i, with the formula $p_{ij}=p\left(w_j\mid w_i\right)=t_{ij}/t_i$. p_{ik}/p_{jk} can be used to represent the correlation between the word w_k and w_i, w_j, when the value of p_{ik}/p_{jk} is greater than 1 it means that w_k is more correlated with w_i and vice versa it means that w_k is more correlated with w_j. The formula for its objective loss function is shown below:

$$F=\sum_{i,j=1}^{|v|}f\left(t_{ij}\right)\left(p_i^Tq_j+b_i+b_j-\log\left(t_{ij}\right)\right)^2\tag{3}$$

where b is the bias, $|v|$ denotes the word list size, and $f\left(t_{ij}\right)$ denotes the weight change function. The value of $f\left(t_{ij}\right)$ varies with the number of word co-occurrences, the higher the frequency of the word the higher the weight, and after a certain point the value is unchanged.

2.2.2. Pre-training language models

Pre-training language modeling uses pre-training in a large-scale corpus, only the model parameters of the pre-trained network are kept, instead of assigning a separate vector representation to each word, and when the whole network is trained, the first half of the network is initialized by the previously saved parameters, which is called "pre-training+fine-tuning", so that This method is called "pre-training + fine-tuning", so that the word vectors can be dynamically represented in different contexts. In the following, we will introduce the current mainstream pre-trained language models: BERT and RoBERTa.

(1) BERT

BERT is a bi-directional coded language representation model released by Google.BERT achieved state-of-the-art results in 11 natural language processing tasks when it was published, and it has now become one of the mainstream pre-training models [27]. The innovation of BERT has two main aspects.Firstly, during the training process of the language model, 15% words are randomly sampled from a piece of text, and these words are masked. These words are then predicted based on the context, which is similar to the CBOW method in Word2Vec, thus realizing the efficient use of text information in both directions and avoiding label leakage. Secondly, on the basis of vocabulary prediction, sentence-to-sentence association prediction is added, i.e., judging the association between the next sentence and the previous sentence. By this method, more semantic features between sentences can be obtained and the applicability of the method to various tasks is enhanced.

(2) RoBERTa

Some scholars proposed RoBERTa pre-training language model in 2019, which modifies and tunes the BERT model in terms of the model's masking strategy and data.The improvement of RoBERTa model on the BERT model is specifically reflected in the following three aspects:(1) RoBERTa adopts the dynamic Masking mechanism to pre-train the word embedding vectors, and first of all, the pre-training data is 10 copies are made and 15% of the Token is randomly selected for masking for each copy of data, so for the same utterance, there will be 10 different masking ways. Then in the training of N epochs, the tokens that are masked for each sequence will keep changing, and the model will gradually adapt to different masking strategies, so as to learn different linguistic representations, and this dynamic masking mechanism will be more flexible compared to the fixed masking in BERT. (2) RoBERTa has a larger batch size, RoBERTa's improvement to the BERT model also includes expanding the batch size. RoBERTa increased the original batch size of 256 to a batch size of 8K to further improve the learning ability of the model, and a larger dataset was used for training data to find the optimal hyperparameters. (3) The Next SentencePrediction (NSP) task was removed because the use of the Next Sentence Prediction task would limit the input length of the model to some extent, making the model less sensitive to information over long distances.

2.3. Neural Networks

Fig. 2 shows the neuron computation process, which details a simple neural network structure, where each small circle represents a neuron and the arrows represent the weights. Where $X=(x_1,x_2,\ldots,x_n)$ is the input on each neuron, $W=(w_1,w_2,\ldots,w_n)$ is the trainable parameter connecting the neurons, b is the bias term, SUM is the summation function, f is the activation function, which can be taken as Softmax, Sigmoid, Tanh, ReLU, etc., and α represents the activation value. Thus the neuron activation process is shown in equation (4):

$$\alpha = f\left(W^T X + b\right) \tag{4}$$

Multiple neurons can be composed into different artificial neural networks by stacking them in different layers. The input of this fully connected feed-forward neural network is a 3-dimensional vector, which is computed by two intermediate layers, the first one with 5 neurons and the second one with 6 neurons, and the output is a 1-dimensional vector.

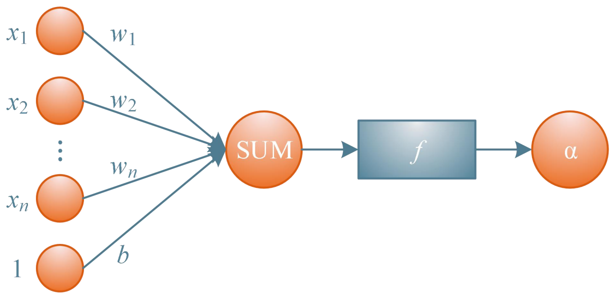

Figure 2. The computational process of neurons.

2.3.1. Convolutional Neural Networks

CNN uses a unique convolutional operation to extract features, which gives it a more powerful spatial feature representation than recurrent neural networks, and at the same time, CNN is more friendly to parallel computation than recurrent neural networks with their temporal computation method. Unlike traditional neural networks, convolutional neural networks contain new concepts such as pooling layers in addition to computational layers.

2.3.2. Recurrent Neural Networks

Figure 3 shows the schematic of RNN network unfolding, RNN is a network model more suitable for sequence data. With weight sharing, RNN can be extended to different styles of data samples. Similarly, RNN uses similar modules for the whole sequence and can generalize a longer sequence to get the desired result.

