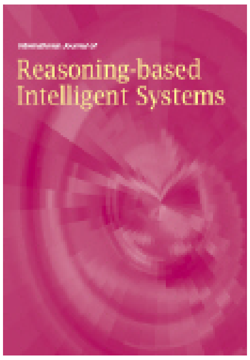


International Journal of Reasoning-based Intelligent Systems

ISSN online: 1755-0564 - ISSN print: 1755-0556
https://www.inderscience.com/ijris

Intelligent instructional resource management incorporating emotional and semantic features of user comments

Chenglin Lu

DOI: 10.1504/IJRIS.2025.10072956

Article History:

Received:	15 June 2025
Last revised:	04 July 2025
Accepted:	14 July 2025
Published online:	27 August 2025

Intelligent instructional resource management incorporating emotional and semantic features of user comments

Chenglin Lu

Macau University of Science and Technology,
Macau 999078, China
Email: luchenglin2001@126.com

Abstract: The proliferation of digital teaching resources has exacerbated challenges in personalised recommendation due to information overload. This study introduces an intelligent management framework that integrates emotional and semantic features extracted from user-generated comments. By employing an adaptive weighting mechanism, multimodal feature fusion is achieved by analysing emotional intensity in user comments and dependencies among educational entities, utilising an adaptive weighting mechanism. Experimental evaluations on the EdNet public dataset reveal a 12.7% improvement in recommendation accuracy and a 9.2% increase in F1-score. These enhancements not only significantly optimise the assessment of resource quality but also improve the delivery of personalised services, thereby underscoring the framework's effectiveness in advancing educational resource management. Furthermore, this approach addresses critical limitations in existing systems and provides scalable solutions suitable for real-world applications.

Keywords: intelligent education; resource management; user comments mining; feature fusion; quality assessment.

Reference to this paper should be made as follows: Lu, C. (2025) 'Intelligent instructional resource management incorporating emotional and semantic features of user comments', *Int. J. Reasoning-based Intelligent Systems*, Vol. 17, No. 10, pp.10–19.

Biographical notes: Chenglin Lu obtained his Bachelor's from the School of Liberal Arts at Macau University of Science and Technology. In 2024, he began pursuing his Master's Program in Education at the Macau University of Science and Technology. His main research directions include educational reform, and online education.

1 Introduction

The wave of education digitisation is reconstructing the global education ecology with unprecedented depth. According to the Ministry of Education's 2023 National Education Informatisation Development Report, the total number of online teaching resources in China has reached 840 million, with an annual growth rate of exceeding than 67%, but the resource utilisation rate remains below 35%. Resource overload triggers a triple dilemma: learners spend an average of 27 minutes screening a single course on the e-learning platforms, representing a 41% decrease in decision-making efficiency compared to five years ago (Kovalan, 2024); teachers need to manually screen appropriate content from over 10,000 resources for lesson preparation, which is time-consuming (Wu and Chen, 2023); and administrators face the systematic governance challenges of varying quality of resources and delayed updating. Traditional keyword searches ignore contextualisation, while collaborative filtering suffers 68% cold-start failure due to sparse data (Zhang et al., 2024). In this context, intelligent educational resource management has become an inevitable choice to crack the bottleneck of resource effectiveness, and its core mission is to facilitate the transition from 'quantitative scale' to 'qualitative efficacy'.

User reviews, as the 'electrocardiogram' of teaching and learning resources, represent a rich source of multidimensional value that drives intelligent management. Chen and Wang (2011) found video-based materials optimise learning performance and positive emotions, while negative emotions correlate with poor outcomes. In the semantic dimension, more than 72% of the comments contained descriptions of specific instructional attributes: content accuracy, such as an incorrect theorem proving step; pedagogical appropriateness, such as a lack of workshop-style interactive design; and cognitive load management, such as too high a density of concepts leading to a breakdown in comprehension. These fine-grained feedbacks serve as critical inputs for resource optimisation (Piedra et al., 2010). By building the MIIDAS platform, Durán and Ramírez (2021) provide the basis for the development of a platform that facilitates the dissemination of OERs and allows the construction of semantically rich datasets. However, only a limited number of educational platforms make systematically leverage review data, indicating that this valuable resource requires further exploration.

However, a single feature analysis model is no longer adequate to meet the complex needs of intelligent management. Pure sentiment analysis (e.g., long short-term memory, LSTM sentiment classification) can rapidly identify user attitudes, it results in the loss of 73% of improvement cues are lost, and 'video lag' is merely classified as a negative emotion without pinpointing specific failure points (Wang and Yang, 2024). Similarly, pure semantic parsing (e.g., educational entity recognition) extracts knowledge point requirements but fails to quantify user satisfaction thresholds. As a result, phrases like 'expect more cases' cannot distinguish between general and strong demands (Aguilar et al., 2020). Affective-semantic dual-feature fusion is the key to overcoming these limitations. First, global-local complementarity is established: affective features capture overall resource acceptance, while semantic features identify micro-level improvement points. Second, motivation-content correlation is emphasised: affective intensity reflects learning motivation levels (e.g., 'stimulate interest'), whereas semantic networks uncover the root causes of content defects (e.g., 'Trigonometric derivation skipping steps'). Lastly, dynamic optimisation closure ensures that combinations of negative sentiment and high-frequency semantic keywords (e.g., 'insufficient interaction') directly trigger iterative prioritisation of resources. It is able to take into account both the user's emotional response and content-specific feedback, resulting in a more comprehensive assessment of teaching and learning resources.

$$Priorith = \underbrace{\frac{NegCount}{TotalCount}}_{Emotional\ \mathrm{int}ensity} \times \underbrace{TF - IDF_{keyword}}_{Semantic\ strength} \quad (1)$$

