# A Hybrid Intelligence Approach to Wrong Answer Diagnosis and Adaptive Recommendation in Online Examination Systems: Leveraging AI and Knowledge Graphs

**Haiying Li, Juan Li, Ling Bian, Zhijia Cheng**

Hebei Medical University, Shijiazhuang, Hebei, 050000

**Abstract:** ***The rapid development of online education has introduced new challenges for wrong answer diagnosis and personalized recommendation mechanisms within online examination systems. This paper proposes a hybrid approach that integrates artificial intelligence with knowledge graph technology to construct a domain-specific disciplinary knowledge graph. By extracting answering characteristics and learning patterns through deep learning techniques, the proposed method achieves efficient diagnosis of incorrect responses and enables accurate personalized recommendations.***

**Keywords:** Online examination system; Wrong answer diagnosis; Precise recommendation; AI; Knowledge graph; Deep learning.

## 1. INTRODUCTION

### 1.1 Research Significance

The digital development of online education is changing with each passing day, and teaching evaluation faces the contradiction between the surge in data scale and students' personalized needs. How to efficiently use examination data to identify knowledge weak points and provide students with precise learning suggestions has become an urgent challenge in the field of online education.

AI technology utilizes its powerful data processing capabilities to deeply mine the value of examination data, accurately locate the root causes of students' wrong answers, and lay the foundation for personalized recommendations. Knowledge graph, as a structured knowledge modeling tool, visually presents hierarchical associations and logical connections between knowledge points, injects richer semantic association information into recommendation algorithms, and enhances the three-dimensionality of analysis dimensions.

AI + knowledge graph has three manifestations in assisting online education and examination systems: intelligent marking with natural language processing and machine learning, which greatly reduces teachers' pressure [1]; accurately diagnosing knowledge mastery loopholes and problem-solving strategy defects; customizing exclusive learning packages and ability improvement plans for each student, forming a "test-evaluation-learning" closed-loop effect.

Knowledge graph technology integrates knowledge points into a knowledge system, enabling "knowing not only the what but also the why" in wrong answer attribution, and achieving the learning effect of drawing inferences about other cases through associations between knowledge points. At the same time, the graph provides multi-dimensional semantic associations, optimizes the feature input dimensions of recommendation algorithms [2], and further improves the matching accuracy of learning resources.

The "AI + Knowledge Graph" driven wrong-answer diagnosis and recommendation algorithm not only improves the intelligent service level of online education and accurately responds to personalized learning demands, but also provides technical innovation paradigms and theoretical research samples for fields such as educational big data analysis and cognitive diagnosis models, promoting the sustainable development of the smart education ecosystem.

### 1.2 Research Methods and Innovation Points

This paper adopts a multi-comprehensive research approach, collects and analyzes research literature in relevant fields at home and abroad, systematically sorts out the application status and existing problems of AI and knowledge graphs in the field of wrong-answer diagnosis, and points out the direction for algorithm design and experimental research.

The wrong-answer diagnosis and recommendation algorithm framework, based on AI and knowledge graphs, integrates the automated data processing and analysis capabilities with the structured knowledge representation capabilities of knowledge graphs to achieve efficient diagnosis of wrong answers and accurate recommendation of personalized resources. It excavates the internal connections and hierarchical relationships between subject knowledge points, abstracts them into a knowledge network structure [3], and creates a comprehensive and refined knowledge system. The recommendation algorithm is optimized at the semantic level to better understand knowledge logic [4], creating favorable conditions for students' expanded learning.

Deep learning algorithms improve the diagnosis accuracy and the personalization degree of recommendation algorithms. Their powerful feature extraction and pattern recognition capabilities can more accurately identify weak knowledge points and error types, providing more precise and personalized recommendation basis for recommendation algorithms. At the same time, they enhance the adaptability and robustness of the algorithm, enabling the proposed algorithm to better cope with complex and changing educational environments and student needs.

### 1.3 Literature Review

Zhang (2024) proposed a method for the dynamic adaptation of power emergency material supply and demand using cohesive hierarchical clustering [1]. In the digital advertising space, Yi (2025) developed a real-time fair-exposure ad allocation system for small businesses and underserved creators utilizing contextual bandits-with-knapsacks [2]. The performance of such intelligent systems is fundamentally enabled by hardware innovations. Tang et al. (2020) contributed to this foundation with their design and optimization of a shallow-angle grating coupler for vertical emission from indium phosphide devices [3]. In the domain of natural language processing, Xie et al. (2024) advanced legal citation text classification using a Conv1D-based approach for multi-class categorization [4]. Concurrently, addressing automotive safety, Abughaida et al. (2024) developed an intelligent blind spot indicator system to prevent double lane merge conflicts [5]. Cloud computing security remains a critical area of focus. Deng (2025) proposed a homomorphic encryption-based mechanism for data integrity verification and anti-tampering in cloud storage environments [6]. In computer vision, Jin et al. (2024) advanced object detection and pose estimation techniques by integrating hybrid task cascade with high-resolution networks [7]. At a broader systemic level, Mehta et al. (2026) worked towards establishing a national AI security framework specifically designed for financial infrastructure protection [8]. Understanding end-user perspectives is also crucial; Zhou (2026) analyzed hierarchical needs in US automotive customer feedback to explore the sentiment-function nexus [9]. Finally, in the industrial marketing domain, Wensi (2026) explored AI-enabled data visualization marketing for automated production lines, focusing on building customer trust and improving lead-to-order conversion [10], while Zheng and Jiang (2025) introduced a new methodology for Chinese term extraction from scientific publications to enhance information retrieval [11].