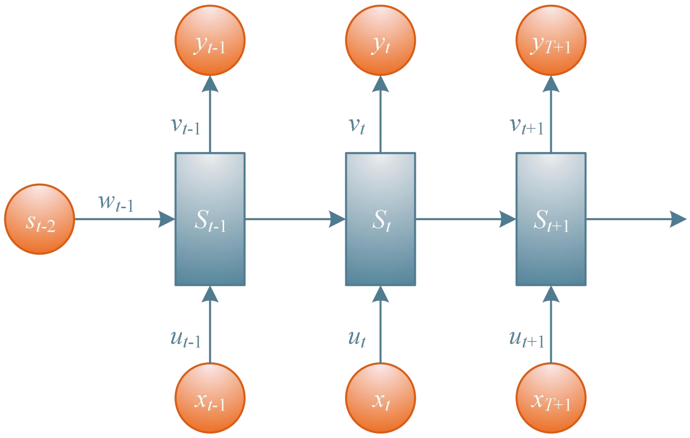

Figure 3 RNN network expansion illustration

The $t-1, t, t+1$ denote different time series, and s_t denotes the hidden layer state at the moment of t, which is computed as shown in Equations (5) and (6):

$$S_t = f\left(U_t X_t + W_t S_{t-1} + b_1\right) \tag{5}$$

$$y_t = f\left(V_t S_t + b_2\right) \tag{6}$$

The formula for each LSTM cell is shown below. The forgetting gate f_t controls how much of the historical state is forgotten through an activation function. The input gate i_t controls how much each cell is updated, and the output gate o_t controls whether or not the internal memory state is input to the next moment.The computation of the LSTM is shown in equations (7) to (11):

$$f_t = \sigma\left(W_f \cdot \left[h_{t-1}, x_t\right] + b_f\right) \tag{7}$$

$$i_t = \sigma\left(W_i \cdot \left[h_{t-1}, x_t\right] + b_i\right) \tag{8}$$

$$c_t = f_t c_{t-1} + i_t \tanh\left(W_f \cdot \left[h_{t-1}, x_t\right] + b_c\right) \tag{9}$$

$$o_t = \sigma\left(W_o \cdot \left[h_{t-1}, x_t\right] + b_o\right) \tag{10}$$

$$h_t = o_t \cdot \tanh C_t \tag{11}$$

where W_i, W_f and W_o are are the weight matrices to be learned, b_i, b_f & b_o are the bias terms, which represent the parameters of the input gates, oblivion gates and output gates respectively. σ & $\tanh$ denote the activation functions, "$\cdot$" denotes the elementary product of the matrix, x_t is the input of the LSTM unit, and h_t is the output of the hidden layer at the current moment.

GRU is a variant of LSTM that combines the forgetting and input gates into a single update gate. Compared with the standard LSTM, the GRU model is simpler and its specific formulas are shown in Eq. (12) to Eq. (15):

$$r_t = \sigma\left(W_r \cdot \left[h_{t-1}, x_t\right]\right) \tag{12}$$

$$z_t = \sigma\left(W_z \cdot \left[h_{t-1}, x_t\right]\right) \tag{13}$$

$$h_t' = \tanh\left(W \cdot \left[r_t \cdot h_{t-1}, x_t\right)\right. \tag{14}$$

$$h_t = z_t \cdot h_t' + \left(1 - z_t\right) \cdot h_{t-1} \tag{15}$$

where σ is the sigmoid activation function, "·" denotes the elementary product of matrices, W_r, W_z and W are the weight matrices to be learned by the GRU network during the training process, x_t is the input to the GRU unit, and h_t is the output of the hidden layer at the current moment.

2.3.3. Graph Attention Neural Networks

Graph Attention Network (GAT) is a neural network model for graph data. Let the feature vector of any node v_i in layer 1 be $h_i \in R^{d^{(l)}}$, where $d^{(l)}$ denotes the length of the feature vector of the node in layer 1. After the aggregation operation through the attention mechanism, a new feature vector $h_i' \in R^{d^{(l+1)}}$ is output for each node, where $d^{(l+1)}$ denotes the length of the feature vector output by the node in the $l+1$ th layer.

For the current central node v_i and a neighbor node v_j, the weight coefficients e_{ij} can be expressed in Equation (16):

$$e_{ij} = a\left(W\vec{h}_i, W\vec{h}_j\right) \tag{16}$$

where $W \in R^{d^{(l+1)} \times d^{(l)}}$ is the weight parameter of the node's feature transformation for the layer, and a is the function to calculate the correlation between two nodes. Attention networks can be designed in several ways, here a single fully connected layer is used with the formula shown in (17):

$$e_{ij} = Leaky\,\mathrm{Re}\,LU\left(\vec{a}^T\left[W\vec{h}_i \| W\vec{h}_j\right]\right) \tag{17}$$

where $\|$ represents the splicing operation, the weight parameter $a \in R^{2d^{(l+1)}}$, and the activation function uses LeakyReLU. in order to better assign the weights, it is necessary to uniformly normalize the correlation between the computed current central node and all the neighboring nodes, where the form of softmax normalization is used, and the specific computation process is shown in Equation (18) shows:

$$\alpha_{ij} = soft\max_j\left(e_{ij}\right) = \frac{\exp\left(e_{ij}\right)}{\sum_{k \in N_i} \exp\left(e_{ik}\right)} \tag{18}$$

where α_{ij} is the weight coefficient, by using the above formula, it can be ensured that the sum of the weight coefficients of all the neighboring nodes of the current central node is 1. The complete weight coefficients computation process is given in the following equation (19):

$$\alpha_{ij} = \frac{\exp\left(Leaky\,\mathrm{Re}\,LU\left(\vec{a}^T\left[W\vec{h}_i \| W\vec{h}_j\right]\right)\right)}{\sum_{k \in N_i} \exp\left(Leaky\,\mathrm{Re}\,LU\left(\vec{a}^T\left[W\vec{h}_i \| W\vec{h}_k\right]\right)\right)} \tag{19}$$

After completing the computation of the above weight coefficients, the new feature vector $\vec{h}_i$ of node v_i is computed according to the weighted attention mechanism as shown in equation (20) below:

$$\vec{h}_i = \sigma\left(\sum_{j \in N_i} \alpha_{ij} W\vec{h}_j\right) \tag{20}$$

To enhance the expressiveness of the attention layer, K independent attention mechanisms can be used and their outputs can be spliced together as shown in Equation (21):

$$\vec{h'}_i = \|_{k=1}^{k} \sigma\left(\sum_{j \in N_i} \alpha_{ij}^k W^k \vec{h}_j\right) \tag{21}$$

where $\|$ represents the splicing operation, α_{ij}^k is the weight coefficient of the kth group of attention mechanisms, and W^k is the weight coefficient of the kth group of attention mechanisms.

