Trends in Social Sciences and Humanities Research
ISSN: 2959-9938
DOI: https://doi.org/10.61784/tsshr3093

APPROACH TO LEARNING BEHAVIOUR ANALYSIS FOR HIGHER-ORDER THINKING DEVELOPMENT: EDUCATIONAL BIG DATA PERSPECTIVE

Ji Li[1], Jin Lu[2*]
[1]*Research Management Office, Shenzhen Polytechnic University, Shenzhen 518055, Guangdong, China.*
[2]*Guangdong Key Laboratory of Big Data Intelligence for Vocational Education, Shenzhen Polytechnic University, Shenzhen 518055, Guangdong, China.*
Corresponding Author: Jin Lu, Email:lujin0808@szpu.edu.cn

Abstract: With the rapid development of information technology, educational big data gradually become an important resource in the field of education. Characteristic by large data volume, diverse types and low value density, educational big data can provide rich information support for education and teaching. Higher-order thinking, as a necessary quality of innovative talents, has been increasingly valued in the field of education. Educational big data is closely linked to the development of higher-order thinking, and through the analysis of educational big data, it can provide an in-depth understanding of students' learning behaviors and provide a basis for the cultivation of students' higher-order thinking. The purpose of this paper is to explore the mechanism of promoting the development of students' higher-order thinking through the analysis of students' learning behaviors in the perspective of education big data. After training, the machine automatically annotated text used in the examples of this paper achieves a score of 4.5 or more, with an accuracy rate close to 0.98, which can meet the needs of classroom applications. It can be seen that when this paper is applied to actual classroom evaluation, it lays an important foundation for carrying out large-scale comparative research on classroom teaching as well as digging into the laws of classroom dialogue, and it can break down the barriers of mathematics teaching between districts and schools, break down the effect of data silos, and provide a relatively uniform basis for the comparison and evaluation of the level of classroom teaching.
Keywords: Big data in education; Higher order thinking; Learning behaviour analysis

1 INTRODUCTION

At present, research on educational big data and higher-order thinking at home and abroad has achieved certain results. In terms of educational big data, research mainly focuses on data collection, processing and analysis, aiming to explore the potential value in educational data and provide support for educational decision-making. In terms of higher-order thinking, the research mainly focuses on the connotation, cultivation method and evaluation system of higher-order thinking, aiming to improve students' innovation and problem-solving ability. However, there are relatively few studies that combine educational big data with the development of higher-order thinking, and further in-depth exploration is needed.

Along with the rise of online learning platforms, prediction studies based on students' online learning behaviors have become increasingly common, involving relevant techniques from several fields, including traditional learning behaviour analysis based on human observation and learning behaviour analysis based on machine learning algorithms. Based on traditional methods, it often starts with academic performance.Hasan et al. collected course data from 22 students and used GPA, sequence course grades, and online module quiz data to predict the final grade level[1]. Burgos et al. collected course data from more than 100 students and used the grades of quizzes, assignments, and other modules in the Module platform to predict whether students could pass the the course[2]. Amrieh et al. predicted students' performance by analyzing data from an online learning platform and obtaining the number of times students spoke in discussions in forums, the number of times they accessed resources, and the amount of time they spent reading the material, which comprised a vector of students' online learning behavioral features, and found that there was a strong correlation between online learning behaviors and students' performance[3]. Cobb et al. predicted the performance of students by observing fourth graders over a 9-day classroom observations of fourth graders, counting the proportion of different student behaviors in the classroom to predict students' reading and spelling scores[4]. McKinney et al. conducted 5-minute classroom observations of children participating in a language arts program every 4 days, predicting performance based on 120 observations of children's behaviors[5].Allen et al. used classroom observation to analyse teacher-student interaction behaviors in the classroom, finding a strong correlation between online learning behaviors and student performance[6]. The manual observation method has become a bottleneck in the research of teaching behaviour due to its low coding efficiency and high labour cost. With the rapid development of machine learning technology, researchers have begun to try to adopt deep learning methods to obtain students' classroom behaviors. Using classroom videos, Bidwell et al. measured student engagement by using a Hidden Markov Model to analyse students' gaze targets in the classroom[7]. Thomas et al. calculated students' attention levels by using OpenFace open-source software to analyse students' facial expressions, head postures, and eye gaze targets in the classroom[8]. Ashwin et al. proposed a hybrid convolutional neural network model to analyse students' facial

