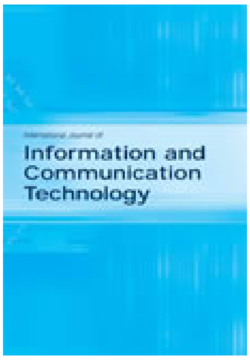


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Imparting traditional wisdom and political knowledge through deep tracking and knowledge graph model

Guoqiang Yu

DOI: 10.1504/IJICT.2025.10072410

Article History:

Received:	24 May 2025
Last revised:	12 June 2025
Accepted:	13 June 2025
Published online:	05 August 2025

Imparting traditional wisdom and political knowledge through deep tracking and knowledge graph model

Guoqiang Yu

School of Marxism,
Hangzhou Vocational and Technical College,
Hangzhou 310000, China
Email: yuguoqiang6660304@163.com

Abstract: In the context of the intelligent transformation of ideological and political education, how to achieve objective prediction of students' knowledge mastery has become a research difficulty. In response to the problem of insufficient model generalisation caused by traditional methods relying on non-public data, this study proposes a prediction model for ideological and political knowledge mastery based on deep tracing and knowledge graph (DKP model). Firstly, extract the temporal characteristics of learning behaviour and the logical correlation between knowledge points. Secondly, design a dynamic knowledge tracking framework and introduce the knowledge graph attention network (KGAT) to model the dialectical relationships and cognitive transfer paths between ideological and political knowledge points. Further, construct an interpretable mastery prediction index system. The cross-domain data fusion paradigm and open-source knowledge graph construction method proposed in the study provide a reproducible technical framework and infrastructure support for educational equity research.

Keywords: ideological and political education; knowledge graph; LSTM; cognitive diagnosis.

Reference to this paper should be made as follows: Yu, G. (2025) 'Imparting traditional wisdom and political knowledge through deep tracking and knowledge graph model', *Int. J. Information and Communication Technology*, Vol. 26, No. 29, pp.59–74.

Biographical notes: Guoqiang Yu obtained his Doctoral degree from Suzhou University in 2022. He is currently a Lecturer at the Institute of Marxism Studies in Hangzhou Vocational and Technical College. His research interests include artificial intelligence risks, digital society, and ideological and political education.

1 Introduction

In the process of digital transformation of higher education and deep integration of ideological and political courses, building an intelligent and interpretable system for predicting the mastery of ideological and political knowledge has become the core challenge in implementing the fundamental task of 'cultivating virtue and nurturing talents'. Although deep learning technology has made significant progress in knowledge tracking in disciplines such as mathematics and programming (Jiao et al., 2021), ideological and political education, due to its unique ideological attributes, dialectical logical structure, and value internalisation goals, has put forward higher requirements for the domain adaptability and interpretability of models (Liu et al., 2023). The current research faces three bottlenecks: at the theoretical level, traditional cognitive diagnostic models (such as IRT and DINA) are difficult to quantify the nonlinear logical relationships between ideological and political knowledge points (such as the philosophical associations of 'contradictory opposition unity' and 'quantitative change to qualitative change'). On a technical level, mainstream knowledge tracking models such as DKT and DKVMN rely on closed data construction, which limits their generalisation ability and makes it difficult to reproduce. On the ethical level, there is a conflict between the black box prediction mechanism and the value guidance goals emphasised in ideological and political education, and an interpretable framework that conforms to the laws of education is urgently needed.

In recent years, educational knowledge graph technology has provided a new path for subject knowledge modelling. Ataeva et al. (2024) used the ontology of mathematics in the LibMeta library as an example to discuss the techniques for generating encyclopedic knowledge graphs and modern applications of mathematics. Discussed the issues of filling graphs, embedding data in graphs, and extracting connections and nodes from graphs. Chen et al. (2018) used neural sequence labelling algorithm to extract teaching concepts from teaching data, and used probabilistic association rule mining to identify the relationship between learning evaluation data and educational significance. Li et al. (2022) proposed a method for automatically constructing multimodal educational knowledge graphs, which integrates speech as a modal resource to promote the reuse of educational resources. Sun et al. (2016) mined various data related to education from massive network data and fused them together. Then, use the extracted named entities and entity relationships to construct a knowledge graph centred around educational events. However, existing methods have not yet solved the problem of structured representation and dynamic tracking of ideological knowledge in ideological and political education.

In response to the above limitations, this study proposes a prediction model for ideological and political knowledge mastery based on deep tracking and knowledge graph, which achieves breakthroughs in the following three aspects:

1 Theoretical framework innovation: Integrating Marxist epistemology and cognitive diagnostic theory, constructing a three-dimensional evaluation system of 'cognitive foundation dialectical thinking practical ability'. By explicitly encoding the logical network of core knowledge points such as 'dialectical materialism' and 'socialist core values' through a knowledge graph, the computable representation of ideological cognitive laws can be achieved.

2 Algorithm design innovation: Propose a hierarchical LSTM-KGAT hybrid architecture, combining the long-term modelling ability of bidirectional LSTM for learning behaviour with the logical reasoning mechanism of knowledge graph attention network (KGAT). KGAT innovatively transforms philosophical relationships (such as 'universality and particularity') into graph attention weights, dynamically capturing cognitive transfer pathways.

3 Explanatory innovation: Based on cognitive diagnostic theory, a multi granularity explanatory index system is constructed to output teaching operable diagnostic

reports such as ‘concept mastery’, ‘logical coherence’, and ‘error attribution’, meeting the dual needs of ideological and political education for value guidance and cognitive analysis.