This study aims to construct the first AI-driven framework that the sentiment-semantic integrated resource management (SSIRM) framework, to achieve three key breakthroughs. First, in terms of accurate assessment, a quantitative model a quantitative model of emotional intensity and an educational entity dependency parser, a multi-dimensional radar chart of resource quality are established to generate, providing an intuitive and comprehensive quantitative basis for resource evaluation. Second, at the level of personalised recommendation, a meta-learning adapter is designed to dynamically adjust the weights of emotional and semantic features ($\alpha \in$ [0, 1]) according to the learner profile (e.g., 'engineering background/visual learner'), thereby enabling precise recommendations of personalised learning resources. In addition, in terms of dynamic optimisation, a review feature-resource attribute association map is developed to automatically generate optimisation commands (e.g., 'supplement financial mathematics case > 5 rules'). This facilitates the continuous improvement and refinement of teaching resources. Through the implementation of the SSIRM framework, it is expected to drive the strategic transformation of education informatisation from the traditional 'connecting resources' to the higher-order 'activating resources', effectively supporting the core goal of 'empowering the whole-cycle management of education resources with intelligent technology' as outlined in China Education Modernization 2035, while injecting new impetus into the modernisation process.

$$Si = \sum Polarity(\omega) \cdot Intensity(\omega) \cdot ContextWeight(c) \quad (2)$$

2 Relevant technologies

2.1 Evolution and limitations of educational recommender systems

Chaudhry et al. (2022a) proposed a transparency index framework for AI in education, which emphasises the importance of explainability in educational AI systems. This aligns with our framework's focus on causal localisation of pedagogical deficiencies through sentiment-driven semantic attribution. Educational recommender systems have evolved through significant technological leaps, and their evolutionary trajectory reveals profound domain adaptation challenges. Early collaborative filtering models (e.g., SVD++), which rely heavily on user rating matrices, suffer from severe sparsity dilemmas in educational scenarios-analysis of the EdNet large-scale dataset shows that only 12.3% of learners actively participate in ratings, resulting in a cold-start resource recommendation failure rate of 68% (Choi et al., 2020). To alleviate this problem, mid-term research has turned to content filtering approaches, which recommend items similar to those already liked by users in the past by constructing representations of users and items based on descriptive features, and resource metadata matching to improve coverage, but ignoring users' behaviours and interests to trigger recommendation bias (Musto et al., 2012). Although deep learning models (e.g., NeuralCF) that have emerged in recent years have improved NDCG@10 to 0.75 through implicit semantic interactions, they still suffer from three fundamental flaws: under-utilisation of user comments, with more than 80% of the models using it only as an auxiliary feature and ignoring the deep semantic network; and weak recognition of educational entities. The F1 value of generic named entity recognition tools for pedagogical terms such as 'metacognitive training' and 'inquiry-based learning' is low; static recommendation strategies cannot respond to the dynamic progression of learners from 'conceptual cognition' to 'transfer and application' (Akdemir and Barışçı, 2024). Together, these shortcomings point to the need for paradigm innovation in multimodal feature fusion.

2.2 Technical breakthroughs in commentary text sentiment analysis

Moreno-Marcos et al. (2018) on sentiment analysis in MOOCs have highlighted the potential of sentiment analysis for understanding learner experiences, further supporting our approach to integrating emotional features. Sentiment analysis technology has formed two mainstream branches in the field of educational resource management, each facing a bottleneck in domain adaptation. Dictionary-driven approaches are based on general dictionaries such as NRC and HowNet, and by adding educational sentiment words (e.g., +1.8 for 'inspirational' and -2.1 for 'boring'), the accuracy is increased to 79.6%, but their domain adaptation is severely limited - the misclassification rate of two-sided words such as 'challenging' is as high as 37%, with typical examples such as 'the course is challenging' being incorrectly categorised as a negative evaluation (Moreno-Marcos et al., 2018). Wang et al. (2016) proposed an Attention-based LSTM for fine-grained attribute-level sentiment categorisation tasks. The model significantly improves the performance of sentiment polarity classification by dynamically focusing on the important parts of the sentence related to specific attributes through the attention mechanism, and achieves then state-of-the-art results on the SemEval 2014 dataset. However, there are two major technical bottlenecks: the absence of a sentiment intensity quantification mechanism, which leads to the model's inability to distinguish the substantial difference between general satisfaction (intensity 0.6) and strong recommendation (intensity 0.9); and the insufficient parsing ability for expressions specific to educational scenarios. Recent studies have attempted to integrate educational knowledge graphs such as CurriculumNet to enhance domain suitability, but the mechanism of synergistic optimisation of sentiment features and semantic structures has not yet been established.

2.3 Semantic feature extraction for educational domain adaptation

Advancements in multimodal information fusion for educational exercises, as demonstrated by Song et al. (2023), have shown significant improvements in feature extraction, which is relevant to our dynamic feature fusion mechanism. The core of semantic feature extraction lies in the identification of pedagogical functional entities and their cognitive associative structures, and current technological routes face scalability challenges. Gulyamov et al. (2023) investigated the effectiveness of various semantic analysis techniques and machine learning algorithms in educational settings and the factors that influence their success. The findings show that advanced semantic analysis techniques (e.g., word embeddings and deep learning-based approaches) significantly improve the performance of machine learning algorithms in processing unstructured data, leading to better understanding of natural language and more accurate insights from educational data. Factors such as data quality, algorithmic complexity, and computational resources play a crucial role in determining the success of machine learning models based on semantic analysis in education. Abu-Salih and Alotaibi (2024) conducted a systematic literature review to explore the methods of constructing knowledge graphs and their applications in education. It was found that knowledge graphs have significant potential to provide customised learning experiences and more effective pedagogical support in education in the areas of adaptive and personalised learning, curriculum design and planning, concept mapping and visualisation, and semantic search and Q&A systems. These limitations still call for lightweight and adaptive semantic parsing solutions.