© By the Author(s) 2024, under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0).

expressions, gestures and body postures to reflect the learning status of students in the classroom, and found that student engagement was positively correlated with paper grades[9]. Gong et al. devised an end-to-end graphical constitutional neural network method, ACKRec, for knowledge concept recommendation in MOOC platforms, which combines students, knowledge points, and other types of of entities, such as courses, videos, and instructors, into a heterogeneous information network that guides the propagation of student preferences based on meta-paths[10]. Subsequent research has found that knowledge graphs can effectively fuse massive heterogeneous data and expert knowledge, and incorporating knowledge graphs into learner personalized recommendation models can help characterize the underlying relationships between entities and increase the interoperability of learning resource recommendations using pathway reverse inference[11]. Deep learning methods can be further classified into two categories according to the spatial-temporal relationship of the model, one category is based on the answer time sequence based on recurrent neural network RNN, which is too sketchy to model the structural information of skills[12]. The other category is based on the structure of skills based on graph neural network GNN[13], the knowledge tracking model based on graph neural network can better express the knowledge, the structure between questions and extend the model by using the edge information, so the knowledge tracking based on graphs has a broader research prospect[14-16].

In online education environments that emphasize self-directed learning, learners' self-regulated learning behaviors and self-determined learning motives have a direct impact on learning outcomes. However, this new type of teaching environment still lacks accurate teaching evaluation and learning behaviour prediction methods and tools, which greatly limits its advantages and restricts its popularity in the education field. Existing research either follows a theory-driven psychometric paradigm or favors data-driven mining of learning behaviour patterns. If these two paths are integrated, it can promote theoretical research related to online learning, examine feasible ways to improve students' online learning effectiveness, and combine them with technological means for personalized intervention. Secondly, traditional online teaching analysis research is often based on annotated data for static descriptive statistics, the output of various types of language or behaviour whether and how often they appear, the static analysis method can not respond to the teacher and student how to build a chain of thinking progression through dialogue, what is the evolutionary pattern of higher-order thinking development and other issues, it is also difficult to provide targeted process guidance for teachers to improve teaching methods. At the same time, the existing research work to build online learning learning community appeared unfair, can not guarantee the knowledge complementary between learners and other problems.

In view of the deficiencies of the existing online teaching behaviour analysis research, this paper is based on the self-developed online course teaching platform, through data collection, data extraction and analysis and mining, constructing educational big data knowledge graph, focusing on the design of model algorithms based on accurate teaching assessment and learning behaviour analysis and prediction and other applied research practices. The relevant research results are expected to provide referable templates for digital online teaching diagnostic platforms, open up new ideas and approaches for learning analytic technology to better serve online learning as well as thinking development, help improve the quality of online teaching interactions, enhance students participation and reflection ability, promote the development of students higher-order thinking, and provide targeted strategies for improving teaching and learning.

2 RESEARCH PROGRAMME DESIGN

Compared with the traditional classroom education environment, online teaching, which has an online education platform with large-scale learners, faces problems such as high dropout rate and low course completion rate, and learning community can be a good solution to the above problems[17]. This paper carries out research on learning community from the perspective of group learning, and will provide a comprehensive overview of the research on learning community in online education environment, and finally construct a learning community for online teaching. The specific construction process is shown in Figure 1.