When performing multi-head attention at the last layer of the network, averaging is usually used instead of cascading, as shown in (22):

$$\vec{h'}_i = \sigma\left(\frac{1}{K}\sum_{k=1}^{K}\sum_{j\in N_i}\alpha_{ij}^{k}W^{k}\vec{h}_j\right) \tag{22}$$

3. Relationship extraction model based on entity information and feature reuse

3.1. Model Architecture and Rationale

3.1.1. Input layer

In the input layer of the model, the vector matrix I^i and the vector matrix M^a are the first to be fed into the language model Roberta, which outputs a hidden layer H. Then the vector matrix M^{e1}, the vector matrix M^{e2}, and the vector matrix M^p are fed into the model to multiply with the hidden layer H , from which information can be extracted about entity pairs with respect to the whole sentence as well as information about cue parts with respect to the whole sentence.

3.1.2. Feature extraction layer

In the feature extraction layer, the hidden layer H is used as the output of Roberta, and the feature vector H_0 contains the information of the whole sentence, and the feature vector H_1 is noted as the feature vector of the sentence information $F^c = \left\{y_1^c, y_2^c, \ldots, y_{L_H}^c\right\}, L_H$ represents the length of the H output of the hidden layer H. The feature vector H_0 is then used as the input to the BIGRU, so that the information of the whole sentence can be extracted through the BIGRU again, thus strengthening the features of the input. For each element in the BIGRU input sequence, each layer is computed as shown in Eqs. (23) to (26):

$$r_t = \sigma\left(W_r\left[h_{t-1}, H_0\right]\right) \tag{23}$$

$$z_t = \sigma\left(W_z\left[h_{t-1}, H_0\right]\right) \tag{24}$$

$$h_t' = \tanh\left(W_{h'}\left[r_t \odot h_{t-1}, H_0\right)\right. \tag{25}$$

$$h_t = z_t \odot h_t + \left(1 - z_t\right) \odot h_t' \tag{26}$$

where σ is the sigmoid activation function, $\odot$ is the term-by-term product, and W_r, W_z and $W_{h'}$ are the parameters of the GRU network. h_t is used as the output of BIGRU, and h_t is denoted as the feature vector $\left\{y_1^b, y_2^b, \ldots, y_{L_H}^b\right\}$ extracted by BIGRU.

Multiplying the vector matrix M^{e1}, the vector matrix M^{e2}, and the vector matrix M^p with H_0, respectively, assume that the vector of H_0 is denoted by $\left\{\beta_1, \ldots, \beta_{L_H}\right\}$, which is given as (27) is shown:

$$M_{sum} = \sum_{i=1}^{L_H} x_i\beta_i = x_1\beta_1^T + x_2\beta_2 + \cdots + x_{L_H}\beta_{L_H} \tag{27}$$

Dividing the computed vector matrix M_{sum} with the length of the masked portion of each vector matrix gives the final average vector, and the result obtained is denoted as $F^{e1} = \left\{y_1^{e1}, y_2^{e1}, \ldots, y_{L_H}^{e1}\right\}, F^{e2} = \left\{y_1^{e2}, y_2^{e2}, \ldots, y_{L_H}^{e2}\right\}$, $F^p = \left\{y_1^p, y_2^p, \ldots, y_{L_H}^p\right\}$. That is, the five feature vectors $F^c, F^b, F^{e1}, F^{e2}, F^p$ obtained in the feature acquisition layer are used as inputs to the next layer.

3.1.3 Feature Reuse Layer

In the feature reuse layer, five features $F^c, F^b, F^{e1}, F^{e2}, F^p$ are taken as inputs, and the features are processed, first by regularizing the input feature vectors. Then to keep the model from overfitting, the model is made to go through a dropout layer again, dropping some neurons randomly, and finally, after passing through the function Tanh, the output is made to be dimensionalized using a linear model. The output obtained from each feature is then spliced together to get O_1. Putting F^c and F^b undergo the same operation again, respectively, to get O_2 and O_3 from the linear layer output. Finally putting F^{e1}, F^{e2} and F^p also gives O_4. Where the specific linear operation formula is shown in (28):

$$O_i = w_i \tanh(F) + b_i \tag{28}$$

Taking the obtained outputs O_1, O_2, O_3, O_4 and passing them through SoftMax to get O_1', O_2', O_3', O_4', as shown in (29):

$$P(y \mid x) = \frac{e^{h(x,y_i)}}{\sum_{j=1}^{n} e^{h(x,y_i)}} \tag{29}$$

Finally, the idea of integrated learning soft voting is used to average the probability of each classification result in each output by adding the probabilities of each classification result to find the result Output of the final relationship classification, and the calculation process is shown in (30):

$$Output = 1/4\left(O_1' + O_2' + O_3' + O_4'\right) \tag{30}$$

3.1.4. Loss function

In this paper, cross entropy is added to the loss function. The final loss function consists of cross-entropy, KL scatter, and negative log-likelihood loss. First, let each batch of data pass through the forward neural network twice before and after, due to the stochastic nature of Dropout, even though it passes through the same model both times, it is different in the forward pass. p1 left path is dropped with a different unit to output distribution and p2 right path is dropped with a different unit to output distribution. For this reason the difference between the two distributions is described in terms of the KL scatter notated as L_{kl}^i, as shown in Eq. (31):

$$L_{kl}^i = \frac{1}{2}\Big(D_{kl}\left(P_1^i\left(y_i \mid x_i\right) \| \, P_2^i\left(y_i \mid x_i\right)\right) + D_{kl}\left(P_2^i\left(y_i \mid x_i\right) \| \, P_1^i\left(y_i \mid x_i\right)\right)\Big) \tag{31}$$

Then the cross-entropy L_{CE}^i and the negative log-likelihood loss L_{NLl}^i are used to find out the one difference value between the two results respectively The computation procedure is shown in Eqs. (32) and (33):

$$L_{CE}^i = -P_1^i\left(y_i \mid x_i\right)\log P_1^i\left(y_i \mid x_i\right) - P_2^i\left(y_i \mid x_i\right)\log P_2^i\left(y_i \mid x_i\right) \tag{32}$$

$$L_{NLl}^i = -\log P_1^i\left(y_i \mid x_i\right) - \log P_2^i\left(y_i \mid x_i\right) \tag{33}$$

Finally, the losses $L_{kl}^i, L_{CE}^i, L_{NLl}^i$ are summed up to obtain a final one loss function L_{loss}^i as shown in Equation (34):

$$L_{loss}^i = (1-\alpha)\left(L_{CE}^i + L_{NLL}^i\right) + \alpha L_{kl}^i \tag{34}$$

3.2. Curriculum system of music majors based on the framework of the Great Civic Policy

3.2.1. Music knowledge acquisition

Knowledge acquisition is the process of acquiring and extracting the required knowledge from multiple data sources. The basic task of music knowledge acquisition is to acquire music domain knowledge and establish a sound, complete and effective music knowledge map to meet the knowledge needs of the music domain.