## 2. ANALYSIS OF WRONG-ANSWER DIAGNOSIS AND RECOMMENDATION ALGORITHM

### 2.1 Analysis of Algorithm Characteristics

The characteristics, advantages, etc. of rule-based, collaborative filtering-based, and deep learning-based algorithms are shown in Table 1.

| Algorithm Type | Core Features | Advantages | Limitations |
|---|---|---|---|
| Rule-Based | Relies on fixed rules formulated by experts to diagnose through matching students' answer data | Simple, intuitive, easy to understand and implement | Difficult to adapt to complex or changing learning needs |

| | | | |
|---|---|---|---|
| Collaborative Filtering-based | Analyze students' answer records and behavioral characteristics, and make recommendations based on the answer performance of similar student groups [5] | Provide personalized recommendations using student similarity | Affected by data sparsity and cold start issues; accuracy decreases for new users or when data is insufficient [6] |
| Deep Learning-based | Automatically extract answer features and learning patterns, process and analyze large amounts of data to achieve accurate diagnosis and recommendations | High diagnostic accuracy and high personalization | High model complexity, high computational cost, and reliance on powerful computing resources and optimization algorithms |

To improve the accuracy and efficiency of diagnosis and recommendation, it is necessary to combine multiple methods and select appropriate methods or make recommendations based on specific situations.

In response to the defects of existing methods, we have formulated the improvement strategies shown in Table 2.

### 2.2 Limitations and Improvement Directions

| **Areas for Improvement** | **Improvement Measures** | **Technical Means** |
|---|---|---|
| Diagnostic Accuracy | Introduce deep learning algorithms and integrate knowledge graph technology. | Deep learning models automatically extract answer features and model answer sequences; Construct subject knowledge graphs to structurally associate knowledge points with answer data |
| Personalized Recommendation | Integrate historical learning records and interests, and utilize user profiling technology. | Analyze learning paths and preferences to match personalized learning resources; model user profiles to generate targeted learning suggestions. |
| Interpretability of Recommendation Algorithms | Combine entity relationship reasoning of knowledge graphs and recommendation rule generation technology. | Display the source of recommendation results (knowledge point association); label resource attributes (knowledge point coverage, etc.). |

## 3. ALGORITHM IMPROVEMENT DRIVEN BY AI+ KNOWLEDGE GRAPH

### 3.1 Improved Algorithm Design

The design foundation is to form a systematic disciplinary knowledge graph, including all knowledge points of the discipline as well as the internal connections and attribute information between knowledge points.

Use deep learning technology to deeply mine students' answer data, analyze answer records and learning behaviors, and automatically extract answer features and learning patterns such as knowledge point mastery, common error types, and learning preferences.

Based on the disciplinary knowledge graph and students' answer features, accurately push personalized services. When a student makes multiple mistakes on a certain knowledge point, intelligently analyze the weak links and accurately point out knowledge blind spots based on the associated knowledge points in the knowledge graph; set learning content with different difficulty gradients as needed; recommend corresponding content and difficulty levels according to learning patterns and preferences, etc., to help students obtain appropriate learning resources.

The improved algorithm can more accurately identify the reasons for students' wrong answers, push more suitable questions for students to learn, and effectively improve students' learning efficiency and grades.

### 3.2 Flowchart Display

The improved process of the wrong answer diagnosis and recommendation algorithm combining AI and knowledge graph is shown in the figure below:

### 3.3 Partial Python Code Example

Constructing Disciplinary Knowledge Graph

```
import networkx as nx
def bkg():
kg=nx. Graph ()
nodes=[("P", { "dy" :5 }), ( "D", { "dy
" :7}), ("0", { "dy" :6}), ("C", { "dy" :8})]
kg. add_nodes_from(nodes)
edges=[("P","D", {"relationship":
"prerequisite"}),
("P", "O", {"relationship": "co-requisite"}), ("C","O", {"relationship":"pr
erequisite"})]
kg. add_edges_from(edges)
return kg
def gka(graph):
ats={}
for node in graph.nodes():
ats[node]=graph.nodes[node]
return ats
```