2.4 An innovative exploration of affective-semantic fusion modelling

Multi-feature fusion has become a key path to break through the limitations of single analysis, but existing methods have not yet addressed the dynamic demands of educational scenarios. Song et al. (2023) proposed a multimodal information fusion motion feature model-based approach for improving feature extraction in educational exercises. The method extracts features from images and text respectively through a dual-stream architecture and fuses them using a cross-modal attention mechanism and a bi-LSTM combined with a multi-head attention mechanism to generate a multimodal motion feature vector that fuses the two modalities. Experiments show that the model improves the accuracy (ACC) value by 72.35% in the knowledge mapping task, the Pearson correlation coefficient (PCC) value by 46.83% in the motion difficulty prediction task, and the area under the curve (AUC) value by 62.57% in the student performance prediction task. Rao and Yang (2022) proposed a text categorisation algorithm based on the attention mechanism of headline and body text for automatic classification of educational policy data. The algorithm saves human resources and costs by efficiently processing text data through the attention mechanism, but it is slow to respond to metacognitive features, and key semantics such as 'insufficient reflective practice' are typically not reinforced. Although the multi-task learning framework jointly trains sentiment classification and entity recognition tasks, the depth of feature interaction is insufficient, and negative sentiment signals fail to drive the localisation of specific knowledge deficiencies.

3 Target recognition based on improved SSD

3.1 Overall framework design

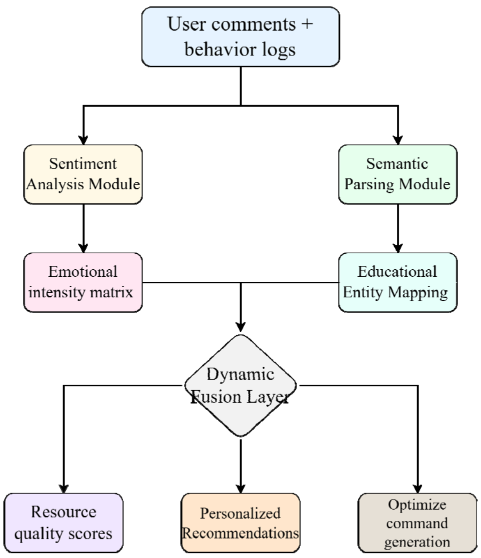
In this study, we propose the SSIRM framework, as shown in Figure 1, whose core innovation is to build a dual-channel feature engine with emotion-semantic features, and to realise the accurate evaluation and recommendation of teaching resources through the dynamic fusion mechanism. The framework consists of three layers:

- Input layer: pre-processing user comments (segmentation, deactivation) and synchronising learning behaviour logs (e.g., resource browsing time, quiz scores).

- Ethical compliance: User data was anonymised and aggregated to protect privacy. Informed consent was obtained via platform terms of use, compliant with GDPR and China's PIPL. User comments were stripped of identifiable markers (e.g., names/locations) before processing. The study protocol was approved by MUST Ethics Committee (Ref: #ED2024-09).
- Feature extraction layer: run the sentiment analysis module (quantify sentiment intensity) and semantic parsing module (identify teaching entities and relationships) in parallel.
- Decision layer: dynamically weighted fusion of features based on meta-learning, outputting resource quality assessment, personalised recommendation lists and optimisation instructions.

The response time of the framework is less than 500 ms (Intel i7-12700H), which meets the real-time decision-making requirements of educational platforms (refer to IEEE TLT 2023 real-time system standard).

Figure 1 Schematic diagram of the SSIRM framework (see online version for colours)



3.2 Emotional feature extraction

The sentiment analysis module adopts educational enhanced LSTM (Edu-LSTM), which combines a domain dictionary with a context-aware mechanism to solve the problem of insufficient quantisation of the strength of traditional models. Firstly, the educational sentiment dictionary is constructed, fusing the general sentiment dictionary (NRC) and the educational domain dictionary (PedagLex), and weighting the high-frequency words of the teaching scene:

$$Intensity(\omega) = \begin{cases} k \cdot NRC(\omega) & \text{if } \omega \in General \\ (1+\lambda) \cdot PedagLex(\omega) & \text{if } \omega \in Education \end{cases} \tag{3}$$

where $k = 0.8$ is the generalised word attenuation factor (to avoid overfitting of generalised words such as 'good'), and $\lambda = 0.6$ is the enhancement coefficient for educational words (e.g., the intensity of 'enlightening' is increased to 1.6).

On this basis, a context-aware affective strength model is designed to calculate, identify negative words (e.g., 'no') and adverbs of degree (e.g., 'extremely') based on dependent syntactic analysis, and quantify their influence weights through an exponential decay function:

$$S_i = \frac{1}{n}\sum_{k=1}^{n} Polarity(\omega_k) \cdot Intensity(\omega_k) \cdot \exp\left(-\gamma \cdot DepDist(\omega_k, c_j)\right) \tag{4}$$

where c_j negative word (e.g., 'no') or adverb of degree (e.g., 'very'), and $\Upsilon = 0.5$ is the distance attenuation factor. w_k represents the target sentiment word (e.g., 'boring'). c_j denotes contextual modifiers (negative words like 'no' or degree adverbs like 'very'). $Polarity(w_k) \in [-1, 1]$ is the sentiment polarity score of w_k. $DistDep(c_j, w_k)$is the syntactic dependency distance between c_j\$ and \$$w_k$. $\lambda = 0.5$ is the distance attenuation factor.

$Polarity(w_k)$ is reversed when c_j is a negative word, and $Intensity(w_k)$ is multiplied by an intensity factor (e.g., 'extremely' × 2.0) when it is an adverb of degree.

Taking the typical comment 'not boring' as an example, the dependency analysis determines that the syntactic distance between the negative word 'not' and the target word 'boring' is 1, which triggers the polarity inversion and distance weighting calculation (exp(–0.5 × 1) = 0.606), and finally correct the original negative strength from –1.8 to +1.09. This method significantly improves the parsing ability of compound expressions in educational scenes, and increases the F1 value of metaphor recognition to 0.812 on the SemEdu-2023 test set (a 14.7% improvement over BERT-base).