(1) Data Sources

Data sources for knowledge acquisition usually include professional literature and data in relevant databases and knowledge bases. When constructing a music knowledge system, the main data sources for knowledge acquisition include music literature (music research literature, chants, scripts), music works of various representation types (audio, video, sheet music), and music information recorded in databases, knowledge bases, web pages, and so on.

(2) Knowledge Extraction

Knowledge extraction refers to the process of extracting the knowledge embedded in the data source by means of identification and understanding. Music content data is a unique knowledge extraction object in the music domain. The three main types of music content data, namely audio, sheet music, and lyrics, correspond to different knowledge extraction methods.

3.2.2. Integration of musical knowledge

Knowledge fusion is a high-level knowledge organization that enables knowledge from different knowledge sources to achieve heterogeneous data integration under the same framework specification, and the steps to achieve this include ontology construction, entity alignment, entity linking, and ultimately, the fusion of data, information, methods, experiences, and ideas to form a high-quality knowledge graph.

(1) Ontology Construction

Ontology construction is a key step in music knowledge fusion, which requires the completion of abstract modeling and structured definition of music domain knowledge. At present, the knowledge information involved in music knowledge mapping can be divided into three major categories: descriptive information of music resources or works, music event information, and music content recording and analysis information. The first two types of information are commonly found in textual knowledge graphs, while the third type of information needs to be obtained after analyzing music content data, which is a multimodal information range, choosing to reuse the ontology or extend the definition of a new ontology.

(2) Entity Alignment

Entity alignment is the process of corresponding entities in different data sources to the same entity to which they jointly point. An important task of knowledge fusion in the music domain is to complete the entity alignment of core entities such as musicians, musical works and musical instruments.

(3) Entity Linking

Ontology construction and entity alignment complete the internal knowledge fusion of the knowledge graph, while entity linking is to link ambiguous entities to external authoritative knowledge bases to realize the knowledge fusion between the knowledge graph and external data sources. The five-star criterion defined by Linked Open Data Cloud (LOD) takes the amount of linked data generated by the knowledge graph with other open knowledge bases as an important criterion for evaluating the quality of the dataset.

4. The construction of music program's ideology and politics under the framework of big ideology and politics education

4.1. Constructing professional Civics courses and paving the way for curriculum Civics

Higher music colleges and universities should establish a professional Civics and Politics course guidance system, and strengthen the linkage between professional course knowledge and course Civics and Politics theories. In the process of developing the course guidance mechanism, the teaching activities are strictly controlled from multiple perspectives, such as talents, lectures, and educational objectives. First of all, professional teachers with excellent professionalism and experience in Civic and Political Education are selected to carry out Civic and Political Education activities, so as to ensure that students can learn the correct theory of Civic and Political Education. Under the pattern of “big ideology and politics”, it is necessary to rely on the power of students to evaluate the ability of teachers' ideology and politics education, to realize the principle of “if there is a change, there is no correction”, and to continuously improve the professional quality of teachers. Secondly, in combination with the requirements of modern ideological education, new educational materials are introduced into the ideological education courses, focusing on Marxism, theories of the new era, socialist core values and other materials to complete the work of ideological and political education.

4.2. Establishing the concept of ideological education and realizing diversified education

Higher music colleges and universities should pay close attention to the ideology of the curriculum in teaching activities, on the one hand, carry out teaching skills “training” competitions, emphasize the important value of the ideology of the curriculum, and establish a more mature and perfect model of guidance for the cultivation of human beings. On the other hand, around the “big ideology and politics” pattern, with the basic goal of cultivating moral integrity, to build a long-term and scientific mechanism for exploring the ideology and politics of the curriculum. In the process of actively completing the music teaching task, professional music teachers should be guided by concepts and actions, centering on the relevant requirements of the “big ideology and politics” pattern, actively exploring scientific methods of integrating curriculum ideology and politics into teaching activities, introducing ideological values and ideological education materials into the whole teaching process, and helping all members to establish new teaching concepts. The program will help all members to establish a new concept of teaching and learning.

4.3. Accelerating the reform of the ideology of the curriculum and improving the quality of human training

Higher music colleges and universities must grasp the fundamental task of cultivating moral integrity, and while accelerating the speed of integration between the curriculum and professional teaching courses, readjust the music teaching mode to realize the organic combination between the curriculum and the cultivation of moral integrity, so that the curriculum and the cultivation of moral integrity can support the professional music teaching activities.

4.4. Diversified integration of curriculum thinking and changing teaching methods

Under the pattern of “big ideology and politics”, the concept of curriculum ideology and politics can be implemented with the help of diversified music activities, re-adjusting the mode of ideological and political education, and strengthening the connection between students and ideological and political education. Music colleges and universities can organize music activities, take the ideological and political thought as the line of action, and link the ideological and political theory with professional music knowledge.

5. Analysis of music relationship extraction results and the effect of civic education

5.1. Entity-relationship extraction dataset

After getting the preprocessed text, this paper needs to store the dataset in accordance with a certain format, specifically entity one, entity two, relationship category and the text containing entity pairs. Set the maximum length of the text length of 128, the length of the text is lower than the maximum value of the text length through the way of zero filling to ensure the consistency of the text length, for the text length of the text length over the maximum value, through the excision of redundant parts to ensure the consistency of the text length. For multiple pairs of entities and relationships in the text, each pair of entities and relationships is categorized when constructing the dataset to minimize the experimental error.

In order to improve the model training and subsequent classification performance, the dataset needs to be processed before inputting into the model, combining the music entity pairs as well as the original text, connecting the entity pairs through $, and replacing the entities involved in the original text with the special character #. For example, “Jay Chou's ### song ## healed many people's hearts in that summer”.