2 Relevant technologies

2.1 Knowledge graph

As a structured semantic network representation method, the core goal of knowledge graph is to systematically organise domain knowledge and reveal its inherent relationships through the triplet form of entities, relationships, and attributes. Compared with traditional knowledge representation models such as ontology or semantic networks, the advantage of knowledge graphs lies in their dynamic scalability and compatibility with heterogeneous sources. In the field of education, knowledge graphs can provide a computable semantic foundation for learning path planning and cognitive diagnosis by explicitly defining the logical relationships between knowledge points, such as prerequisite dependencies, hierarchical inclusions, and causal relationships (Zhong et al., 2023). For example, the knowledge graph in mathematics education can construct a cognitive chain through the progressive relationship of ‘function derivative integral’, while ideological and political education needs to deal with complex logical networks with ideological characteristics such as ‘materialism dialectics epistemology’ (Peng et al., 2023). This structured representation not only supports the visual presentation of knowledge, but also provides an interpretable reasoning framework for machine learning models.

The construction of educational knowledge graphs typically follows a domain driven design paradigm. Firstly, it is necessary to define core knowledge points and their relationship types based on subject curriculum standards and authoritative textbooks. For example, in ideological and political education, it is necessary to cover the correlation between modules such as ‘basic principles of Marxism’ and ‘theoretical system of socialism with Chinese characteristics’. Subsequently, knowledge units are extracted from multimodal data sources (textbook texts, instructional videos, question banks) using natural language processing techniques such as entity recognition and relationship extraction, and the educational effectiveness of logical relationships is ensured through double-blind validation by domain experts (Shen et al., 2022). In this process, expert verification is particularly crucial to avoid logical fallacies or cognitive hierarchy misplacement caused by automated extraction. For example, when constructing a subgraph of ‘socialist core values’, it is necessary to strictly distinguish the attribution relationship between ‘national level value goals’ and ‘individual level value norms’ to prevent conceptual confusion. After completing knowledge extraction, discrete knowledge points are mapped into a low dimensional vector space using graph embedding techniques such as TransE and GraphSAGE, which preserves semantic similarity and reflects cognitive dependencies in the topological structure (Zhu et al., 2022).

The application mechanism of knowledge graph in educational cognitive diagnosis is mainly reflected in its support for cognitive state modelling and transfer path inference. The knowledge mastery state of learners can be mapped to the activation intensity of nodes in the graph. For example, if a student’s activity is high at the ‘contradiction universality' node but weak at the 'contradiction specificity' node, it indicates that their dialectical thinking has a local discontinuity. In addition, graph neural networks (GNNs) can capture implicit associations between knowledge points through message passing mechanisms (Chandak et al., 2023). For example, when students frequently make mistakes at the nodes of 'practice determines cognition' and 'cognition reacts on practice', the model can automatically infer that they have not understood the higher-order concept of 'dialectical relationship between cognition and practice' (Santos et al., 2022). This reasoning ability enables knowledge graphs to not only statically describe knowledge structures, but also dynamically track cognitive development trajectories, providing a basis for personalised interventions.

The construction of a knowledge graph for ideological and political education faces unique challenges. Firstly, ideological knowledge has distinct dialectical logical characteristics and requires explicit encoding of philosophical relationships such as 'unity of opposites' and 'quantitative and qualitative changes', which is fundamentally different from the linear progressive cognitive path in the STEM field. For example, the interactive relationship between productivity and production relations needs to be designed with bidirectional dependency edges rather than unidirectional pre-existing relationships. Secondly, the goal of internalising the value of ideological and political knowledge requires the integration of a cross modal relationship between theoretical identification and practical application into the graph, such as the correlation analysis between the emotional characteristics of text in classroom discussions and post class practical behaviour data. Therefore, this study proposes a three-layer graph architecture of 'semantics cognition value', which describes the logical relationship of knowledge points at the basic semantic layer, maps learning behaviour patterns at the cognitive layer, and associates emotional attitudes and behavioural practice data at the value layer, thus fully supporting the dynamic tracking of ideological knowledge.

The theoretical development of knowledge graph provides important support for educational intelligence, but its deep application in ideological and political education still needs to address domain adaptation issues. For example, traditional graph embedding algorithms assume relationship symmetry by default, which is difficult to adapt to philosophical propositions such as 'the economic base determines the superstructure' with strong one-way correlations. In response to this, this study introduces a direction sensitive graph attention mechanism that accurately characterises the asymmetric dependency characteristics of ideological knowledge by dynamically adjusting relationship weights. This innovation not only expands the theoretical boundaries of knowledge graphs, but also lays the methodological foundation for the LSTM-KGAT hybrid model proposed in subsequent chapters.

2.2 Knowledge graph attention network

KGAT as an important branch of GNN focuses on dynamically capturing the semantic correlation strength between entities and relationships in the knowledge graph through attention mechanisms. Unlike traditional graph convolutional networks (GCNs) that rely on fixed neighbourhood aggregation patterns, KGAT assigns differentiated importance measures to different adjacent nodes through learnable attention weights (Liu et al., 2021). This characteristic enables it to effectively handle the asymmetric logical relationships commonly present in the knowledge graph of ideological and political education (such as 'determination reaction' and 'universality specificit'), thereby more accurately representing the dialectical structure of ideological knowledge. For example, in the correlation analysis of ‘productivity and production relations’, KGAT can automatically identify the unidirectional dominant relationship of ‘productivity determines production relations’ instead of simply symmetrising it, which provides key support for subsequent cognitive state inference.