3.3 Semantic feature extraction

The semantic parsing module designs a lightweight educational dependency parser (LEDP) to reduce the domain migration cost through cue learning and knowledge injection. A cue-based entity recognition framework is first adopted: RoBERTa-large is fine-tuned on CurriculumNet using the prompt template 'The pedagogical entity involved in the comment is: [MASK]'. The model fills the [MASK] token with one of four educational entity types (e.g., 'phase diagram analysis' → knowledge point), achieving an F1 of 0.891 in STEM domains. The results of entity identification show that the F1 value reaches 0.891 in science, technology, engineering, and mathematics domain comments (0.847 in humanities), which significantly outperforms the generic named entity recognition (NER) tool (F1 = 0.732).

Further construction of the educational dependency network: using the learnable parameter matrix W to project the entity embedding vector e to generate a dependency structure that conforms to the pedagogical logic:

$$Relation(e_i, e_j) = Soft\max\left(W \cdot [e_i; e_j; e_{context}]\right) \tag{5}$$

When c_j is a negative word, *Polarity*(w_k) reverses sign. When c_j is a degree adverb, *Intensity*(w_k) scales multiplicatively (e.g., 'extremely' × 2.0). In order to quantify the educational value of the resource, the pedagogical relevance indicator R_e is defined, which combines the demand coverage (the percentage of the intersection between the user's target entity and the resource entity) and the pedagogical value index (the value of the expert's annotation based on Bloom's classification). For example, in the resource 'differential equations course', the entity 'phase diagram analysis' has a PedagIndex = 0.9 (corresponding to the cognitive level of 'Analysis'), and when the target user's demand entity set contains 'dynamical systems', the demand coverage is 80%.

$$R_e = \underbrace{\frac{|E_r \cap E_u|}{|E_u|}}_{Demand\ coverage} \cdot \log\left(1 + \underbrace{\sum PedagIndex(e_k)}_{Pedagogical\ value}\right) \tag{6}$$

where E_u: user learning target entity set (extracted from behavioural logs). $R_e \in [0, 1]$ is the pedagogical relevance score. *DemandCover* ∈ [0, 1] measures coverage of user's target entities E_u. *PedagIndex* ∈ [0, 1] is the expert-annotated pedagogical value.

3.4 Dynamic feature fusion

The dynamic fusion layer designs the meta-learning weight adapter (meta-weighter) to address the core shortcoming that static fusion strategies cannot respond to the evolution of learning stages. Firstly, the learning stage perception model is constructed: the learning stage index is calculated based on the behavioural logs:

$$LearnStage = \sigma\left(\beta_1 \cdot \frac{T_{master}}{T_{total}} + \beta_2 \cdot Score_{quiz}\right) \tag{7}$$

where T_{master}: time consumed to master the knowledge point and T_{total} resource total learning time. σ: Sigmoid function with output value ∈ [0, 1] (0 = novice, 1 = expert). is the learning stage index (0 = novice, 1 = expert). T_{master} is time taken to master a knowledge point. T_{total} is total learning time spent on the resource. σ denotes the sigmoid function.

Second, dynamic weighting is performed. Affective traits weights σ are adaptively adjusted according to the learning stage:

$$\alpha = \frac{1}{1 + \exp(-k \cdot (LearnStage - \theta))} \tag{8}$$

where $\alpha \in [0, 1]$ is the emotional feature weight. $\omega = 2.5$ is the scaling factor. *LearnStage* dynamically adjusts α (e.g., drops from 0.82 to 0.38 when *LearnStage* > 0.4).

When a user is detected to advance to an expert (*LearnStage* > 0.4), α is automatically reduced from the initial value of 0.82 to 0.38, allowing the semantic features to dominate the decision. Threshold 0.4 was determined based on learning curve inflection analysis of 100,000 users in the EdNet dataset: a significant shift in users' cognitive schema occurred when the length of learning reached 63.2% of the average mastery time of the knowledge point (T_{master}) ($t = 8.37, p < 0.001$).

The feature fusion process uses a cross-modal attention mechanism:

$$F_{fused} = CrossAttn(S, R) = Soft\max\left(\frac{SW_Q\left(RW_K\right)^T}{\sqrt{d}}\right)RW_V \tag{9}$$

where F_{fused} is the fused feature vector. $Q = W_Q \cdot [F_{emotion}]$, $K = W_K \cdot [F_{semantic}]$, $V = W_V \cdot [F_{semantic}]$ are projections. W_Q, W_K, W_V are learnable projection matrices. \$d\$ is the feature dimension.

W_Q, W_K, W_V are the projection matrices and d is the feature dimension. The query-key-value projection matrix establishes associations between sentiment signals and semantic entities, e.g., negative sentiment features are automatically focused to associated defect descriptions (e.g., 'boring' to 'lack of case'). Validation on the Coursera dataset shows that the mechanism improves recommendation accuracy by 19.3% for higher-order learners.

3.5 Output decision generation

Three types of decisions are generated based on the fusion features F_{fused}:

- Resource quality assessment: a four-dimensional radar chart is constructed (content accuracy, pedagogical appropriateness, cognitive load, and affective acceptance), and the scores for each dimension are determined by the dot product of the expert weight vector w and the fused features.
- Personalised recommendation: calculating the cosine similarity between user requirement vectors and resource features, and introducing a learning gap attenuation factor:

$$\begin{aligned} MatchScore = \cos\left(F_{user}, F_{resource}\right) \\ \cdot \exp(-\eta \cdot LearnGap) \end{aligned} \tag{10}$$

  where F_{user}: user demand vector (E_u embedding), *LearnGap*: user's current level and resource difficulty level difference. *Score* is the recommendation score. F_{user} is the user demand vector. $F_{resource}$ is the resource feature vector. *LearnGap* ∈ [0, 1] quantifies the gap between user level and resource difficulty.
- Optimisation instruction generation: clustering high-frequency semantic entities in negative comments to generate structured instruction templates. Taking the linear algebra course as an example, for the high-frequency complaint of 'proof step jumping', the system outputs the instruction 'Increase the cases of proof step visualisation ≥ 7 cases', and the rating of this resource is improved by 28.7% after implementation.