As for the construction of the relationship extraction dataset, a total of 5,565 textual data were collated and divided into training and test sets in the ratio of 8:2. The statistics of the number of relationship categories are shown in Fig. 4, with a maximum number of 1,560 entity relationship categories for Artist-Songs, which decreases from left to right, and a minimum number of 100 entity relationships for the Final Arrangement.

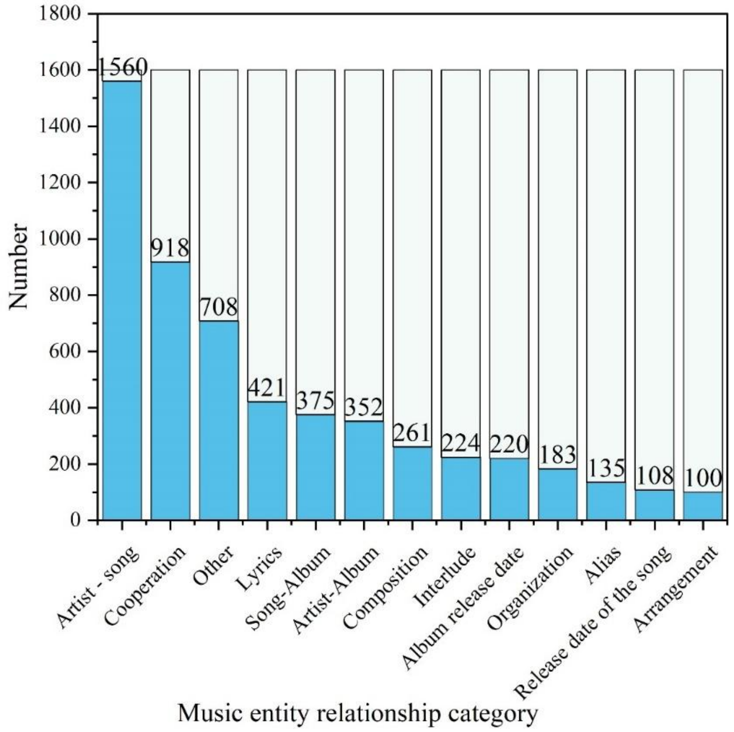

Figure 4. Statistics of the number of music entity relationship categories.

Figure 5 shows the text length frequency statistics of the relationally extracted dataset. In this section, the text length as well as the frequency number of occurrences are also counted, and similar to the named entity recognition task, the length of the text is mostly clustered between 10 and 80, and the frequency number ranges from 1 to 183.

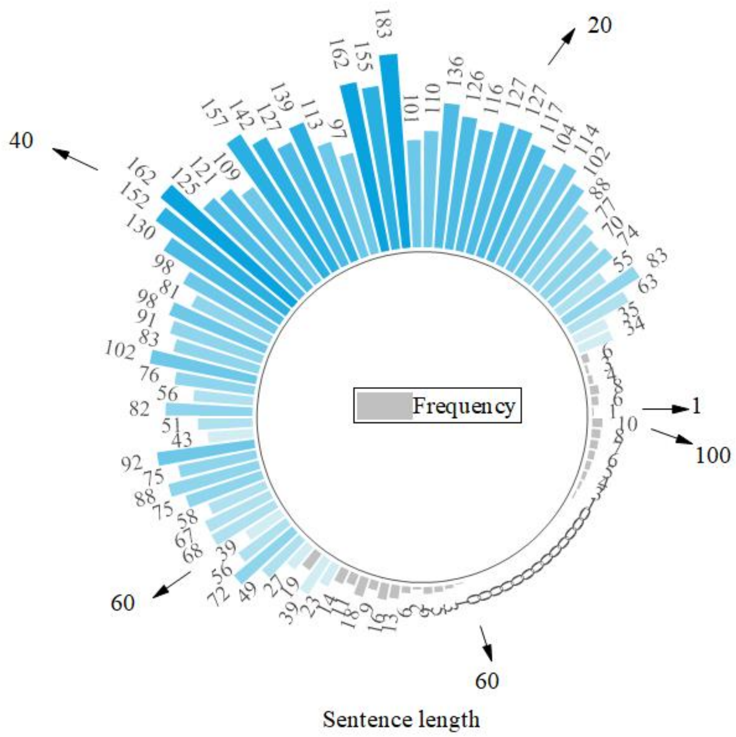

Figure 5. The relationship extracts data set text length frequency statistics.

For a multicategorization problem, it is usually processed by transforming it into multiple dichotomous problems, i.e., by comprehensively examining the evaluation metrics on n dichotomous confusion matrices. Therefore, in this paper, the macro mean (Macro) is chosen for statistics.

The result of the experiment. The experimental results of the relationship categories among different entities are shown in Figure 6. For the predefined different entity relationship categories, the performance of text classification varies. It is not difficult to find that the experimental effects of the lyric and composition relationship categories are relatively poor, which is related to the fact that there is less data of this type of relationship category in the dataset. In addition, since the same relationship category has diverse expressions, it will to a certain extent affect the classification results of the model. For example, for lyrics, the text has various expressions, such as "write down", "fill in words", "compose words", "lyricist", and so on. When different expression texts of this type of relationship category are added to the dataset, the performance of the model is improved to a certain extent. Overall, the average values of P-value, R and F1 of the model are 0.932, 0.945 and 0.939 respectively, indicating that the performance of the model is relatively good.

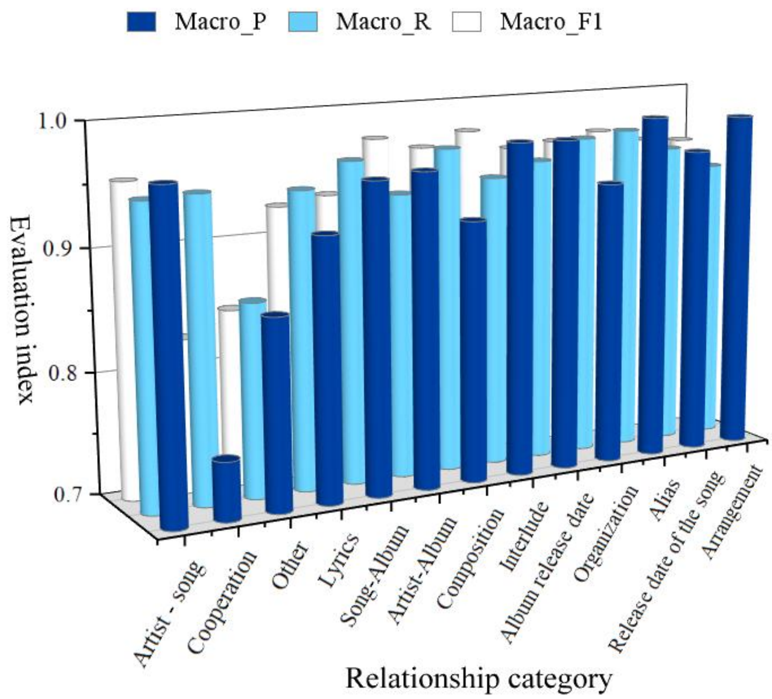

Figure 6. Category of relationships between entities.