In the field of education, the application value of KGAT mainly lies in its ability to dynamically model cognitive transfer pathways. The process of learners’ knowledge acquisition is not a linear progression, but presents a multi-path, non-uniform cognitive transition characteristic. KGAT can explicitly reveal the potential cognitive dependencies between knowledge points through attention weight distribution (Li et al., 2021). For example, when students frequently make mistakes at the node of ‘practice is the only criterion for testing truth’, the model can discover a weak correlation with the node of ‘repetitiveness and infinity of cognition’ through attention weight analysis, and infer that students have failed to establish a holistic cognitive framework of the ‘practice cognition’ dialectical relationship. This dynamic reasoning mechanism breaks through the dependence of traditional knowledge tracking models on preset cognitive paths and is more in line with the emergent characteristics of human cognition.

The design of KGAT needs to closely integrate the semantic characteristics of knowledge graphs. In the context of ideological and political education, the logical relationships between knowledge points have distinct hierarchical and non-Euclidean characteristics (Du et al., 2022). To this end, KGAT introduces a hierarchical attention mechanism: at the local level, the semantic similarity weights of knowledge points within the same value level are calculated. At the global level, capture the overall consistency of value objectives through cross layer attention gating modules. This hierarchical design not only enhances the model’s ability to represent complex ideological structures, but also provides teachers with interpretable teaching diagnostic criteria through visual output of attention weights.

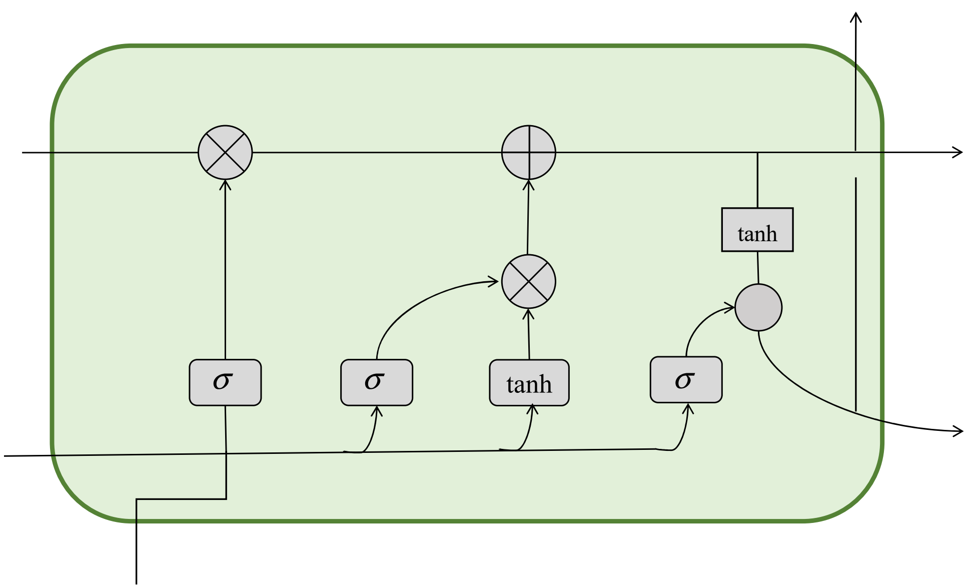
It is worth noting that the application of KGAT in the ideological and political knowledge graph needs to address the issue of relational heterogeneity. Traditional graph attention networks typically assume that the types of relationships in the graph are single or homogeneous, while ideological and political knowledge graphs contain multiple logical relationships (such as ‘included’, ‘contradictory unity’, ‘historical inheritance’, etc.). To address this challenge, a relationship aware attention computing framework is proposed: firstly, different types of relationships are encoded into independent semantic spaces through relationship embedding vectors. Secondly, when calculating the attention coefficient between nodes, relationship types are introduced as dynamic adjustment factors. For example, when dealing with the opposing relationship between ‘materialism and idealism’, the model generates negative attention weights to reflect their conflicting attributes, while when dealing with the inheritance relationship between ‘Marxism with Chinese characteristics’ and ‘socialism with Chinese characteristics’, it assigns high positive weights to strengthen cognitive continuity. This mechanism enables KGAT to more finely depict the inherent contradictions and unified laws of ideological knowledge.

The theoretical development of KGAT still needs to be deeply integrated with the cognitive laws of education. Cognitive diagnostic research shows that the process of knowledge internalisation by learners presents a phased characteristic of ‘contact understanding criticism practice’. To this end, KGAT is extended to a dynamic attention network in the temporal dimension: by introducing LSTM modules to capture learning behaviour sequences, the cognitive state changes at time steps are mapped to dynamic adjustments of attention weights. For example, when students deepen their understanding of the 'mass historical view' through case analysis at a certain stage, the model will simultaneously enhance the attention connection strength between this node and related knowledge points such as 'social existence determines social consciousness'. This spatiotemporal coupling attention mechanism enables the model to simultaneously capture the static logic of knowledge structure and the dynamic trajectory of cognitive development, laying a theoretical foundation for personalised learning path recommendation.

2.3 LSTM

LSTM as a special type of recurrent neural network (RNN), aims to solve the problems of gradient vanishing and exploding that traditional RNNs face in modelling long sequence data (Lindemann et al., 2021). Unlike ordinary RNNs that rely solely on simple recurrent units to transmit information, LSTM achieves selective memory and forgetting of temporal information by introducing gating mechanisms (forget gate, input gate, output gate) and cell state (cell state). the structure is shown in Figure 1. This feature enables it to effectively capture long-range dependencies in learning behaviour sequences.