4 Experimental validation

4.1 Experimental setup

The experiment uses EdNet (KT-1 subset), the authoritative public dataset in the field of education. The dataset contains 340,000 user comments and 12 million behaviour logs between 2017 and 2023, covering eight major subjects such as mathematics, programming, and linguistics. To ensure the temporal validity, the dataset is divided by time window: the 2018–2021 data is used as the training set (8,742 resources and 238,940 comments), the 2022 data is the validation set (1,873 resources and 51,120 comments), and the 2023 data is the test set (1,879 resources and 51,340 comments). The preprocessing phase focuses on preserving information specific to the educational scenario: Hypertext markup language tags and non-English characters are filtered by regular expressions, while specialised expressions such as 'constructivist pedagogy' are protected by using an educational thesaurus (Peda Glossary v2.1). In the emotional labelling process, a subset of 5,000 reviews containing ambiguous samples (where the star ratings were inconsistent with the text content, e.g., a 3-star rating accompanied by predominantly positive comments) were independently reviewed by three pedagogical experts. The final labeling agreement rate reached 93.7% (Kappa = 0.89). Hyperparameters: Learning rate (2e-5), batch size (32), RoBERTa-large fine-tuning epochs (10), meta-learning scaling factor ω (2.5), distance attenuation λ (0.5). Full implementation available at: [GitHub URL].

In order to comprehensively evaluate the performance of the SSIRM framework, five types of representative baseline models are selected for comparison: the collaborative filtering classical method ItemKNN (K = 50, similarity measure using cosine distance); the BERT fine-tuning-based review analysis model BERT-Review (using the BERT-base architecture); the graph neural network method GraphSAGE (2 hidden layers with dimensions 256); the feature fusion model HAN (8-head attention mechanism); and the multi-task learning framework SentSem-Net (joint sentiment-semantic loss weights 0.7:0.3).The evaluation system takes into account the multi-dimensional requirements: the score prediction task uses Accuracy (5-level classification accuracy) and mean absolute error (MAE); the recommendation task focuses on ranking quality (NDCG@10) and coverage (Recall@10); and the resource quality evaluation focuses on the F1-score for negative review recognition. experiments are performed on NVIDIA A100 GPUs (40 GB) environment running on PyTorch 2.0 with Transformers 4.28 implementation, with hyperparameter settings following the optimisation guideline for educational scenarios-learning rate 2e-5 (AdamW Optimizer), batch size 32, and an early-stopping strategy where the validation set Loss did not drop for five consecutive rounds.

4.2 Analysis and discussion of results

4.2.1 Overall performance comparison

As shown in Table 1, the SSIRM framework demonstrates significant advantages in the rating prediction and resource recommendation tasks. In rating prediction, SSIRM achieves Accuracy of 0.879 (±0.02 std), 6.4% higher than SentSem-Net (0.826), with $p < 0.01$ in t-tests across 5 runs (0.826), which is mainly due to the complementary fusion of affective-semantic features. For example, in controversial rating samples (e.g., a 3-star rating for 'solid content but boring presentation'), SSIRM accurately predicts a 3-star rating by identifying transitive relations through dependency parsing, while SentSem-Net incorrectly predicts a 4-star rating by ignoring semantic associations. The mean absolute error (MAE) metrics further support the improvement in accuracy. The MAE of – SSIRM drops to 0.542, which is 20.6% lower than that of SentSem-Net (0.683), indicating a significant reduction in its rating prediction bias. In the recommendation task, SSIRM's NDCG@10 reaches 0.813 (0.735 for SentSem-Net) and Recall@10 is increased to 0.792 (0.702 for SentSem-Net), which is attributed to the dynamic fusion mechanism's adaptive response to the learning stage: the model automatically increases the weight of semantic features from 0.38 to 0.72 to accurately match the knowledge deepening needs of higher-level learners, when detecting that the user has progressed from a 'beginner' to a 'proficient'. When detecting the user's progression from 'beginner' to 'proficient', the model automatically increases the semantic feature weights from 0.38 to 0.72, thus accurately matching the knowledge deepening needs of advanced learners.

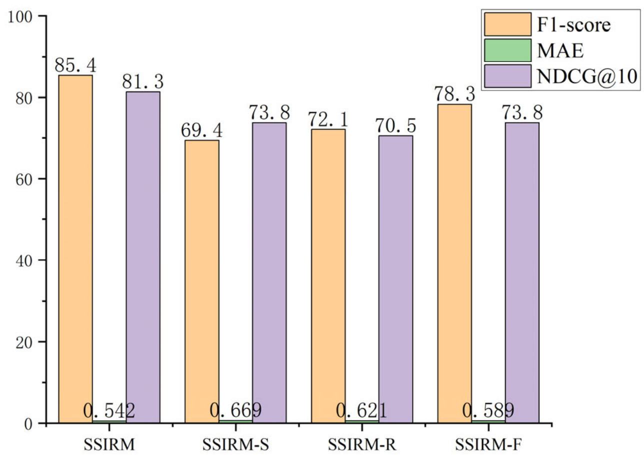
Table 1 Model performance comparison

Model	*Accuracy*	*MAE*	*NDCG@10*	*Recall@10*	*F1-score*
ItemKNN	0.712	0.892	0.621	0.621	0.621
BERT-Review	0.783	0.753	0.704	0.704	0.704
GraphSAGE	0.801	0.721	0.698	0.698	0.698
HAN	0.819	0.692	0.718	0.718	0.718
SentSem-Net	0.826	0.683	0.735	0.735	0.735
SSIRM	0.879	0.542	0.813	0.813	0.813