5.2. Knowledge graph presentation

The knowledge storage module reads out and transforms the results of the knowledge fusion module, transforms entities and their information into entity data, relational data, and stores them using Neo4j. At present, the music domain knowledge graph in this paper has been applied in the multi-round dialog music search system as a powerful knowledge base in the search system. Next show the music domain knowledge graph designed in this paper: figure 7 shows the music domain knowledge graph display, the round nodes represent the entities, and the connecting edges between the entities represent the relationships between the entities. This representation of entities and inter-entity relationships in the form of graphs makes the relationships of entities visualized, and the relationships between entities are also inferrable, and it is also convenient to query the multi-hop relationships between entities in the query process.

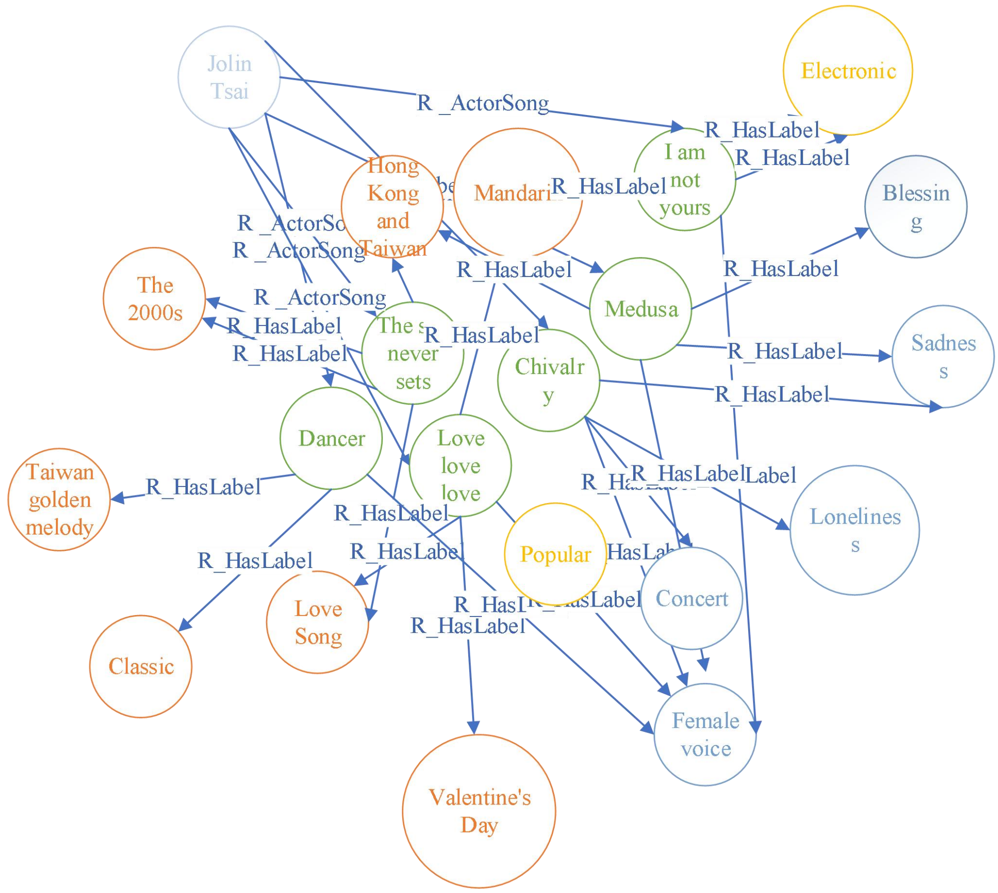

Figure 7. Music field knowledge map display.

5.3. Effectiveness of Civics Teaching in Music Program

In order to further understand the ideological awareness of music students in the Master of Arts program, this paper conducted a questionnaire survey in six classes in the first year of study. Three of the classes served as experimental classes and the other three as control classes. The experimental classes combined the study of Civics and Politics in the music curriculum under the framework of Greater Civics and Politics Education in their music teaching, and actively incorporated Civics and Politics elements into their teaching to cultivate students' national sentiment, national culture self-confidence, collectivism, and to promote the enhancement of students' moral qualities. The control class does not deliberately incorporate elements of Civics and Politics.

After one academic year of teaching, a questionnaire survey on students' consciousness of Civics was conducted. After recycling and organizing, it was found that the validity rate of the final questionnaire was higher than 95%, which can guarantee that the questionnaire data has a high research value.

The questionnaire involves four dimensions: patriotism, national cultural confidence, moral character, and collectivism. Each dimension has 2-4 questions. After the survey, after one year of teaching folk songs with the integration of the elements of ideology and politics, the four dimensions are improved in the order of patriotism>moral quality>national cultural confidence>collectivism.

There are four topics in the patriotism dimension, involving students' concern for the development of the country, their feelings about the growing strength of the motherland, their enthusiasm for singing the national anthem when the flag is raised, and their perception of the spirit of patriotism in music. On the dimension of patriotism, the comparison of the effect of course contemplation based on knowledge mapping music teaching is shown in Table 1.

After integrating the elements of Civics and Politics in music teaching in the music classroom, the proportion of students in the experimental class, compared with students in the control class, in terms of caring about the country, the proportion of students who cared about national events and the development of the country on a regular basis was increased by 10.99%, and the proportion of students who could very obviously feel that the motherland was becoming stronger day by day was increased by 9%. In the national flag-raising ceremony, the proportion of students who enthusiastically and actively sang the national anthem also increased by 9.47%, and the proportion of students who were able to quickly and intuitively feel the spirit of musical patriotism increased by 6.48%. Overall, after one year of teaching with the integration of Civics and Politics elements in music teaching, the students in the experimental class had a significant improvement in their sense of patriotism.