Figure 1 LSTM structure diagram (see online version for colours)



The core innovation of LSTM lies in the design of its gating mechanism. The forget gate determines which information in the previous cell state needs to be retained or discarded through an activation function (such as Sigmoid). For example, when a student correctly answers the same knowledge point related question multiple times in a row, the model can automatically reduce the forgetting probability of that knowledge point. The input gate is responsible for filtering the parts of the current input information that need to be updated to the cellular state, for example, the cognitive changes that students experience after watching the 'dialectical materialism' teaching video can be encoded as new memory units. The output gate further controls which information in the current cell state will be transmitted to the hidden state, thereby affecting the subsequent prediction results. This refined control mechanism enables LSTM to simultaneously handle short-term behavioural fluctuations (such as temporary memory reinforcement) and long-term cognitive development (such as gradual improvement of dialectical thinking ability) in ideological and political learning (Wen and Li, 2023).

In the field of education, the temporal modelling capability of LSTM is particularly suitable for learning path analysis and knowledge state inference. Taking ideological and political education as an example, students' learning behaviour usually exhibits non-uniformity and multi-scale characteristics: at the micro level, the order of answering a single classroom test reflects the degree of instant knowledge mastery. At the macro level, changes in cross semester learning engagement reflect the overall development trend of cognitive abilities. LSTM can naturally fuse behavioural features of different time granularities through multi-level hidden states. For example, when analysing the learning trajectory of students in the course of 'basic principles of Marxism', the model can capture the score fluctuations of weekly thematic tests through short-term memory units, while using long-term memory units to track the overall internalisation level of their worldview methodology (Tatsunami and Taki, 2022). This multi-scale modelling capability is significantly superior to traditional time series analysis methods such as ARIMA, especially exhibiting stronger robustness when dealing with sparse and irregular online learning behaviour data.

However, standard LSTM still faces domain adaptation challenges in ideological and political knowledge tracking tasks. On the one hand, the ideological attribute of ideological and political knowledge requires models to not only capture statistical patterns of behaviour sequences, but also understand the logical connections between knowledge points. For example, students' performance on the two knowledge points of 'historical materialism' and 'surplus value theory' may be influenced by their underlying philosophical connections, and traditional LSTM is difficult to explicitly model such semantic relationships. On the other hand, the goal of internalising the value of ideological and political education requires the model to output explanatory indicators with educational significance, rather than relying solely on end-to-end black box predictions. Therefore, research needs to combine LSTM with modules such as knowledge graph and attention mechanism to construct a hybrid architecture to compensate for its theoretical limitations (Seabe et al., 2023). For example, by jointly training the hidden states of LSTM with knowledge graph embedding vectors, the model can simultaneously utilise temporal behavioural features and semantic associations of knowledge points, thereby enhancing its ability to model complex cognitive patterns.

It is worth noting that the gating mechanism of LSTM has inherent compatibility with educational cognitive theory. According to the cognitive load theory, learners' knowledge absorption efficiency is limited by the capacity of working memory, and cognitive resource allocation needs to be optimised through information screening and integration. The forget gate and input gate mechanisms of LSTM precisely simulate this cognitive process. In addition, bidirectional LSTM (BiLSTM) models sequence data in both forward and backward time dimensions, which can capture the 'forward promotion' and 'backward transfer' effects in ideological and political knowledge learning (Nguyen et al., 2021). For example, after mastering the theory of socialism with Chinese characteristics, students may deepen their understanding of the theory of the primary stage of socialism in reverse, and this reverse cognitive path can be effectively represented through a bidirectional hidden state.

The theoretical development of LSTM provides a powerful tool for analysing educational temporal data, but its deep application in ideological and political education still needs to address domain specific issues. For example, traditional LSTM assumes equal time intervals by default, while real learning behaviour often presents non-uniform time intervals (such as sudden reviews and long-term interruptions). In response to this, a time aware LSTM variant is proposed, which dynamically adjusts memory strength by introducing a time decay factor to more realistically simulate human memory patterns. This improvement enables the model to quantify the 'learning interval effect' – the impact of the time interval between two learning behaviours on knowledge retention, providing a theoretical basis for the generation of personalised review plans (Shahid et al., 2021).

2.4 Mastery and prediction of ideological and political knowledge

As a core component of higher education, ideological and political education has a distinct ideological attribute and complex logical structure in its knowledge system. Traditional assessment methods rely heavily on static testing and subjective evaluation, making it difficult to dynamically track the multidimensional transformation process of 'theoretical identification value internalisation practical application' in students' cognitive development. With the development of artificial intelligence technology, knowledge tracking models based on deep learning provide a new path for the evaluation of ideological and political education, but their application needs to be rooted in the deep integration of Marxist epistemology, cognitive science theory, and educational laws. This article systematically explains the core logic and innovative value of predicting the mastery of ideological and political knowledge from a theoretical perspective.

The hierarchical and networked characteristics of the ideological and political knowledge system require predictive models to break through the limitations of traditional linear cognitive paths. The knowledge graph constructs a semantic network through entity relationship triplets (<knowledge points, relationships, knowledge points>), providing a foundation for the computability of ideological knowledge. Traditional RNN models are difficult to capture cross level collaborative effects. Graph embedding algorithms (such as TransR) map discrete knowledge points to a low dimensional space, transforming propositions such as 'social existence determines social consciousness' into geometric constraints in vector space (such as translation invariance), providing structured prior knowledge for deep learning models.

