

Tree-Evidence-Guided Multi-Layer Learning Behavior Graph Framework for Student Performance Prediction in Online Learning

Kerangka Graf Tingkah Laku Pembelajaran Berbilang Lapisan Berpandukan Bukti Model Pokok untuk Ramalan Prestasi Pelajar dalam Pembelajaran Dalam Talian

Wang Dan¹², Rodziah Binti Latih^{1*}

¹Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Malaysia

²Faculty of Artificial Intelligence, Neijiang Normal University, Neijiang, China

*Corresponding author: Rodziah Binti Latih (e-mail: rodziah.latih@ukm.edu.my)

ABSTRACT

Online learning platforms can record students' learning process behaviors in many aspects and can assist in grade prediction, but directly splicing variables is not conducive to the interpretation of associations. This study constructs a Tree-Evidence-Guided Multi-Layer Learning Behavior Graph (TEG-MLBG) framework. First, it integrates multi-source behaviors into four layers, uses tree models as the main predictors to extract evidence, converts it into a relational graph with similarity, and uses graph residuals to supplement tree predictions to obtain the final result. In the test folds of 25, the indicators, the mean absolute error of this framework, are almost the same as the baseline, MAE of 5.122 ± 0.320 . Taking trees as the main components for prediction, graph modeling as residuals for improvement, and comprehensively considering multiple dimensions provide an interpretable method for online learning grade prediction.

Keywords: online learning analytics; student performance prediction; prediction-aware student relation graph; tree-based predictive evidence; graph residual correction

ABSTRAK

Platform pembelajaran dalam talian boleh merekodkan tingkah laku proses pembelajaran pelajar dalam pelbagai aspek dan boleh membantu dalam ramalan gred, tetapi penyambungan pembolehubah secara langsung tidak kondusif untuk tafsiran perkaitan. Kajian ini membina rangka kerja Tree-Evidence-Guided Multi-Layer Learning Behavior Graph (TEG-MLBG). Pertama, ia mengintegrasikan tingkah laku berbilang sumber ke dalam empat lapisan, menggunakan model pokok sebagai peramal utama untuk mengekstrak bukti, menukarkannya kepada graf hubungan dengan persamaan, dan menggunakan baki graf untuk menambah ramalan pokok bagi mendapatkan hasil akhir. Dalam lipatan ujian 25, penunjuk, ralat mutlak min rangka kerja ini, hampir sama dengan garis dasar, MAE iaitu 5.122 ± 0.320 . Mengambil pokok sebagai komponen utama untuk ramalan, pemodelan graf sebagai baki untuk penambahbaikan, dan mempertimbangkan pelbagai dimensi secara komprehensif menyediakan kaedah yang boleh ditafsirkan untuk ramalan gred pembelajaran dalam talian.

Kata kunci: analitik pembelajaran dalam talian; ramalan prestasi pelajar; graf perhubungan pelajar yang menyedari ramalan; bukti ramalan berasaskan pokok; pembetulan baki graf

INTRODUCTION

Online learning platforms can record students' course activities, covering many aspects such as task completion, video viewing, interaction, and assignment submission. When talking about learning analytics, the display of students' learning participation, with learning process traces, is better than just the final situation or background information. Predicting students' performance based on online learning behaviors is not only a task at the technical level but also related to the identification of learning risks and

the interpretable learning for teaching support (Long & Siemens 2011; Gašević et al. 2015; Romero & Ventura 2020).

Multi-source learning activities have diverse educational meanings. Task achievement reflects the implementation of the curriculum, video viewing is related to the application of learning materials, chapter tests record the situation of knowledge acquiring, and interaction presents the social participation situation, etc. Integrating these variables into a single feature table may achieve a reasonable prediction accuracy, but difficult to clarify the reasons on which the prediction depends, and this is integrating multi-source online behaviors into the education semantic-guided behavior layer (Albreiki et al. 2021; Al-Zawqari et al. 2022; Wise & Shaffer 2015).

Graph learning offers a way to represent student connections, but building these relationship graphs is challenging. Many models build graphs from original behavioral similarities. However, such graphs capture proximity in the raw behavioral feature space but may not reflect similarity in the prediction process. For structured online-learning data, tree-based models create nonlinear partitions (Chen & Guestrin 2016; Ke et al. 2017). If two students are in the same or similar leaf nodes in the trained model more than once, they may have similar prediction-relevant decision patterns. Conversely, students with different raw behavioral profiles may still be assigned to similar leaf nodes by a trained tree-based model, indicating similar prediction-relevant decision patterns. Therefore, the association between students can not only be defined from the original behavior distance but also from the prediction structure learned inside the trained model (Kipf & Welling 2017; Zhou et al. 2020; Li et al. 2022).