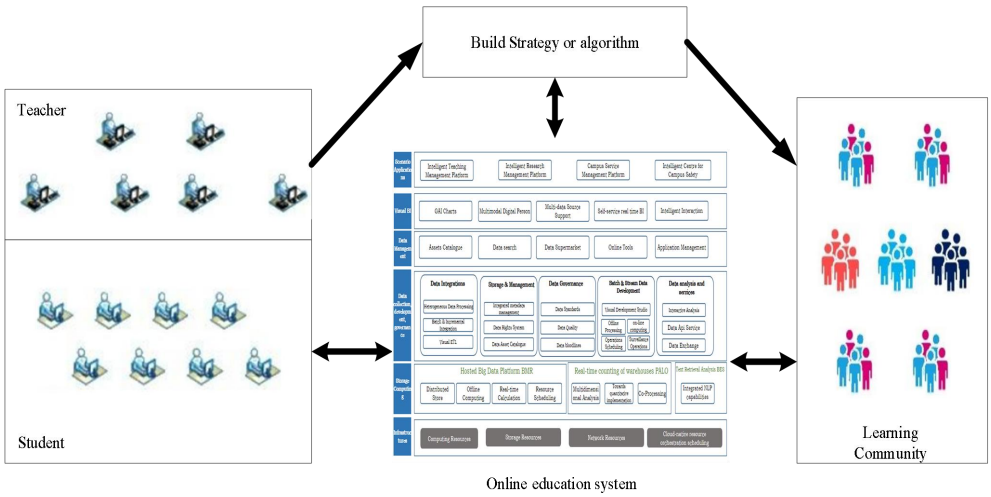

Figure 1 Schematic Diagram of the Process of Building a Learning Community in an Online Education Environment

In this paper, we propose a hybrid neural network-based automatic annotation method for online teaching, combining the advantages of CNN in being able to extract local information such as words and phrases in the discourse and Bi-LSTM in being able to understand the global feature information in the sequence data in combination with the contextual content, to learn the annotated data and perform model training. The specific architecture is shown in Figure 2. Explore the use of intelligent recording and broadcasting equipment to collect data on the online teaching process, and transcribe the online teaching discourse information captured by audio and video into text form to establish an online teaching corpus. A linguistic representation model (BERT) is constructed to preprocess the text by representing the online teaching discourse text as continuous vectors and applying it to the downstream computational task of neural networks to prepare for automatic annotation. The linguistic representation model within the embedding layer is constructed to convert the text information into continuous vectors, and a constitutional neural network layer is used to extract local features from the continuous vectors, while a bi-directional long and short-term memory neural network layer is applied to extract global features from the continuous vectors, and encode and annotate the text information based on the local features and global features.

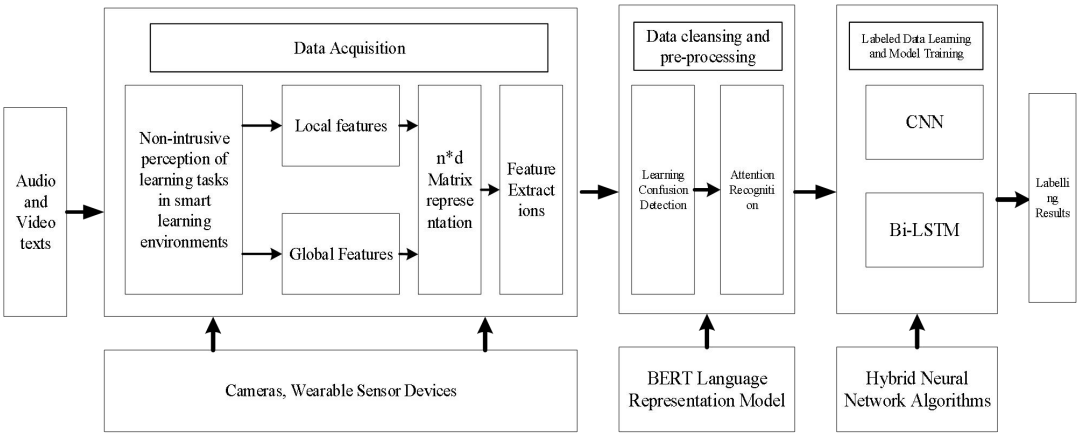

Figure 2 Automatic Annotation Method for Online Teaching Based on Video Text

The core of online learning facilitation mechanism is to design a friendly recommendation model for educational resources[18-19]. This project proposes to propose a group recommendation model based on a hierarchical attention mechanism, where planning consists of a two-part attention network, where the first part learns the preference weights of each member in the group decision-making process; and the second part learns the interactions between the members in the group in order to dynamically adjust the influence weights of the members, so as to merge the preferences of the different members in the group decision-making process in order to achieve the recommendation of appropriate resources to the group. The specific architecture is shown in Figure 3.

The first layer is the embedding layer. The purpose of the embedding layer is to transform high-dimensional sparse vectors into low-dimensional dense vectors to improve the computational efficiency, in which the learner and learning resource one-hot coding are converted into low-dimensional dense vectors, and the embedding vectors of the learner and learning resource are denoted as c and v, respectively.