The knowledge mastery status of learners exhibits significant temporal evolution characteristics. Behavioural learning theory suggests that the reinforcement gap effect and forgetting curve jointly shape cognitive trajectories. The LSTM network simulates this process through gate mechanisms (forget gate, input gate, output gate): the forget gate dynamically filters out invalid information (such as accidental answering), the input gate selectively integrates new knowledge (such as cognitive transitions after thematic learning), and the output gate controls the explicit expression of cognitive states. In the context of ideological and political education, bidirectional LSTM further captures the effects of forward promotion and backward transfer: for example, learning the 'Thought on socialism with Chinese characteristics for a new era' may deepen the understanding of the 'theory of the primary stage of socialism' in reverse. The model encodes cognitive cumulative effects across time steps through hidden state vectors, transforming the learning loop of theory practice reflection into a quantifiable state transition matrix.

The ideological orientation of ideological and political education requires predictive models to avoid black box risks. The attention mechanism explicitly reveals the decision-making basis of the model through weight distribution. For example, when students perform weakly at the 'mass historical view' node, the KGAT module can locate their weak correlation with the 'people's subject status' and generate diagnostic recommendations for 'weak foundation of historical materialism'. This mechanism coincides with the emphasis on 'visualisation of cognitive conflicts' in constructivist learning theory, enabling teachers to design cognitive scaffolds in a targeted manner. Meanwhile, relationship aware attention follows the 'three principles of educational technology ethics': transparency (traceable weights), controllability (manually corrected knowledge edges), and fairness (avoiding potential bias amplification), providing ethical guarantees for technology empowered education.

The theoretical innovation of mastering and predicting ideological and political knowledge is essentially the dialectical unity of educational laws and technological logic. Through the guidance of Marxist epistemology in model design, knowledge graph reconstruction of cognitive topology, and deep learning simulation of cognitive dynamics, this study provides a methodological framework that combines scientific and ethical considerations for ideological education evaluation. Future research needs to further explore the theoretical adaptability in cross-cultural scenarios, as well as the potential application of technologies such as federated learning in protecting student privacy, and continue to promote the deep integration of artificial intelligence and ideological and political education.

3 Hybrid model based on LSTM-KGAT

This chapter proposes a hybrid model that combines LSTM and KGAT, aiming to achieve dynamic prediction of ideological and political knowledge mastery through a collaborative mechanism of temporal behaviour modelling and knowledge association reasoning. The overall architecture of the model is shown in Figure 2, and its core process includes four stages: cross source data fusion, knowledge graph embedding, temporal feature extraction, and attention inference.

3.1 Cross source data fusion and feature representation

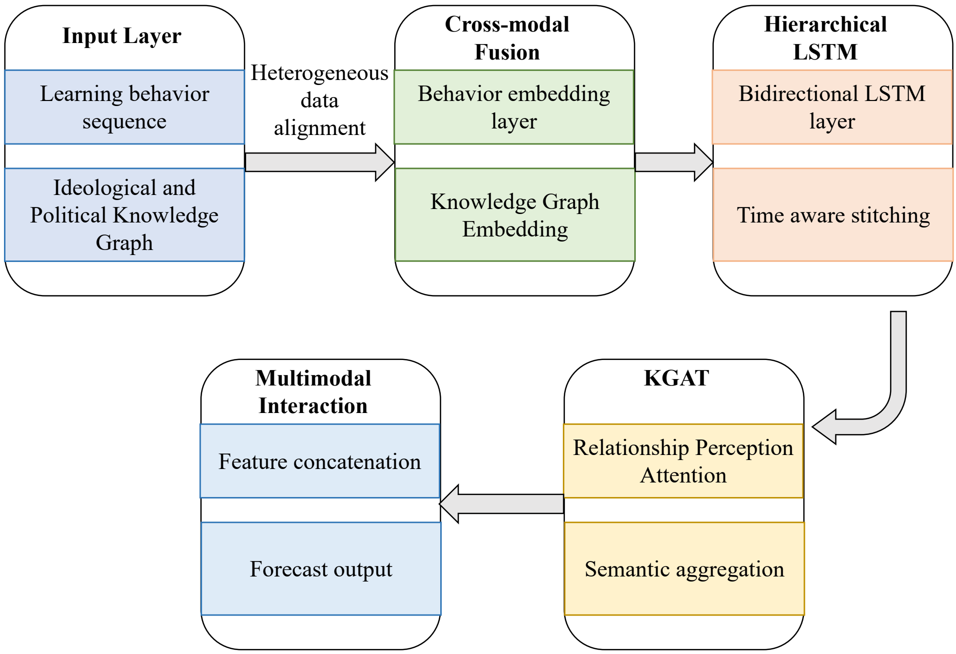
Assuming the input data includes a learning behaviour sequence $X = \{x_1, x_2, \ldots, x_T\}$ and an ideological and political knowledge graph $G = \{V, \varepsilon\}$, where $x \in \mathbb{R}^d$ represents the behaviour feature vector of the t^{th} time step (such as answer correctness, learning duration, interaction type), V is the set of knowledge points and $\varepsilon \subseteq V \times R \times V$ is the set of logical relationships (R is the relationship type). Firstly, perform joint encoding on heterogeneous data:

Behaviour sequence embedding maps the original behaviour features to the hidden space through a fully connected layer:

$$h_t^{(0)} = W_x x_t + b_x \tag{1}$$

where $W_x \in \mathbb{R}^{d_h \times d}$ and $b_x \in \mathbb{R}^{d_h}$ are learnable parameters, and d_h is the hidden layer dimension.

Figure 2 Model architecture diagram (see online version for colours)



The knowledge graph embedding uses the TransR algorithm to generate vector representations of knowledge point entities and relationships:

$$e_h^{(r)} = e_h M_r,\ e_t^{(r)} = e_t M_r \tag{2}$$

$$score(h, r, t) = \left\| e_h^{(r)} + r - e_t^{(r)} \right\|_2^2 \tag{3}$$

Among them, $e_h, e_t \in \mathbb{R}^k$ is entity embedding, $M_r \in \mathbb{R}^{k \times d_r}$ is relational projection matrix, and $r \in \mathbb{R}^{d_r}$ is relational embedding vector.