Table 1. Comparison of the dimension of patriotism.

Serial number	Question content	Options	Experimental Class	Control class	The effect is significant
1	Do you take active care to develop relevant information about the country?	Frequent concern	55.68%	44.69%	More significant
		Hyperthermal relation	39.44%	45.15%	
		None of your heart	4.88%	10.16%	
2	Do you take active care to develop relevant information about the country?	Very obvious	61.48%	52.48%	More significant
		Some obvious	35.48%	39.48%	
		Not obvious	3.04%	8.04%	
3	Will you sing the national anthem on Monday?	Sing with enthusiasm	70.95%	61.48%	Somewhat significant
		Initiative	21.69%	26.15%	
		Not very active	7.35%	12.37%	
4	Whether you can quickly feel the patriotism of the new music?	Very capable of	75.96%	69.48%	General
		General	19.98%	21.69%	
		Not sure	4.06%	8.83%	

There are four topics in the cultural self-confidence dimension, involving students' exposure to music in daily life, their initiative in music learning, their self-confidence in music culture and their willingness to pass on national culture. In the cultural self-confidence dimension, a comparison of the effect of curriculum civics on music teaching based on knowledge mapping is shown in Table 2.

After the integration of Civics elements in music classroom teaching, compared with the control class students, in the two dimensions of students' exposure to music in daily life and initiative in music learning, the changes between the experimental class and the control class are not obvious, the proportion of students who often listen to music increases by less than 3%, and the proportion of students who take the initiative to learn about culture increases by only 4.32%. However, in the two aspects of cultural self-confidence and awareness of promoting culture, the experimental class changed significantly compared with the control class, in which the proportion of students who were very confident in culture increased by 9.38%, and the proportion of students who were not very willing to promote culture decreased by 9.68%. On the whole, after one year of teaching music with the integration of Civic and Political elements, the students in the experimental class had a certain degree of improvement in cultural self-confidence consciousness, but it was not as obvious as the patriotism dimension.

Table 2. Comparison of the dimension of moral quality and confidence.

Serial number	Question content	Options	Experimental Class	Control class	Whether the effect is significant
1	Do you listen to music in your daily life	Often	28.77%	25.48%	Not significant
		Occasionally	40.49%	38.26%	
		Little	30.74%	36.63%	

2	When you learn music, you will take the initiative to understand the culture and the story	Compare initiative	34.52%	30.18%	Not significant
		General	45.39%	48.36%	
		Not very active	20.09%	21.46%	
3	Do you become more and more confident in music culture	Be confident	60.24%	50.86%	Relatively significant
		Some confidence	33.48%	34.24%	
		Not confident	6.28%	14.90%	
4	Are you willing to learn and carry on culture	Very willing	36.85%	30.45%	Relatively significant
		General	37.24%	33.96%	
		Be reluctant to	25.91%	35.59%	

Table 3 shows the comparison of moral character confidence dimension, students in the experimental class compared with students in the control class in the degree of understanding of socialist core values, the current universal core values education and related activities, so the master of arts is not unfamiliar with the core values, but the proportion of students who are more aware of the connotation of the core values will not be high, after the teaching of music civic education, it can be seen that a very good understanding of the core values of the student's proportion has increased by 6.14%, and the proportion of those who don't understand has decreased by 7.38%, the effect is relatively significant. In the degree of feeling of excellent qualities such as integrity and kindness, the proportion of students who are very capable of feeling excellent qualities has increased by 5.35%, and the effect is relatively significant.

Table 3. Contrast of moral quality confidence dimension.

Serial number	Question content	Options	Experimental Class	Control class	Whether the effect is significant
1	Do you understand the connotation of the core values of social nonitarianism	Know very well	32.48%	26.34%	Relatively significant
		General	38.45%	37.08%	
		Out of line	29.20%	36.58%	
2	Can you feel good, integrity, quality unity and courage from music	Very capable of	40.59%	35.24%	· Relatively significant
		General	29.24%	40.00%	
		Be not able to	30.17%	24.76%	
3	Do you think you can improve your moral qualities through music learning	Know very well	25.36%	23.39%	Not significant
		General	37.96%	38.41%	
		Out of line	36.68%	38.20%	

Table 4 shows the comparison of collectivism dimension, after integrating the elements of Civics in music teaching in the music classroom, the students in the experimental class, compared with the students in the control class, there is no significant enhancement effect of the students' sense of collectivism and perception of the spirit of collectivism, the enhancement of the magnitude of 1.21% and 1.21%, respectively. This does not negate the pedagogical practice of this music evidence, which occurs because there is so little music content that deals with the spirit of collectivism and because the development of the spirit of collectivism is mostly dependent on various types of practical activities.

Table 4. he collective dimension is compared.

Serial number	Question content	Options	Experimental Class	Control class	Whether the effect is significant
1	Do you have a high concentration awareness	High	28.69%	27.48%	Not significant
		General	38.42%	39.48%	
		Indeterminate	32.89%	33.04%	
2	Can you feel the spirit of collectivism from music	Very capable of	40.36%	39.15%	Not significant
		General	42.96%	43.85%	
		Be not able to	16.68%	17.00%	

6. Conclusion

Based on the related technology of knowledge graph construction, this paper constructs a relationship extraction model for music entity information and feature reuse, and establishes a music professional curriculum system in the context of the Great Civic and Political Education. Through the combination of simulation experiment and empirical analysis, the results of music relationship extraction and the effect of Civic and Political education are analyzed. The experimental results show that the number of entity relationship categories of artist-song is the most 1560, and the number of entity relationships of arrangement is the least, only 100. Statistical relationship extraction dataset text length frequency, the length of the text is mostly concentrated in the range of 10 to 80, and the frequency is between 1 and 183. The data obtained from the above experiments were transformed to obtain information such as entity data and relationship data, to construct a knowledge graph in the field of music, and to realize the visual display of music entity relationships. In order to further understand the music course learning of masters of arts in the context of big ideology and politics, a controlled experiment was carried out on the subject class. Patriotism has the most significant improvement effect in the four dimensions, the proportion of students who care about national events and national development, can very obviously feel that the motherland is getting stronger and stronger enthusiastically and positively sing the national anthem, and can quickly and intuitively feel the spirit of patriotism in music have been improved by 10.99%, 9%, 9.47%, and 6.48%, respectively, and the experimental class has a significant increase in the awareness of patriotism.