This study focuses on three aspects: RQ1: how to organize multi-source online learning behaviors into multi-layer representations with educational significance; RQ2: how to construct student relationship graphs to reflect predictive similarity; RQ3: when the powerful tree model has stable predictions, how the graph structure can be used as a supplement to relationship residuals.

The proposed Tree Evidence-Guided Multi-layer Learning Behavior Graph (TEG-MLBG) framework addresses the problem with four related parts: the first is to separate online learning behaviors into levels, such as task participation. The second is to use tree-based models as the main predictors and sources of prediction evidence, including feature importance. The third is how to transform leaf nodes and leaf-weighted graphs of shared leaf nodes into student relationships that can be predicted and perceived. The fourth is how the tree-guided graph residual correction mechanism estimates relationship residual information that is added to the tree model prediction under verification control. The goal is not to completely replace tree models with asymmetric graph models, but to build a tree prediction framework that can predict and perceive student associations and supplement and complement graph residuals.

RELATED WORK

Student Performance Prediction in Online Learning

The prediction of students' performance is a key task in educational data mining and learning analytics. According to the systematic review by Albreiki et al. (2021), online learning platforms are among the main data sources. Realinho et al. (2022) used a data set of 4,424 students and 35 variables to predict dropout and academic achievement. Research by Al-Zawqari et al. (2022) shows that it is still difficult to predict high-achieving students, best-performing model achieved an accuracy of 81% and a precision of 69%. Wang and Yu (2025) found that there is a close relationship between students' grades and repeated resource viewing.

Multi-Source Behavioral Representation and Explainable Prediction

Online learning platforms can collect multi-source data of learning behaviors, and the educational meanings of these data should not be overlooked. Gašević et al. (2015) emphasized, learning analytics needs to focus on "learning" itself. Wise and Shaffer (2015) proposed that when faced with many recordable variables, we should use theories to clarify the data trajectories of meaningful learning activities and transform micro-level analysis into teaching strategies. Tiukhova et al. (2024) indicate that, when predicting students' grades, learning analytics indicators will vary with the course context, and general activities and regularity indicators have high stability, which means that relevant behavior

indicators are more suitable as the basis for prediction, compared with isolated log variables. Ujkani et al. (2024) used explainable artificial intelligence to predict the success of courses, and the accuracy of the model reached 94%. Mustofa et al. (2025) carried out a hybrid model to explore the risk of dropping out, and after processing the HLRNN model, the accuracy reached 96%. These results highlight the need to connect the prediction results with interpretable behavioral bases. The research of Herrmann and Weigert (2024) shows that the meaningful variables are the behavioral evidence reflecting the learning process. This study organizes such evidence into behavior layers, such as task engagement, resource engagement, assessment performance, and social interaction, in order to preserve the educational differences among learning behaviors.

Tree-Based Models as Predictive Evidence Sources

In the studies of students' performance prediction, tree-based models usually shows a stable performance. Hakkal and Ait Lahcen (2024) carried out a study to explore XGBoost to predict students' performance, and the AUC value reached 0.88. Chen and Guestrin (2016) proposed XGBoost to improve computational efficiency. Ahmed et al. (2025) carried out a comparative study of ten regression models in two datasets, the root mean square error of the ensemble voting regressor on the first dataset was 0.1050, the mean absolute error was 0.0837, while the coefficient of determination reached 0.9890; the coefficient of determination on the second dataset was 0.7716.

These studies suggest that because of the integrate nonlinear modeling and interpretable feature mechanisms, the tree-based models are stable in structured educational prediction. However, most of the current research only treats them as independent predictors rather than to build student relationship models. There is a research gap: tree-based models have predictive information beyond the traditional "feature importance", but they are not used to build student relationship models for prediction-oriented student relation modeling.