3.2 Hierarchical LSTM time series modelling

To capture the long-term and short-term dependencies of learning behaviour, a bidirectional LSTM network is designed to extract temporal features (Sagheer et al., 2021). Let $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ represent the forward and reverse hidden states, respectively:

$$\overrightarrow{h_t} = LSTM\left(h_t^{(0)}, \overrightarrow{h_{t-1}}\right) \tag{4}$$

$$\overleftarrow{h_t} = LSTM\left(h_t^{(0)}, \overleftarrow{h_{t+1}}\right) \tag{5}$$

After bidirectional state concatenation, a time aware representation is obtained:

$$h_t^{lstm} = \left[\overrightarrow{h_t}; \overleftarrow{h_t} \right] \tag{6}$$

This step dynamically filters noisy data through a forget gate and reinforces key behavioural patterns using an input gate.

3.3 Knowledge graph attention reasoning

Design a graph attention mechanism for relationship perception to model the dialectical logic between knowledge points. For knowledge point v_i, its neighbourhood set $N(i) = \{(v_j, r_{ij}) \mid (v_i, r_{ij}, v_j) \in \varepsilon\}$, the attention coefficient is calculated as follows:

$$\alpha_{ij}^{(r)} = \frac{\exp\left(\sigma\left(a_r^T \left[W_g e_i; W_g e_j^{(r)}\right]\right)\right)}{\sum_{k \in N(i)} \exp\left(\sigma\left(a_r^T \left[W_g e_i; W_g e_k^{(r)}\right]\right)\right)} \tag{7}$$

where $W_g \in \mathbb{R}^{d_g \times k}$ is the shared weight matrix, $a_r \in \mathbb{R}^{2d_g}$ is the attention vector related to the relationship, and σ is the LeakyReLU activation function. The aggregated knowledge points are characterised as:

$$z_i \sum_{j \in N(i)} \alpha_{ij}^{(r)} \cdot \left(W_r e_j^{(r)} + b_r\right) \tag{8}$$

This mechanism dynamically adjusts attention weights through relationship type r, such as assigning negative weights to ‘unity of opposites’ relationships to characterise conflict effects.

3.4 Multimodal feature interaction and prediction

Cross modal interaction between temporal behaviour feature h_t^{lstm} and knowledge semantic feature z_i:

$$z_i \sum_{j \in N(i)} \alpha_{ij}^{(r)} \cdot \left(W_r e_j^{(r)} + b_r\right) \tag{9}$$

where $W_c \in \mathbb{R}^{d_c \times (d_h + d_g)}$ is the interaction matrix. The final predicted probability for students to master knowledge point v_i is:

$$p_{ti} = sigmoid\left(W_p u_{ti} + b_p\right) \tag{10}$$

3.5 Loss function and optimisation

Joint optimisation of knowledge graph embedding loss and temporal prediction loss:

$$L = \sum_{(h,r,t) \in \varepsilon} \max\left(0, \gamma + score(h, r, t) - score(h', r, t')\right) + \lambda \sum_{t=1}^{T} \sum_{i=1}^{|V|} y_{ti} \log p_{ti} \tag{11}$$

The first term is the interval loss of TransR (γ is the boundary hyperparameter), the second term is the cross entropy prediction loss, and λ is the equilibrium factor. Using Adam optimiser for end-to-end training, while introducing gradient pruning to prevent semantic distortion of ideological and political knowledge embedding.

This framework transforms ideological and political principles into computable dynamic tracking processes through rigorous mathematical formalisation, providing a methodological tool that combines accuracy and interpretability for the intelligence of ideological and political education.

4 Experimental results and analysis

4.1 Experimental setup

The experiment was conducted based on the public dataset EdNet and the subset of ideological and political education from the China Education Panel Studies (CEPS). Baseline models including DKT, DKVMN, GKT, and IRT-DNN were compared. The evaluation indicators include AUC-ROC, F1 score, Kappa coefficient, and early detection rate (EDR) of weak knowledge points. The hyperparameters of the model were optimised through grid search. The dimension of the LSTM hidden layer was set to 64, and the number of KGAT attention heads was 2. During training, the early stopping method (patience value = 10) was used to prevent overfitting, and the learning rate was set to 0.001.