The second layer is the aggregation layer. This layer is to aggregate the preferences of learners in the learning community to obtain the preference representations of different learning communities. The purpose of the aggregation layer is to obtain the preference representation of the learning community, which is composed of two parts: the fused representation of the learner's preferences and the subject preferences of the learning community. The preference fusion of learners in the learning community firstly utilises the hierarchical attention mechanism to obtain the corresponding embedded representation of learners based on different learning resources, and on the basis of which the embedded fusion representation of learners is obtained. The weight of the influence of different historical interactions on the learning community topic preference representation is learnt through the attention mechanism and the embedded representation of preferences is obtained.

The third layer is the pooling layer. This layer is used to capture the linear relationship between learning communities and learning resources. In it, a tracking investigation and longitudinal mediation analysis of the relationship between strategy metacognitive knowledge, learning motivation, and maths performance is conducted, focusing on the long-term effects of autonomous motivation on positive strategy metacognitive knowledge (cognitive/metacognitive strategies, competence enhancement strategies), and control motivation on negative strategy metacognitive knowledge.

The fourth layer is the full connectivity layer. This layer is used to capture nonlinear higher-order relationships between learning communities and learning resources.

The fifth layer is the prediction layer. This layer is used to compute the learning community's predicted preference scores for candidate learning resources, the final output of the model.

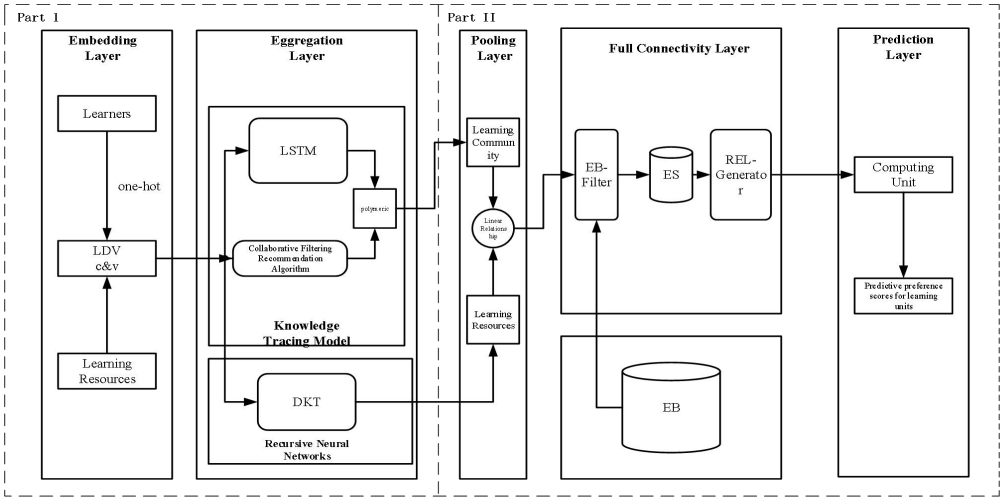

Figure 3 Group Recommendation Model Based on Hierarchical Attention Mechanism

3 EXPERIMENTAL PROCEDURE AND DATA ANALYSIS

Based on MOOCs, an online course teaching platform, and large-scale online education data, this project empowers interactive online teaching analysis and higher-order thinking development, focusing on models and methods of teaching evaluation, student behaviour analysis, knowledge tracking, and personalized recommendation of teaching resources, etc. The main research contents include: online teaching evaluation system oriented to the cultivation of thinking; analysis of online learning behaviors based on the theory of self regulation; study of online learning facilitation mechanism and learning community based on knowledge tracking and resource recommendation, etc. The main research includes: online teaching evaluation system for the cultivation of thinking; online learning behaviour analysis based on self-regulation theory; online learning facilitation mechanism and learning community research based on knowledge tracking and resource recommendation. The framework of the specific experimental steps is shown in Figure 4.

Firstly, to address the lack of online teaching evaluation system, biometric method is used to extract the indicators with high usage rate in the literature of online teaching evaluation, and construct the evaluation system of offline teaching oriented to higher-order thinking. Research on automatic annotation technology for online teaching based on video text, and automatic annotation and effective extraction of key features related to the cultivation of thinking in online teaching based on the evaluation system. Carry out research on the definition and classification system of elements of classroom teaching activities.