Student Relation Graphs and Graph Learning

Graph learning can construct the dependence relationship of data and treat samples as related structures. Zhou et al. (2020) pointed out that graph neural networks characterize nodes by aggregating information from neighborhoods through message passing. Li et al. (2020) proposed R²GCN to construct a heterogeneous student interaction problem network; the evaluation was on a real dataset of 104,113 mouse trajectories for over 4,000 students and 1,631 problems. The research shows that the graph structure can characterize the interaction between students and problems on a large scale and is superior to traditional independent sample modeling in performance prediction. Li et al. (2022) proposed Study-GNN to construct a student graph through methods with similarity learning, since many elements can be characterized as nodes or connected entities in the graph. Huang and Zeng (2024) proposed a dual graph neural network model that combines interaction and attribute graph structures, their passing/failing prediction reaches 83.96%, and the accuracy in passing/withdrawing prediction reaches 90.18%, indicating that combining interaction and attribute relationships can improve the academic performance prediction effect. Zhang et al. (2025) developed GNN-TINet for multi-label student performance prediction, achieving a prediction consistency score of 0.92 and an accuracy of 98.5%. Through exploratory analysis, meaningful associations such as the average grade point average are identified, and it is speculated that the student graph may need to go beyond the proximity of the original features and be incorporated into the prediction framework of model learning.

Research Gaps and Positioning of This Study

Recent research focuses on improving the prediction performance of students' performance by means of different modeling strategies. There are operations such as conducting early prediction from learning management system logs and optimizing random forests (Santos & Henriques 2023; Chen & Liu 2024; Wang 2024; Huang & Chen 2024). This reflects the value of online learning trajectories and advanced machine learning models in prediction. However, most research focuses on prediction accuracy or model optimization, and few pay attention to organizing multi-source learning behaviors into educationally meaningful layers and constructing relevant student relationships by using the internal decision-making structure of tree-based models.

In order to fill the gap, this study constructs a multi-layer learning behavior graph architecture led by tree-evidence. This architecture integrates online learning behaviors from multiple sources into the educational behavior level, takes the tree model as a stable prediction tool and a path-level evidence source, transforms the tree-leaf co-occurrence into a student relationship graph, and uses the validation-controlled graph residual correction to correct and improve the prediction of the tree. This study does not aim to replace strong tree-based predictors with graph models. Instead, it positions graph modeling as a relational residual supplement within a tree-based prediction framework.

METHODOLOGY

TEG-MLBG is the modeling framework. Figure 1 illustrates the overall methodological framework of the proposed Tree-Evidence-Guided Multi-Layer Learning Behavior Graph for student performance prediction. As shown in the figure, there are four modules: construction of the semantic-driven teaching layer, tree-like topic prediction and evidence extraction, leaf prediction to construct the student relationship graph, and residual correction of the verification control graph. In the empirical evaluation, this framework is implemented as a program that performs prediction, and it is then compared with based benchmarks.

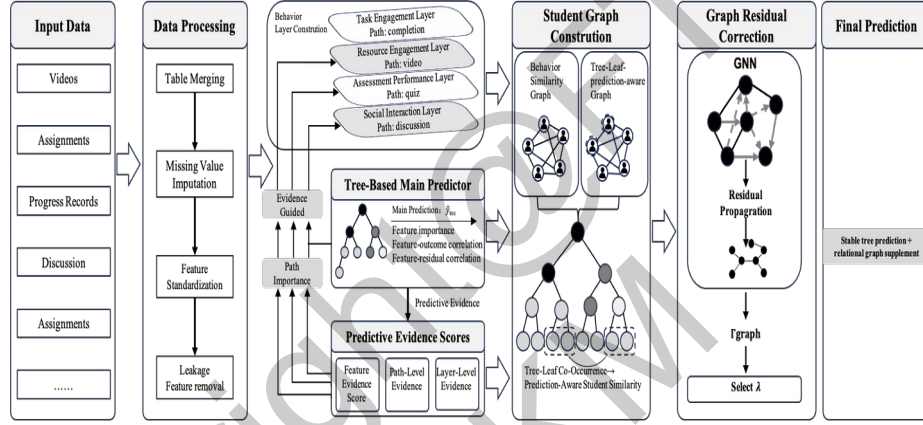


FIGURE 1. Overall Methodological Framework of Tree-Evidence-Guided Multi-Layer Learning Behavior Graphs for Student Performance Prediction.

The learner group of online courses is marked as S , and this group includes learners from $s_1, s_2, \dots, s_n$. The behavior of each learner can be represented by $x_i \in \mathbb{R}^d$. TEG-MLBG does not regard all variables as a single flat set of features, but builds up four types of educational behavior levels: indicators such as task completion belong to the task participation level; video viewing and so on are included in the resource participation level; indicators such as chapter tests are in the evaluation performance level; interaction situations such as posting are within the social interaction level.