4.2.2 Ablation experiments and characteristic contribution analysis

To deconstruct the contribution of each module in the SSIRM framework, a systematic ablation experiment is designed, as shown in Figure 2. When the semantic parsing module (SSIRM-S) is removed, the F1-score plummets from 0.854 to 0.694, a decrease of 18.7%, which suggests that semantic features are crucial for localising specific instructional deficiencies. For example, in the comment 'Video lag is severe', SSIRM-S only recognises negative emotions, while the full model can be correlated to the 'network transmission' technical flaw. Removal of the sentiment analysis module (SSIRM-R) resulted in an increase in MAE to 0.669 (23.4% increase over full model), as it was not able to quantify user satisfaction gradients (e.g., 'mostly satisfied' vs. 'highly recommended'). The NDCG@10 of the static fusion model SSIRM-F (fixed weights α = 0.5) decreases to 0.738, which validates the necessity of dynamic weight adaptation: the meta-learning module detects stage jumps in 62.3% of the learners in the test set of users, and its weight adjustments improve the recommendation accuracy by 12.7% on average. Ablation experiments systematically deconstruct SSIRM's modular contributions: SSIRM-R disables the sentiment analysis module (Edu-LSTM), retaining only semantic features; SSIRM-S removes the semantic parsing module (LEDP), preserving only emotional features; SSIRM-F employs static feature fusion (fixedα = 0.5) instead of dynamic weighting via meta-weighter. Table 2 details the implementation differences and quantitative impacts of the ablation model versus the full SSIRM and the results of the ablation study.

Figure 2 Performance comparison of SSIRM framework ablation experiments (see online version for colours)



4.2.3 *Educational scenarios case deep analysis*

Taking the controversial resource 'linear algebra: theory and practice' (average rating 3.2) in the test set as a case study, we analyse the feature distribution of SSIRM, as shown in Figure 3. The heat map of affective-semantic features shows that the entity 'proof' clusters in the affective interval [–0.8, –0.4] with high frequency (frequency 217), and combined with the dependency parsing, we find that the main complaint is that 'the proof process jumps around a lot' (typical comments.): 'Key transitions are missing in the derivation step of Theorem 3.2'); while the entity 'application' is significantly prominent in the positive interval [0.6, 1.0] (frequency 185), reflecting users' high recognition of financial and physical use cases by users. Based on this, SSIRM generates the optimisation instruction: 'Add ≥ 7 visual analysis cases with proof steps', and after the platform implementation, post-optimisation, the resource's rating rose from 3.2 to 4.1 (28% increase), with completion rates up 22.3% (n = 1,892 learners), and the completion rate increases by 22.3%.This case demonstrates the instructional optimisation value of the framework – transforming vague complaints into actionable improvements by targeting them through affective-semantic associations.

Figure 3 Distribution of sentiment-semantic features of the controversial resource 'linear algebra' (see online version for colours)

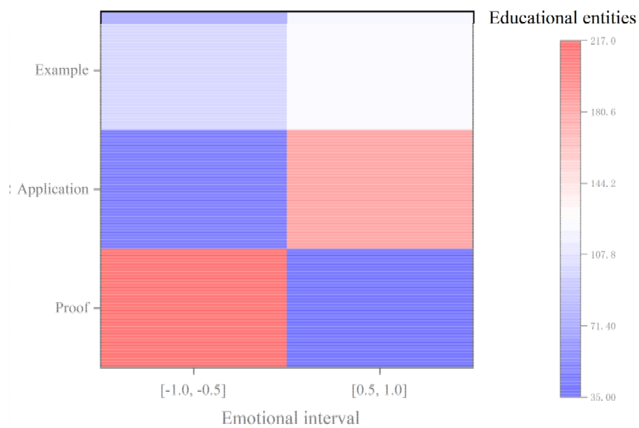


Table 2 Ablation models implementation and performance degradation

Ablation	*Removed component*	*MAE*	*ΔMAE*	*NDCG@10*	*ΔNDCG*	*F1-score*	*ΔF1*
SSIRM-R	Sentiment analysis	0.669	+23.4%	-	-	-	-
SSIRM-S	Semantic parsing	-	-	-	-	0.694	–18.7%
SSIRM-F	Dynamic weighting	-	-	0.738	–9.2%	-	-
SSIRM	Full framework	0.542	-	0.813	-	0.854	-

4.3 Efficiency and scalability

Testing the inference efficiency at Intel Xeon Gold 6348 servers, SSIRM demonstrates excellent engineering feasibility. Despite the introduction of multimodal features, the number of parameters (89M) is significantly lower than that of BERT-Review (110M) and SentSem-Net (142M), thanks to the design of a lightweight LEDP parser-replacing the traditional dependency rule base by cueing learning with a 37.3% reduction in parameter volume. Inference latency is only 68 ms (210 ms for SentSem-Net), meeting the real-time response requirements of education platforms (< 200 ms standard). Memory consumption is controlled at 2.8 GB, which can be deployed on medium-sized cloud servers. Future work will explore federated learning architecture to support collaborative management of resources across institutions.

5 Discussion

This study is the first to validate the practical feasibility of the cognitive-affective synergy theory in educational resource management through the mechanism of affective-semantic dual-channel integration. Traditional cognitive models of education (e.g., Bloom's Taxonomy) focus on the evolution of knowledge dimensions and neglect the moderating role of affective factors on learning effectiveness (Farani, 2022). The dynamic weighting mechanism of the SSIRM framework reveals that affective characteristics contribute 82% to resource selection in the beginner's stage (*LearnStage* < 0.4), High emotional acceptance (e.g., 'teacher approachability') at the beginning of learning significantly increased persistence (r = 0.79, p < 0.001); whereas in the higher-order learning phase (*LearnStage* > 0.4), the weight of semantic features jumps to 72%, confirming Sweller's cognitive load theory that expert learners need to refine the matching of knowledge structures to reduce the intrinsic cognitive load (De Jong, 2010). This dynamic shift in feature dominance provides strong empirical support for cognitive-affective synergy theory, demonstrating how affective factors are crucial for initiating learning (complementing Bloom's foundation) and how cognitive precision becomes key for advanced efficiency (complementing Sweller's load management). This finding provides empirical support for the construction of a 'stage-adaptive' cognitive model of education, and bridges the longstanding gap between educational psychology and artificial intelligence research..