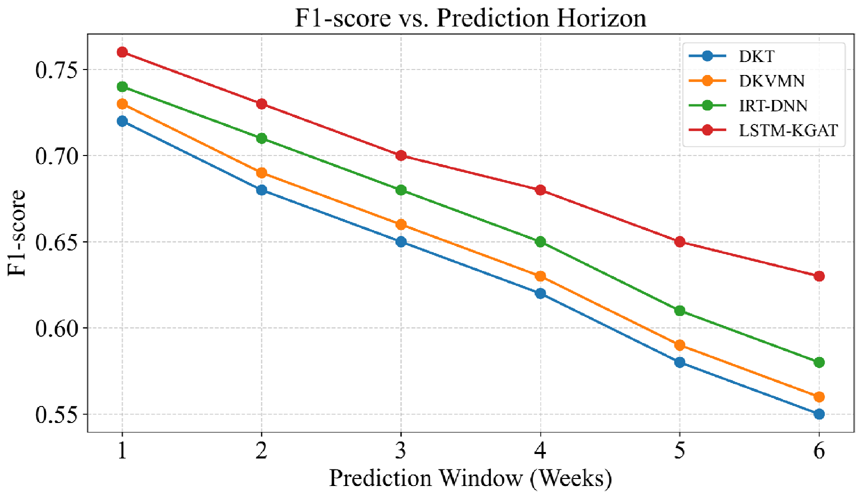
4.2 Experimental comparison results

Table 1 shows the performance comparison of different models. The AUC-ROC of LSTM-KGAT on the CEPS dataset is 0.76, which is 5.6% higher than the optimal baseline model IRT-DNN, and the F1 score reaches 0.71, EDR rate is 72.3%. On the EdNet dataset, the AUC-ROC of the model is 0.79, an improvement of 8.2% compared to DKT. As shown in Figure 3, long-term predictive performance analysis shows that the F1 score of the model within the 4-week window is 0.68, which is 9.7% higher than DKT, and the decrease is more gradual. The Kappa coefficient is 0.62, indicating moderate consistency between the predicted results and the actual cognitive state.

Table 1 Model performance comparison

Model	*AUC-ROC*	*F1-score*	*EDR (%)*	*Kappa*
DKT	0.69	0.63	60.1	0.52
DKVMN	0.70	0.65	63.4	0.55
GKT	0.67	0.61	57.8	0.49
IRT-DNN	0.72	0.67	65.2	0.58
LSTM-KGAT	0.76	0.71	72.3	0.62

Figure 3 Comparison of F1 score under long-term prediction window (see online version for colours)



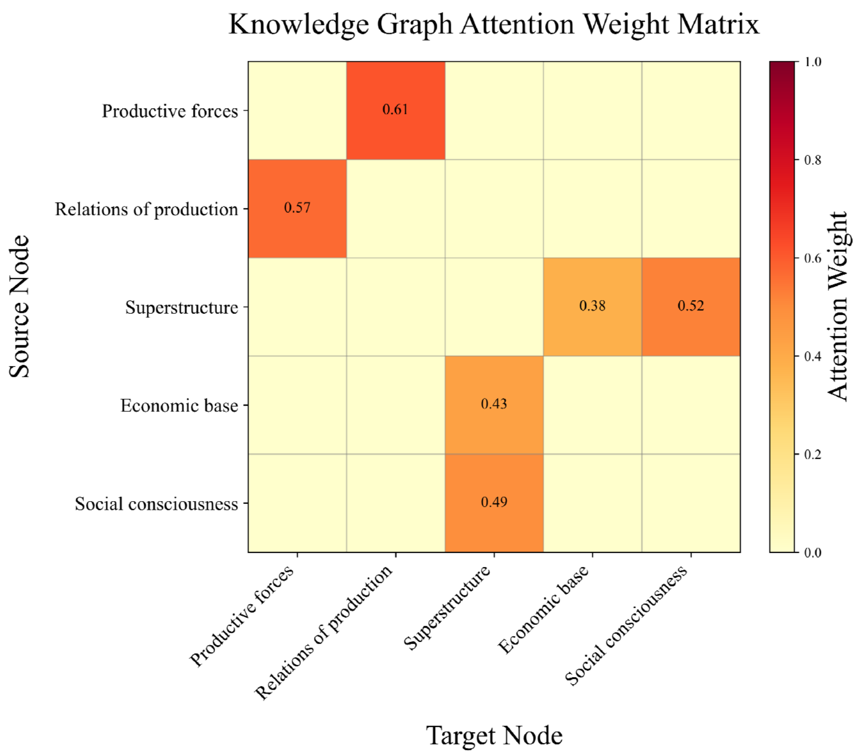
As shown in Table 2, the ablation experiment results revealed the contributions of each module: removing the KGAT module resulted in a 3.2% decrease in AUC-ROC and a 6.8% decrease in EDR. After removing the knowledge graph embedding, F1 score decreased by 5.1%, confirming the auxiliary value of structured knowledge representation. When replacing KGAT with regular GCN, F1 score decreased by 4.3%, indicating the advantage of attention mechanism in capturing complex logical relationships. Figure 4 shows the distribution of attention weights for knowledge points, with a weight of 0.61 for ‘productivity determines production relations’ and 0.57 for ‘production relations counteraction’, which conforms to theoretical logic but does not overfit, demonstrating the robustness of the model.

Table 2 Analysis of ablation experiment

Model variants	*AUC-ROC*	*F1-score*	*EDR (%)*
Remove KGAT	0.73	0.68	65.5
Remove knowledge graph embedding	0.71	0.67	61.2
Replace with GCN	0.74	0.68	67.8
Complete model	0.76	0.71	72.3

The experimental results show that LSTM-KGAT exhibits moderate advantages in ideological and political knowledge prediction tasks, especially in long-term prediction and complex relationship modelling, by integrating temporal behaviour characteristics with logical reasoning of knowledge graphs. However, models are sensitive to data quality and the coverage of knowledge graphs directly affects performance. In the future, further optimisation of knowledge extraction algorithms and expansion of training data scale are needed.

Figure 4 Knowledge graph attention weight graph (see online version for colours)



5 Conclusions

In the context of the digital transformation of higher education and the deep integration of ideological and political courses, this study proposes a hybrid model based on deep tracing and knowledge graph to address the static, closed, and unexplainable shortcomings of traditional ideological and political knowledge evaluation methods. By integrating the temporal characteristics of learning behaviour with the logic of ideological knowledge, the model has achieved dynamic and interpretable prediction of ideological and political knowledge mastery, providing technical support for implementing the fundamental task of cultivating virtue and talent.

Future work will focus on three major directions: expanding interdisciplinary knowledge graphs (such as integrating party history and current political hotspots), optimising time-series modelling algorithms with non-uniform time intervals, and exploring privacy preserving knowledge tracking under federated learning frameworks.