Tree-based models are then trained as the main predictors. Given the behavior feature matrix X and performance label y , the tree model learns:

$$\hat{y}_{tree} = f_{tree}(X) \quad (1)$$

The tree residual is defined as:

$$\hat{r}_{tree} = y - \hat{y}_{tree} \quad (2)$$

Use the tree model as an evidence generation device, obtain feature importance, the association between features and results, and residuals, summarize into evidence at the path and hierarchy levels, help construct the association between the behavior layer and performance prediction, and show the under-explained behavior path information.

For student relation construction, the framework uses tree-leaf co-occurrence. Suppose the tree model contains T trees, and $\text{leaf}_t(i)$ denotes the leaf node reached by student s_i in the t -th tree. The Tree-leaf similarity between students s_i and s_j is defined as:

$$A_{ij}^{\text{leaf}} = \frac{1}{T} \sum_{t=1}^T \mathbb{1}(\text{leaf}_t(i) = \text{leaf}_t(j)) \quad (3)$$

A weighted version further combines leaf co-occurrence with tree prediction proximity:

$$A_{ij}^{\text{weighted}} = A_{ij}^{\text{leaf}} \cdot \exp\left(-\frac{|\hat{y}_{\text{tree},i} - \hat{y}_{\text{tree},j}|}{\sigma}\right) \quad (4)$$

The picture shows the similar situation of the learned prediction path and the prediction result, not just the similar situation within the original behavior space.

Finally, TEG-MLBG applies tree-guided graph residual correction. The graph component estimates relational residual information from neighboring students:

$$\hat{r}_{\text{graph},i} = \sum_{j \in N(i)} \omega_{ij} r_{\text{tree},j} \quad (5)$$

The final prediction is:

$$\hat{y}_{\text{final}} = \hat{y}_{\text{tree}} + \lambda \hat{r}_{\text{graph}} \quad (6)$$

When selecting λ in the validation set, when $\lambda = 0$, it is a degenerate tree model; when λ is greater than 0, specific folding relationships are supplemented with the help of the graph structure. This design originates from empirical observations. The correction of graph residuals is controlled by validation and depends on folding, which is not a conventional method for effective performance improvement.

RESULTS AND DISCUSSION

In this experiment, with the multi-source behavior records of the online learning platform of the modern educational technology course of a certain university in China, the final course performance is taken as the prediction object, and the platform process behavior is taken as the input feature. Variables directly related to the final score, etc., are removed before modeling.

In the 25-time test folding evaluation, Table 1 compares the predictive performance of TEG-MLBG with several baseline models. The mean absolute error (MAE) of TEG-MLBG is 5.122 ± 0.320 , the root mean square error (RMSE) is 8.758 ± 0.814 , and the coefficient of determination (R^2) is 0.876 ± 0.024 ; the MAE of the tree model baseline is 5.124 ± 0.321 , the RMSE is 8.758 ± 0.815 , and the R^2 is 0.876 ± 0.024 . The results of the two are similar, which means that the strong tree baseline captures the main nonlinear prediction signal of the structured online learning behavior data.

Pairwise statistical tests show that the instantiated TEG-MLBG achieves a perceptible improvement in MAE at the statistical level compared with the hyper-tree baseline, with an average improvement of 0.0027 and a Wilcoxon p-value of 0.026. But the improvement is not large, and the improvement in root mean square error (RMSE) and coefficient of determination (R^2) is not prominent. The existing evidence cannot show that all indicators of the graph components are always better than the tree model, and more careful interpretation is needed: the tree graph derived from leaf prediction and the graph residual correction play a limited supplementary role in some data partitions.

TABLE 1. Predictive Performance Comparison of TEG-MLBG and Baseline Models.

Model	MAE	RMSE	R ²
TEG-MLBG	5.121782026	8.757844666	0.875614646
Tree model / LightGBM_or_HGBR	5.124448562	8.758070882	0.875604308
ExtraTrees	8.792917009	13.32179051	0.713327729
MLP	9.963040173	16.14992462	0.5769445
Ridge Regression	10.91509124	16.27300491	0.572059527
Path-aggregate MLP	11.16015374	16.98401205	0.533208391

The sensitivity of λ supports this interpretation. In 25-fold cross-validation, as shown in Table 2, 10 folds select λ as 0, and the validation data in this fold does not tend to add graph residual correction; another 10 folds select the upper limit λ up to 3.00, which indicates that the graph structure can obtain residual relationship information when partially divided; the remaining 5 folds select other positive values. This distribution shows that graph residual correction is not a fixed improvement mechanism, but a supplement to verification control, and only plays a role in part of data division.