References

1. Gao, Y. (2023). Exploring the Value Orientation of College Art Students Based on Marxist philosophy. Cultura: International Journal of Philosophy of Culture and Axiology, 20(3).
2. Zhang, Y. (2020). Research and Analysis on the Ideological Status of Contemporary Young College Students. Frontiers in Educational Research, 3(4).
3. Liu, J. (2019, August). Research on Ideological and Political Education in Colleges and Universities under the Background of Art Practice. In 1st International Symposium on Innovation and Education, Law and Social Sciences (IELSS 2019) (pp. 14-18). Atlantis Press.
4. Sarapik, V. (2017). Ideology and Art. Kunstiteaduslikke Uurimusi, 26(01+ 02), 7-14.
5. Yang, D. (2020, October). Art Education and College Students' Ideology Education Method under Information Technology. In 2020 International Conference on Computers, Information Processing and Advanced Education (CIPAE) (pp. 108-110). IEEE.
6. Liang, J., & Zheng, H. (2023). The Challenges and Countermeasures of Ideological and Political Education Theory Course Teaching for Art Students in Universities in the New Media Era. Journal of Education and Educational Technologies, 3(11), 11-16.
7. Wang, M., & Ning, Q. (2020, November). Research on the Way to Carry Out Ideological and Political Education for the "Post-2000s" of Colleges in the New Media Era. In 2020 5th International Conference on Modern Management and Education Technology (MMET 2020) (pp. 727-733). Atlantis Press.

8. Fu, K. (2023). The Significance of Ideological and Political Education in College Physical Education. Journal of Theory and Practice of Social Science, 3(9), 1-7.
9. Springer, D. G. (2020). Students' perspectives on the master's degree in music education. Journal of Music Teacher Education, 30(1), 79-92.
10. Bebetsos, E. (2015). Prediction of Participation of Undergraduate University Students in a Music and Dance Master's Degree Program. International Journal of Instruction, 8(2), 165-176.
11. Rao, T. (2021). The exploration of integrating piano teaching into ideological and political education from the perspective of morality building and people cultivating. Region-Educational Research and Reviews, 3(1), 1-5.
12. Zhou, J., Wang, D., & Li, Z. (2025). Exploring the Integration of Ideological and Political Education into Vocal Music Courses in Higher Education Institutions under the Framework of Music Education Certification. Education Insights, 2(9), 61-66.
13. Han, Z. (2025). Research on Collaborative Education System of Music in Colleges and Universities under Ideological and Political Education Courses. Research and Commentary on Humanities and Arts, 2(10).
14. Lu, X. (2024). Exploration Methods and Approaches for the Integration of Vocational Music Education and Curriculum Ideology and Politics. International Journal of Social Science and Education Research, 7(10), 358-363.
15. Zhang, Z. (2025). Discussion on ideological and political elements in music education under the background of "three complete education". Proceeding in Humanities, Education and Social Sciences, 4, 1-2.
16. Zhang, Q., & He, Q. (2023). Integrating ideological and political education into music education on college students' mania. CNS Spectrums, 28(S2), S41-S41.
17. Dang, F. R., Tang, J. T., Pang, K. Y., Wang, T., Li, S. S., & Li, X. (2021). Constructing an educational knowledge graph with concepts linked to Wikipedia. Journal of Computer Science and Technology, 36(5), 1200-1211.
18. Li, N., Shen, Q., Song, R., Chi, Y., & Xu, H. (2022). MEduKG: a deep-learning-based approach for multi-modal educational knowledge graph construction. Information, 13(2), 91.
19. Wang, P., & Lin, J. (2021, June). Construction and Application of Knowledge Graph of Teaching Arrangement for Cross-disciplinary Majors. In 2021 IEEE 3rd International Conference on Computer Science and Educational Informatization (CSEI) (pp. 85-88). IEEE.
20. Yue, H., Lin, H., Jin, Y., Zhang, H., & Cai, K. (2022). Opening knowledge graph model building of artificial intelligence curriculum. International Journal of Emerging Technologies in Learning (iJET), 17(14), 64-77.
21. Huang, Z., Tian, R., & Fu, G. (2024). Mining and Teaching Practice of Ideological and Political Course Elements Based on Text Extraction and Knowledge Graph. Journal of Combinatorial Mathematics and Combinatorial Computing, 119, 95-104.
22. Zhu, G., & Tao, T. (2022, December). Application of Knowledge Graph in Ideological and Political Teaching. In 2nd International Conference on Internet, Education and Information Technology (IEIT 2022) (pp. 858-864). Atlantis Press.
23. Weng, J., Gao, Y., Qiu, J., Ding, G., & Zheng, H. (2019, August). Construction and application of teaching system based on crowdsourcing knowledge graph. In China conference on knowledge graph and semantic computing (pp. 25-37). Singapore: Springer Singapore.
24. Qiao, L., Yin, C., Chen, H., Sun, H., Rong, W., & Xiong, Z. (2019, December). Automated constmction of course knowledge graph based on China MOOC platform. In 2019 IEEE International Conference on Engineering, Technology and Education (TALE) (pp. 1-7). IEEE.
25. Xing Qi,Xiaoyu Shen,Yucheng Zhang,Bo Yang,Keqiang Xie & Nan Dong. (2025). A manufacturing knowledge graph completion method based on a lightweight dual encoding model. Applied Intelligence,55(15),1009-1009.https://doi.org/10.1007/S10489-025-06909-0.
26. Bing Liu. (2020). Text sentiment analysis based on CBOW model and deep learning in big data environment. Journal of Ambient Intelligence and Humanized Computing,11(2),451-458.https://doi.org/10.1007/s12652-018-1095-6.
27. Fei Chen,Zhenling Zhang,Yangli Jia,Ruchao Jia,Haitao Wang & Xinyu Cao. (2025). Research on the similarity calculation of short text in the terminology domain based on siamese BERT model. Scientific Reports,15(1),36954-36954.https://doi.org/10.1038/S41598-025-20908-8.