This study provides an innovative and practical solution for the intelligentisation of ideological and political education, and its methodological framework can be extended to other fields of humanities and social sciences. It has positive significance for promoting the ethical and humanistic development of ‘artificial intelligence + education’.

Acknowledgements

This work is supported by the Zhejiang Philosophy and Social Sciences Planning Project (No. 25GXSZ017YBM).

Declarations

All authors declare that they have no conflicts of interest.

References

Ataeva, O., Serebryakov, V. and Tuchkova, N. (2024) 'ontology-driven knowledge graph construction in the mathematics semantic library', *Pattern Recognition and Image Analysis*, Vol. 34, No. 3, pp.448–455.

Chandak, P., Huang, K. and Zitnik, M. (2023) 'Building a knowledge graph to enable precision medicine', *Scientific Data*, Vol. 10, No. 1, p.67.

Chen, P., Lu, Y., Zheng, V.W., Chen, X. and Yang, B. (2018) 'Knowedu: a system to construct knowledge graph for education', *IEEE Access*, Vol. 6, No. 3, pp.31553–31563.

Du, K., Yang, B., Wang, S., Chang, Y., Li, S. and Yi, G. (2022) 'Relation extraction for manufacturing knowledge graphs based on feature fusion of attention mechanism and graph convolution network', *Knowledge-Based Systems*, Vol. 255, No. 4, p.109703.

Jiao, L., Wang, D., Bai, Y., Chen, P. and Liu, F. (2021) 'Deep learning in visual tracking: a review', *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 34, No. 9, pp.5497–5516.

Li, N., Shen, Q., Song, R., Chi, Y. and Xu, H. (2022) 'MEduKG: a deep-learning-based approach for multi-modal educational knowledge graph construction', *Information*, Vol. 13, No. 2, p.91.

Li, Q., Wang, D., Feng, S., Niu, C. and Zhang, Y. (2021) 'Global graph attention embedding network for relation prediction in knowledge graphs', *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 33, No. 11, pp.6712–6725.

Lindemann, B., Maschler, B., Sahlab, N. and Weyrich, M. (2021) 'A survey on anomaly detection for technical systems using LSTM networks', *Computers in Industry*, Vol. 131, No. 12, p.103498.

Liu, X., Xiantong, Z. and Starkey, H. (2023) 'Ideological and political education in Chinese Universities: structures and practices', *Asia Pacific Journal of Education*, Vol. 43, No. 2, pp.586–598.

Liu, Y., Yang, S., Xu, Y., Miao, C., Wu, M. and Zhang, J. (2021) 'Contextualized graph attention network for recommendation with item knowledge graph', *IEEE Transactions on Knowledge and Data Engineering*, Vol. 35, No. 1, pp.181–195.

Nguyen, H.D., Tran, K.P., Thomassey, S. and Hamad, M. (2021) 'Forecasting and anomaly detection approaches using LSTM and LSTM autoencoder techniques with the applications in supply chain management', *International Journal of Information Management*, Vol. 57, No. 5, p.102282.

Peng, C., Xia, F., Naseriparsa, M. and Osborne, F. (2023) 'Knowledge graphs: opportunities and challenges', *Artificial Intelligence Review*, Vol. 56, No. 11, pp.13071–13102.

Sagheer, A., Hamdoun, H. and Youness, H. (2021) 'Deep LSTM-based transfer learning approach for coherent forecasts in hierarchical time series', *Sensors*, Vol. 21, No. 13, p.4379.

Santos, A., Colaço, A.R., Nielsen, A.B., Niu, L., Strauss, M., Geyer, P.E., Coscia, F., Albrechtsen, N.J.W., Mundt, F. and Jensen, L.J. (2022) 'A knowledge graph to interpret clinical proteomics data', *Nature Biotechnology*, Vol. 40, No. 5, pp.692–702.

Seabe, P.L., Moutsinga, C.R.B. and Pindza, E. (2023) 'Forecasting cryptocurrency prices using LSTM, GRU, and bi-directional LSTM: a deep learning approach', *Fractal and Fractional*, Vol. 7, No. 2, p.203.

Shahid, F., Zameer, A. and Muneeb, M. (2021) 'A novel genetic LSTM model for wind power forecast', *Energy*, Vol. 223, No. 11, p.120069.

Shen, T., Zhang, F. and Cheng, J. (2022) 'A comprehensive overview of knowledge graph completion', *Knowledge-Based Systems*, Vol. 255, No. 5, p.109597.

Sun, K., Liu, Y., Guo, Z. and Wang, C. (2016) 'Visualization for knowledge graph based on education data', *International Journal of Software & Informatics*, Vol. 10, No. 3, p.23421.

Tatsunami, Y. and Taki, M. (2022) 'Sequencer: Deep LSTM for image classification', *Advances in Neural Information Processing Systems*, Vol. 35, No. 6, pp.38204–38217.

Wen, X. and Li, W. (2023) 'Time series prediction based on LSTM-attention-LSTM model', *IEEE Access*, Vol. 11, No. 3, pp.48322–48331.

Zhong, L., Wu, J., Li, Q., Peng, H. and Wu, X. (2023) 'A comprehensive survey on automatic knowledge graph construction', *ACM Computing Surveys*, Vol. 56, No. 4, pp.1–62.

Zhu, X., Li, Z., Wang, X., Jiang, X., Sun, P., Wang, X., Xiao, Y. and Yuan, N.J. (2022) 'Multi-modal knowledge graph construction and application: a survey', *IEEE Transactions on Knowledge and Data Engineering*, Vol. 36, No. 2, pp.715–735.