TABLE 2. Sensitivity Analysis of the Residual Correction Coefficient λ .

Lambda	Fold Count	MAE Mean	RMSE Mean	R ² Mean	Fold Ratio
0	10	5.1359778	8.751159922	0.873702749	0.4
3	10	5.166314373	9.012056559	0.869715685	0.4
0.25	1	5.219473662	9.033288481	0.877101116	0.04
0.79	1	4.669078666	7.36878111	0.916057644	0.04
1.04	1	4.696247768	8.272394202	0.892179973	0.04
2.19	1	5.424994633	9.610525009	0.854067242	0.04
2.78	1	5.066234559	7.983140069	0.903361148	0.04

In addition to prediction indicators, the current interpretable output provides initial basis for behavior layer design. Other behavior levels assist the prediction process with different weights. The results at the path level show that the prediction of performance is not dominated by a single behavior path, but rather is more inclined to be supported by the combination of multiple relatively weak behavior paths. This model conforms to the educational motivation of TEG-MLBG: online learning behavior should be constructed as a behavior framework that can track the learning process, rather than being reduced to a list of undifferentiated features.

The current results support three prudent claims: the strong tree model is a stable main predictor of structured online learning behavior data; leaf prediction can sense the student graph and graph residual correction can bring a small but perceptible MAE supplement, and its effect depends on the selection of data division and verification; multi-layer behavior representation and path evidence can provide interpretive clues in education. Currently, because graph diagnosis and other aspects still need to be supplemented, the conference paper establishes TEG-MLBG as an interpretable modeling framework, and there is evidence that non-graph models are generally stronger than strong trees in prediction.

CONCLUSION

In this study, a tree evidence-guided multi-layer learning behavior graph architecture for predicting students' performance in online learning is constructed. To deal with the flattened situation of heterogeneous online learning behaviors, educational behavior layers are integrated by integrating task engagement and so on. To overcome the limitations of the original similarity student graph, leaves and

weighted graphs are constructed from the shared leaf node membership relationship of the tree model. To solve the problem of the separation of the tree model and the graph model, the main prediction of the tree model and the residual correction of the graph are integrated.

The results produced by the current experiment should be carefully analyzed. A robust tree-like basic framework can stably carry out prediction. After instantiating TEG-MLBG, the improvement of its MAE is negligible. The λ sensitivity analysis shows that the residual correction of the graph depends on the folds and is restricted by the verification procedure, and is not fixedly effective. The graph components can be regarded as a supplementary to the relation residual and a strengthening structure of interpretability, not a substitute for the tree model.

The key contributions of this work are to construct a framework, which involves the hierarchical processing of educational semantics, the application of tree-based prediction evidence, the predictive perception modeling of student relationships, and the correction operation of graph residuals. Future work will check the diagnostic situation of the graph structure, the stable situation of layer contributions, additional ablation settings, the robustness across data partitions, and the verification situation across courses, so as to evaluate the prediction performance and educational value of the framework.

ACKNOWLEDGEMENT

The author would like to thank the students and instructors involved in the online course for providing the learning behavior data used in this study. The author also appreciates the valuable support in data collection, data preprocessing, and experimental analysis during the development of this research.