Secondly, to address the problem of low accuracy in the analysis of online teaching behaviors, we collect log data from the intelligent tutoring system, adopt the research paradigm of educational data mining, model sequential learning behaviors using the two-layer Hidden Markov Model (TL-HMM), and carry out the exploration of sequential pattern mining and the law of thought progression. Explore the construction of a relational network diagram of student-teacher relationship and student behaviour in online course platforms. Combine the self-regulated learning process with the comparison of captured learning behaviour patterns in order to describe effective learning behaviour patterns and try to make academic predictions from learning behaviour models.

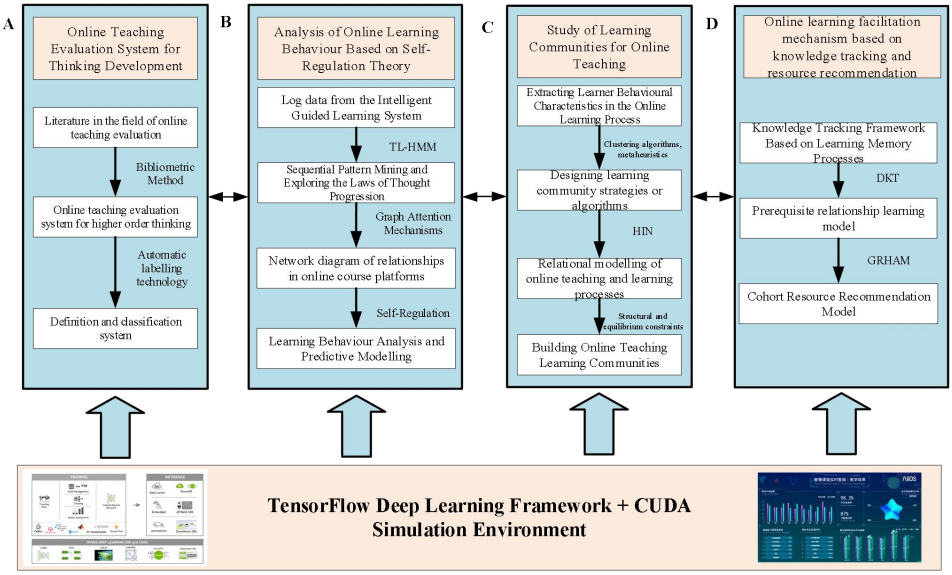

Figure 4 Experimental Framework

Thirdly, to address the problem of learners records, learning preferences and other behavioral characteristics in the online learning process, and then design learning community strategies or algorithms. Modelling is oriented towards entities such as learners, courses, videos and knowledge concepts and their relationships in the online education process, and structural constraints and balancing constraints are defined separately for the information of learners' knowledge structures and protective attributes (e.g. gender, profession, habits, etc.). Exploring the construction of learning communities.

Finally, to address the problems of 'resource overload and information disorientation' that exist in the recommendation of online teaching resources, we carry out research on knowledge tracking technology in the field of intelligent education and explore the construction of a knowledge tracking framework based on the learning memory process. Relying on the educational resources (texts, videos, courses) and the concepts they contain in the online education platform as a resource-concept map, we construct a learning model of the prerequisite relationship between the knowledge concepts of massive online educational resources. Construct a group resource recommendation model for online education learners based on the learning behaviour analysis of large-scale online learners.

Based on the Matlab R2021a platform for data testing, the database is selected as the WOS database, the specific experimental results are shown in Figure 5. Pre-operation using OpenCV and other algorithms to process video images, will be through the establishment of a learning rate optimization based on the structure of the deep learning framework for expression feature recognition, combined with the Tensenflow platform for training models, the use of Pandas and other deep learning algorithms to analyse and process the data and visual display, and ultimately, can improve the multimodal education collaborative database, for the analysis of the learner's s emotional state.