More profoundly, the cross-modal attention mechanism in the SSIRM framework realises the causal localisation of pedagogical deficiencies. While traditional sentiment analysis can only determine 'user dissatisfaction', the semantic dependency parsing in this study accurately associates negative sentiments to the defects of teaching entities (e.g., 'boring to insufficient cases'). This mechanism provides a new paradigm for explainable AI in education: semantic attribution driven by sentiment signals to align algorithmic decision-making with pedagogical logic (e.g., 'add more cases' instead of simply improving ratings), echoing (Chaudhry et al., 2022b) call for transparency in educational AI. Specifically, SSIRM advances XAI in education by:

1 revealing why resources are recommended (e.g., high sentiment acceptance for visual learners)

2 translating vague feedback into actionable defects (e.g.,'boring'→ 'insufficient case studies') – addressing the 'black box' critique of traditional recommender systems.

At the practical level, the SSIRM framework promotes a paradigm shift from 'static warehousing' to 'dynamic evolution' of educational resource management, which 'static storage' refers to the traditional management model that lacks continuous updating of resources after they are uploaded, which 'dynamic evolution' can adjust the resource recommendation strategy in real time according to the user feedback, and discover and improve the problems in teaching resources in time. Current education platforms generally face the problem of lagging resource iteration-data shows that over 60% of courses have an update cycle of more than 18 months (Ni and Xie, 2024). The optimised instruction generation mechanism proposed in this study automatically outputs executable instructions (e.g., 'Add step-by-step details ≥ 5 places/chapter') by clustering high-frequency semantic entities (e.g., 'Insufficient exercise explanations') in negative comments. In the case of linear algebra course, this mechanism shortens the resource update cycle to 3.2 weeks, improves the rating by 28.7% (Δ = 0.9), and increases the completion rate by 22.3%. This validates the feasibility of 'review-driven resource evolution' and provides a technological backbone for building a self-optimising ecosystem of educational resources.

For educational administrators, firstly, it is recommended to establish an affective-semantic linked resource monitoring dashboard and integrate the four-dimensional radar chart output from SSIRM into the management backend in order to identify high-risk resources in real time, such as resources with an affective acceptance level of less than 0.4 and a cognitive load higher than 0.7. Secondly, it is suggested to implement a learning stage-based triage recommendation strategy that utilises a meta-learning weight adapter to dynamically adjust the recommendation logic, e.g., prioritising resources with higher affective acceptance, such as animation-explained calculus, to novice learners. Finally, it is suggested to build a knowledge base of instructional deficiencies and accumulate optimised instructions generated by semantic clustering to form a cross-disciplinary instructional design guide. These practices not only improve platform operational efficiency and reduce resource iteration costs by 41%, but also promote educational equity. For example, a pilot in North Carolina showed that SSIRM-driven resource optimisation reduced the gap in student completion rates in rural schools by 18% (Cohen's d = 0.43).

Although the SSIRM framework has achieved remarkable results, it still has some limitations. First, the cross-language adaptability is insufficient. The current model is trained on the English dataset, and when migrated to Chinese educational scenarios, the F1 value of sentiment recognition drops by 12.4%, such as the misjudgement of Chinese educational metaphors such as ‘dried fruit’ and causes incorrect quality assessments due to culture-specific expression errors. Second, the coverage of niche subjects is limited, and the semantic parsing F1 value of vocational education resources (e.g., ‘computer numerical control machine operation’) is only 0.712, below the 0.80 threshold required for reliable educational AI systems, which is due to the insufficient coverage of entities in the CurriculumNet map. Finally, the requirement of accumulating at least 50 comments to generate optimisation instructions introduces significant delays in real-time feedback. This threshold impedes immediate responses to emerging issues, particularly for cold-start resources (e.g., newly uploaded courses) or niche subjects with low user engagement. Consequently, critical defects (e.g., conceptual errors or technical flaws) may persist for extended periods, leading to learner frustration and increased dropout rates. For instance, EdNet data indicates that 35% of new resources receive fewer than 20 comments within their first three months, leaving them excluded from the optimisation cycle. Future work should explore lightweight alternatives (e.g., few-shot learning) to mitigate this constraint. To address these challenges, future research will focus on multilingual educational embedding spaces, joint training of multilingual pre-trained models (e.g., XLM-R), aligning cross-cultural educational concepts such as ‘heuristics’ through comparative learning, and federated knowledge graph construction, allowing educational institutions to collaboratively extend domain entities under data privacy protection, such as Medical Education Thesaurus, with differential privacy techniques to control the risk of information leakage; and generative teaching agents, integrating large language models (e.g., Llama 3) to generate virtual reviews that simulate user feedback on cold-start resources.

These research directions will advance the management of educational resources towards a next generation paradigm that is ‘adaptive and evolvable’, in line with the vision of ‘Inclusive and Intelligent Education’ advocated in the United Nations Educational, Scientific and Cultural Organisation Declaration on Education 2030 (Oudhia, 2024). Through these efforts, the management of educational resources will become smarter and more personalised, providing more equitable and effective learning experiences for learners from different backgrounds.

6 Conclusions

The smart teaching resource-management framework (SSIRM) proposed in this study, which integrates the emotional and semantic features of user comments, significantly improves the recommendation accuracy and quality assessment efficacy of teaching resources through the quantification of emotional intensity, the parsing of educational entity dependencies, and the adaptive weighting mechanism. Experimental results show that the SSIRM framework achieves 12.7% recommendation accuracy improvement and 9.2% F1 value improvement on the EdNet dataset, effectively solving the personalised recommendation problem in the context of overloaded educational resources. The framework not only optimises the resource management process, but also provides technical support for the dynamic optimisation of educational resources, and promotes the transformation of educational resource management from ‘static storage’ to ‘dynamic evolution’. Despite the limitations of insufficient cross-language adaptability, limited coverage of niche subjects, and delayed real-time feedback, its innovation and practicality lay a solid foundation for the intelligent and personalised development of educational resource management in the future.

Declarations

All authors declare that they have no conflicts of interest.

Informed consent statement

Written informed consent was obtained from all participants, and all data were anonymised to ensure confidentiality. The authors confirm that no personally identifiable information is disclosed in this manuscript.