REFERENCES

- Ahmed, W., Wani, M. A., Pławiak, P., Meshoul, S., Mahmoud, A., & Hammad, M. 2025. Machine learning-based academic performance prediction with explainability for enhanced decision-making in educational institutions, *Scientific Reports* 15 (2025): 26879.
- Albreiki, B., Zaki, N., & Alashwal, H. 2021. A systematic literature review of student performance prediction using machine learning techniques, *Education Sciences* 11, No. 9 (2021): 552.
- Al-Zawqari, A., Peumans, D., & Vandersteen, G. 2022. A flexible feature selection approach for predicting students' academic performance in online courses, *Computers and Education: Artificial Intelligence* 3 (2022): 100103.
- Chen, M., & Liu, Z. 2024. Predicting performance of students by optimizing tree components of random forest using genetic algorithm, *Heliyon* 10, No. 12 (2024): e32570.
- Chen, T., & Guestrin, C. 2016. XGBoost: A scalable tree boosting system, *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2016): 785–794.
- Gašević, D., Dawson, S., & Siemens, G. 2015. Let's not forget: Learning analytics are about learning, *TechTrends* 59, No. 1 (2015): 64–71.
- Hakkal, S., & Ait Lahcen, A. 2024. XGBoost to enhance learner performance prediction, *Computers and Education: Artificial Intelligence* 7 (2024): 100254.
- Herrmann, L., & Weigert, J. 2024. AI-based prediction of academic success: Support for many, disadvantage for some?, *Computers and Education: Artificial Intelligence* 7 (2024): 100303.
- Huang, Q., & Chen, J. 2024. Enhancing academic performance prediction with temporal graph networks for massive open online courses, *Journal of Big Data* 11 (2024): 52.
- Huang, Q., & Zeng, Y. 2024. Improving academic performance predictions with dual graph neural networks, *Complex & Intelligent Systems* 10 (2024): 3557–3575.

- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. 2017. LightGBM: A highly efficient gradient boosting decision tree, *Advances in Neural Information Processing Systems* 30 (2017): 3146–3154.
- Kipf, T. N., & Welling, M. 2017. Semi-supervised classification with graph convolutional networks, *5th International Conference on Learning Representations* (2017).
- Li, H., Wei, H., Wang, Y., Song, Y., & Qu, H. 2020. Peer-inspired student performance prediction in interactive online question pools with graph neural network, *Proceedings of the 29th ACM International Conference on Information and Knowledge Management* (2020): 2589–2596.
- Li, M., Wang, X., Wang, Y., Chen, Y., & Chen, Y. 2022. Study-GNN: A novel pipeline for student performance prediction based on multi-topology graph neural networks, *Sustainability* 14, No. 13 (2022): 7965.
- Long, P., & Siemens, G. 2011. Penetrating the fog: Analytics in learning and education, *EDUCAUSE Review* 46, No. 5 (2011): 31–40.
- Mustofa, S., Emon, Y. R., Mamun, S. B., Akhy, S. A., & Ahad, M. T. 2025. A novel AI-driven model for student dropout risk analysis with explainable AI insights, *Computers and Education: Artificial Intelligence* 8 (2025): 100352.
- Realinho, V., Machado, J., Baptista, L., & Martins, M. V. 2022. Predicting student dropout and academic success, *Data* 7, No. 11 (2022): 146.
- Romero, C., & Ventura, S. 2020. Educational data mining and learning analytics: An updated survey, *WIREs Data Mining and Knowledge Discovery* 10, No. 3 (2020): e1355.
- Santos, R. M., & Henriques, R. 2023. Accurate, timely, and portable: Course-agnostic early prediction of student performance from LMS logs, *Computers and Education: Artificial Intelligence* 5 (2023): 100175.
- Tiukhova, E., Vemuri, P., Flores, N. L., Islind, A. S., Óskarsdóttir, M., Poelmans, S., Baesens, B., & Snoeck, M. 2024. Explainable learning analytics: Assessing the stability of student success prediction models by means of explainable AI, *Decision Support Systems* 182 (2024): 114229.
- Ujkani, B., Minkovska, D., & Hinov, N. 2024. Course success prediction and early identification of at-risk students using explainable artificial intelligence, *Electronics* 13, No. 21 (2024): 4157.
- Wang, J., & Yu, Y. 2025. Machine learning approach to student performance prediction of online learning, *PLOS ONE* 20, No. 1 (2025): e0299018.
- Wang, K. 2024. Optimized ensemble deep learning for predictive analysis of student achievement, *PLOS ONE* 19, No. 8 (2024): e0309141.
- Wise, A. F., & Shaffer, D. W. 2015. Why theory matters more than ever in the age of big data, *Journal of Learning Analytics* 2, No. 2 (2015): 5–13.
- Zhang, X., Zhang, Y., Chen, A. L., Yu, M., & Zhang, L. 2025. Optimizing multi label student performance prediction with GNN-TINet: A contextual multidimensional deep learning framework, *PLOS ONE* 20, No. 1 (2025): e0314823.
- Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., & Sun, M. 2020. Graph neural networks: A review of methods and applications, *AI Open* 1 (2020): 57–81.