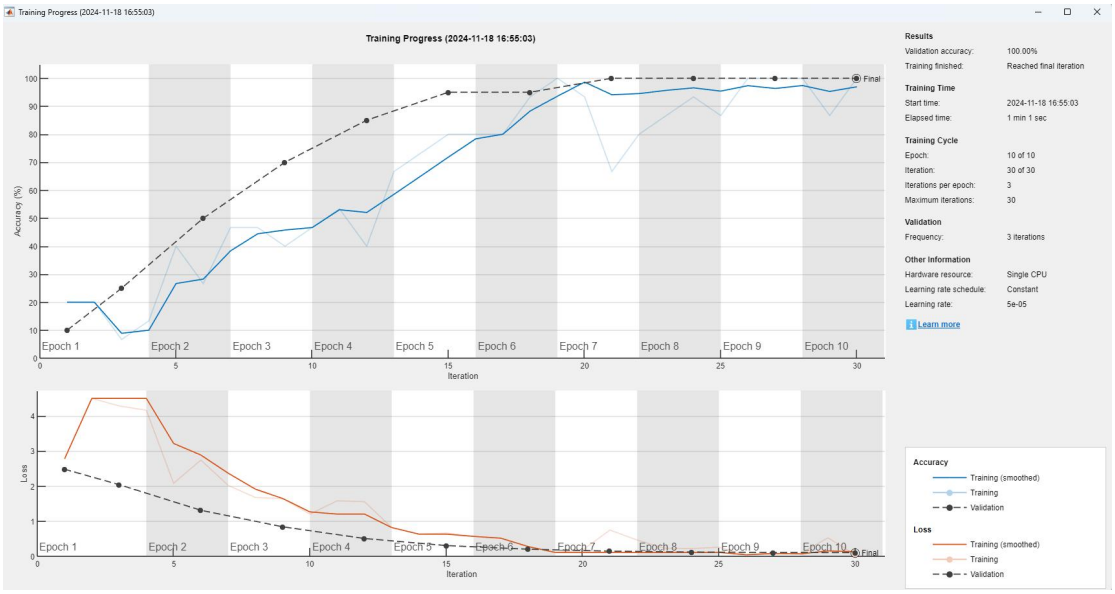

Figure 5 Experimental Results

As can be seen from the figure, the CNN+BiLSTM used has the highest parameters and the best training effect. After training, the score of the machine automatically annotating the text used in this embodiment reaches more than 4.5, and the accuracy rate is close to 0.98. This indicates that the machine annotation has high reliability and validity, and the machine automatically annotates each lesson, which greatly improves the analysis and annotation speed and accuracy, and reaches the level of large-scale application and analysis of classroom dialogues.

4 CONCLUSION

At present, teacher training and teaching and research activities are often based on the subjective experience of expert teachers or researchers, which is random and dispersed. This project will provide scientific and prospective guidance and support for teaching and research activities and training activities, which will help to promote the transformation of online teaching and learning, develop interactive, inspirational and inquiry-based teaching and learning, and contribute to the cultivation of innovative and practicing talents, and to enhance the effectiveness of mathematics teacher training, teaching and research activities. This paper develops a sequence pattern mining technique adapted to the online teaching field, which reveals the processional developmental patterns and thinking progression laws of online teaching and learning, and is expected to promote mutual communication, resource sharing, and mutual collaboration among learners and learners, learners and learners' assistants in order to complete the learning tasks. In this paper, we propose an automatic annotation method based on CNN-BiLSTM hybrid neural network model, which is well adapted to the online teaching and learning field, and is expected to be able to accurately categorize and encode online teaching and learning discourses. Secondly, it is able to quickly process multi-type and multi-level online teaching process data, which

provides a guarantee for scaled online teaching analysis. Finally, the model relies on evaluation indicators to effectively identify and refine the thinking features such as analysis and interpretation, summation, transfer and innovation, etc., which are reflected in online teaching, and is expected to be able to dig out the hidden semantic information in the dialogue text and reflect the thinking features in the online teaching process. The results of this paper can be applied to the evaluation of online teaching quality, laying an important foundation for large-scale comparative research on online teaching as well as mining online plus some dialogue laws, breaking down the barriers of mathematics teaching between different regions and schools, eliminating the effect of data silos, and providing a relatively uniform basis for the comparison and evaluation of the level of online teaching.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

This work is partially supported by 2024 Guangdong Province Education Science Planning Project (Higher Education Special Project) "Research on Smart Classroom Teaching Behavior Analysis Methods Based on Scene Semantic Understanding and Deep Learning Characteristic Representation (2024GXJK766)", 2023 Guangdong Provincial Higher Vocational Education Teaching Reform Research and Practice Project "Research on Online Teaching Quality Evaluation Method Based on Multimodal Affective State Analysis (2023JG277)", 2024 Shenzhen Polytechnic University Quality Engineering Project "Research on Classroom Scene Understanding and Behavior Analysis Method Based on Multimodal Attention Mechanisms (7024310268)", 2024 Higher Education Scientific Research Planning Project of the Chinese Society of Higher Education "Research on the Analysis of Teaching and Learning Deep Interaction Characteristics in Smart Classroom Environment Supported by Multimodal Data (24XH0407)", 2023 Shenzhen Education Science Planning Project"Research on the Evolutionary Mechanism and Intervention of Interpersonal Relationships among College Students Driven by Multimodal Data (rgzn23003)".