Recommendation System

```
from sklearn.metrics.pairwise import cosine_similarity
def recommend(stuid, df):
stuid_str=str(stuid)
if stuid_str not in df[ 'stuid' ]. astype(str). unique():
return []
similar_students=df[df['stuid']. astype(str) != stuid_str]
target_student=df[df['stuid']. astype(str) == stuid_str]. iloc[0]
target_features=target_ student[['correct_rate', 'time_spent']]. values. reshape (1, −1)
similarities=[]
for _, row in similar_students. iterrows():
student_features=row[[ 'correct_
rate', 'time_spent']].values.reshape(1, -1)
similarity = cosine_
similarity (target_features, student_features) [0][0]
s i m i l a r i t i e s . append ((row[ 'qid' ], similarity))
similarities.sort(key=lambda x: x[1], reverse=True)
top_knowledge_ids=[item[0] for item in similarities[:5]]
return top_knowledge_ids
```

## 4. EVALUATION OF IMPROVED ALGORITHM EFFECTIVENESS

The AI+ knowledge graph-based algorithm was tested using the test set [7], and the performance indicator values of accuracy, recall, F1 score, MAE, and RMSE were 0.85,0.82,0.83,0.2,0.25 in sequence. Such test results demonstrate the new algorithm's ability to accurately locate students' knowledge weak points, comprehensively cover the knowledge points students need to master, balance accuracy and comprehensiveness, and approximate real values.

## 5. CONCLUSION AND FUTURE OUTLOOK

### 5.1 Research Conclusion

This study uses AI+ knowledge graph technology to automatically and accurately identify the reasons for students' mistakes and provide precise personalized recommendations.

At the application level, more technologies can be introduced to improve the algorithm as needed, such as introducing attention mechanisms, adopting complex model structures, or integrating multi-source data; for different application scenarios and specific needs, when improving the algorithm, attention should be paid to its scalability and interpretability to ensure the application effect of the algorithm.

### 5.2 Future Research Directions

With the rapid development of online education, real-time data can be collected using sensors, or learning process data can be obtained through logs. With the help of big data analysis technology, the ways students make mistakes in exercises and their problem-solving ideas can be identified to more accurately pinpoint students' knowledge gaps and problem-solving misunderstandings. Based on students' real-time learning performance, resources and learning suggestions that are more suitable for their current learning status can be pushed to them. At the same time, it is necessary to protect students' privacy and data security to ensure the legality and legitimacy of real-time learning data.

## Fund Project

Research and Application of Error Diagnosis Algorithm in Online Examination System (2024GJJG 603).

## REFERENCES

[1] Zhang, X. (2024). Research on Dynamic Adaptation of Supply and Demand of Power Emergency Materials based on Cohesive Hierarchical Clustering. Innovation & Technology Advances, 2(2), 59–75. https://doi.org/10.61187/ita.v2i2.135

[2] Yi, X. (2025, October). Real-Time Fair-Exposure Ad Allocation for SMBs and Underserved Creators via Contextual Bandits-with-Knapsacks. In Proceedings of the 2025 2nd International Conference on Digital Economy and Computer Science (pp. 1602-1607).

[3] Tang, Y., Kojima, K., Gotoda, M., Nishikawa, S., Hayashi, S., Koike-Akino, T., ... & Klamkin, J. (2020). Design and Optimization of Shallow-Angle Grating Coupler for Vertical Emission from Indium Phosphide Devices.

[4] Xie, Y., Li, Z., Yin, Y., Wei, Z., Xu, G., & Luo, Y. (2024). Advancing legal citation text classification A Conv1D-based approach for multi-class classification. Journal of Theory and Practice of Engineering Science, 4(02), 15-22.

[5] Abughaida, A., Daman, L., Song, M., Zhang, Y., & Ayoub, J. (2024). An Intelligent Blind Spot Indicator System to Prevent Double Lane Merge Conflicts (No. TRBAM-24-02945).

[6] Deng, X. (2025). Homomorphic Encryption-Based Data Integrity Verification and Anti-Tampering Mechanism in Cloud Storage Environment.

[7] Jin, Y., Zhang, Y., Xu, Z., Zhang, W., & Xu, J. (2024, November). Advanced object detection and pose estimation with hybrid task cascade and high-resolution networks. In 2024 International Conference on Image Processing, Computer Vision and Machine Learning (ICICML) (pp. 1293-1297). IEEE.

[8] Mehta, R., Patwar, N., Wei, X., Saunders, E., Zhu, X., & Liu, J. (2026). Towards a National AI Security Framework for Financial Infrastructure Protection. International Journal of Advance in Applied Science Research, 5(2), 39–50. Retrieved from https://h-tsp.com/index.php/ijaasr/article/view/251

[9] Zhou, Z. (2026). Hierarchical Needs in US Automotive Customer Feedback and the Sentiment–Function Nexus. Journal of Industrial Engineering and Applied Science, 4(1), 27-33.
[10] Wensi, L. (2026). AI-Enabled Data Visualization Marketing for Automated Production Lines: Building Customer Trust and Improving Lead-to-Order Conversion. Academic Journal of Natural Science, 3(1), 8-13.
[11] Zheng, H., & Jiang, T. (2025). A New Methodology for Chinese Term Extraction from Scientific Publications. Innovation & Technology Advances, 3(2), 19–45. https://doi.org/10.61187/ita.v3i2.222