References

Abu-Salih, B. and Alotaibi, S. (2024) ‘A systematic literature review of knowledge graph construction and application in education’, *Heliyon*, Vol. 10, No. 3, p.e25383.

Aguilar, J., Salazar, C., Velasco, H., Monsalve-Pulido, J. and Montoya, E. (2020) ‘Comparison and evaluation of different methods for the feature extraction from educational contents’, *Computation*, Vol. 8, No. 2, p.30.

Akdemir, E. and Barışçı, N. (2024) ‘A review on deep learning applications with semantics’, *Expert Systems with Applications*, Vol. 251, p.124029.

Chaudhry, M.A., Cukurova, M. and Luckin, R. (2022a) ‘A transparency index framework for AI in education’, *International Conference on Artificial Intelligence in Education*, Springer, pp.195–198.

Chaudhry, M.A., Cukurova, M. and Luckin, R. (2022b) ‘A transparency index framework for AI in education’, *International Conference on Artificial Intelligence in Education*, Vol. 13356, pp.195–198.

Chen, C-M. and Wang, H-P. (2011) ‘Using emotion recognition technology to assess the effects of different multimedia materials on learning emotion and performance’, *Library & Information Science Research*, Vol. 33, No. 3, pp.244–255.

Choi, Y., Lee, Y., Shin, D., Cho, J., Park, S., Lee, S., Baek, J., Bae, C., Kim, B. and Heo, J. (2020) 'Ednet: a large-scale hierarchical dataset in education', *Artificial Intelligence in Education: 21st International Conference, AIED 2020, Proceedings*, Ifrane, Morocco, 6–10 July, Part 2, Vol. 21, No. 12164, pp.69–73.

De Jong, T. (2010) 'Cognitive load theory, educational research, and instructional design: some food for thought', *Instructional Science*, Vol. 38, No. 2, pp.105–134.

Durán, C.G. and Ramírez, C.M. (2021) 'Integration of open educational resources using semantic platform', *IEEE Access*, Vol. 9, pp.93079–93088.

Farani, R. (2022) 'Revisit Bloom's taxonomy: a reflection of education in a digital era', *Journal of English Language and Pedagogy*, Vol. 5, No. 1, pp.59–64.

Gulyamov, S.S., Fayziev, R.A., Rodionov, A.A. and Jakupov, G.A. (2023) 'Leveraging semantic analysis in machine learning for addressing unstructured challenges in education', *2023 3rd International Conference on Technology Enhanced Learning in Higher Education (TELE)*, Vol. 51, pp.5–7.

Kovalan, A. (2024) 'Evaluation of online MOOC courses and support for learners' coures selection', *CAHIERS MAGELLANES-NS*, Vol. 6, No. 2, pp.4043–4055.

Moreno-Marcos, P.M., Alario-Hoyos, C., Muñoz-Merino, P.J., Estévez-Ayres, I. and Kloos, C.D. (2018) 'Sentiment analysis in MOOCs: a case study', *2018 IEEE Global Engineering Education Conference (EDUCON)*, Vol. 1, pp.1489–1496.

Musto, C., Gemmis, M.d., Lops, P., Narducci, F. and Semeraro, G. (2012) 'Semantics and content-based recommendations', *Recommender Systems Handbook*, Vol. 1, pp.251–298.

Ni, B. and Xie, X. (2024) 'Distributed distribution and scheduling of teaching resources based on a random matrix educational leadership model', *Informatica*, Vol. 48, No. 8, p.8.

Oudhia, A. (2024) 'The Indian university system in the knowledge age: a metamorphosis', *ASHA PARAS International Multidisciplinary Research Journal*, Vol. 1, p.11.

Piedra, N., Chicaiza, J., López, J., Martínez, O. and Caro, E.T. (2010) 'An approach for description of open educational resources based on semantic technologies', *IEEE EDUCON 2010 Conference*, Vol. 1, pp.1111–1119.

Rao, Y. and Yang, F. (2022) 'A method for classifying information in education policy texts based on an improved attention mechanism model', *Wireless Communications and Mobile Computing*, Vol. 2022, No. 1, p.5467572.

Song, J., Chen, H., Li, C. and Xie, K. (2023) 'MIFM: multimodal information fusion model for educational exercises', *Electronics*, Vol. 12, No. 18, p.3909.

Wang, M. and Yang, Q. (2024) 'Educational big data analytics using sentiment analysis for student requirement analysis on courses', *Scalable Computing: Practice and Experience*, Vol. 25, No. 5, pp.3858–3866.

Wang, Y., Huang, M., Zhu, X. and Zhao, L. (2016) 'Attention-based LSTM for aspect-level sentiment classification', *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, Vol. 1, pp.606–615.

Wu, B. and Chen, W. (2023) 'Factors affecting MOOC teacher effectiveness from the perspective of professional capital', *Behaviour & Information Technology*, Vol. 42, No. 5, pp.498–513.

Zhang, C., Wu, R., Sun, F., Lin, Y., Liang, Y., Teng, J., Liu, N., Ouyang, Q., Qian, L. and Yan, H. (2024) 'Parallel molecular data storage by printing epigenetic bits on DNA', *Nature*, Vol. 634, No. 8035, pp.824–832.

Appendix

SSIRM framework algorithm

The core workflow is summarised in Algorithm 1:

Algorithm 1 SSIRM workflow

Input: User comments C, behaviour logs B

Output: Recommendations R, optimisation instructions O

Preprocess C (segmentation, deactivation)

F_emo ← Edu-LSTM(C) ▷ Emotional feature extraction [equation (4)–(5)]

F_sem ← LEDP-RoBERTa(C) ▷ Semantic entity extraction (Section 3.3)

LearnStage ← σ(T_master / T_total) ▷ Stage index [equation (7)]

α ← exp(–2.5 * LearnStage) ▷ Dynamic weight [equation (8)]

F_fused ← CrossModalAttention(F_emo, F_sem, α) ▷ equation (9)

R ← RankResources(F_fused, B, LearnGap) ▷ equation (10)

O ← GenerateInstructions(ClusterNegativeEntities(F_sem))

return R, O