REFERENCE

[1] Hasan R, Palaniappan S, Raziff A R A, et al. Student academic performance prediction by using decision tree algorithm//2018 4th international conference on computer and information sciences (ICCOINS). IEEE, 2018: 1-5.
[2] Burgos C, Campanario M L, de la Peña D, et al. Data mining for modeling students' performance: A tutoring action plan to prevent academic dropout. Computers & Electrical Engineering, 2018, 66: 541-556.
[3] Amrieh E A, Hamtini T, Aljarah I. Mining educational data to predict student's academic performance using ensemble methods. International journal of database theory and application, 2016, 9(8): 119-136.
[4] Cobb J A. Relationship of discrete classroom behaviors to fourth-grade academic achievement. Journal of Educational Psychology, 1972, 63(1): 74.
[5] McKinney J D, Mason J, Perkerson K, et al. Relationship between classroom behavior and academic achievement. Journal of Educational Psychology, 1975, 67(2): 198.
[6] Allen J, Gregory A, Mikami A, et al. Observations of effective teacher–student interactions in secondary school classrooms: Predicting student achievement with the classroom assessment scoring system—secondary. School psychology review, 2013, 42(1): 76-98.
[7] Bidwell J, Fuchs H. Classroom analytics: Measuring student engagement with automated gaze tracking. Behav Res Methods, 2011, 49(113).
[8] Thomas C, Jayagopi D B. Predicting student engagement in classrooms using facial behavioral cues//Proceedings of the 1st ACM SIGCHI international workshop on multimodal interaction for education. 2017: 33-40.
[9] TS A, Guddeti R M R. Automatic detection of students' affective states in classroom environment using hybrid convolutional neural networks. Education and information technologies, 2020, 25(2): 1387-1415.
[10] Wang X, Zhao S ,Guo L , et al. Graph CA: Learning From Graph Counterfactual Augmentation for Knowledge Tracing. IEEE/CAA Journal of Automatica Sinica, 2023, 10(11): 2108-2123.
[11] Gong J, Wang S, Wang J, et al. Attentional graph convolutional networks for knowledge concept recommendation in moocs in a heterogeneous view//Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval. 2020: 79-88.
[12] Hongwei Wang, Miao Zhao, Xing Xie, et al. Knowledge Graph Convolutional Networks for Recommender Systems. In The World Wide Web Conference (WWW '19). Association for Computing Machinery, New York, NY, USA, 2019, 3307-3313.
[13] Shimizu R, Matsutani M, Goto M. An explainable recommendation framework based on an improved knowledge graph attention network with massive volumes of side information. Knowledge-Based Systems, 2022, 239: 107970.
[14] Bing Q, Zhu Q, Dou Z. Cognition-aware knowledge graph reasoning for explainable recommendation//Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining. 2023: 402-410.
[15] Corbett A T, Anderson J R. Knowledge tracing: Modeling the acquisition of procedural knowledge. User modeling and user-adapted interaction, 1994, 4: 253-278.

[16] Piech C, Bassen J, Huang J, et al. Deep knowledge tracing. Advances in neural information processing systems, 2015, 28.
[17] Nakagawa H, Iwasawa Y, Matsuo Y. Graph-based knowledge tracing: modeling student proficiency using graph neural network//IEEE/WIC/ACM International Conference on Web Intelligence. 2019: 156-163.
[18] Yang Y, Shen J, Qu Y, et al. GIKT: a graph-based interaction model for knowledge tracing//Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2020, Ghent, Belgium, September 14-18, 2020, Proceedings, Part I. Springer International Publishing, 2021: 299-315.
[19] TONG H, WANG Z, ZHOU Y, et al. HGKT: Introducing hierarchical exercise graph for knowledge tracing. 2006. https://arxiv.org/abs/2006.16915.