EDITORIAL

Inaugural Editorial for the *ICCK Transactions on Educational Data Mining*

Zongwen Fan[1,*]

[1] College of Computer Science and Technology, Huaqiao University, Xiamen 361021, China

Abstract

This editorial presents the motivations underlying the establishment of the *ICCK Transactions on Educational Data Mining* (*TEDM*), an international, peer-reviewed journal dedicated to advancing theoretical, methodological, and applied research in Educational Data Mining (EDM). The journal is conceived as a platform to bring together researchers, educators, and practitioners from diverse disciplines, fostering cross-disciplinary dialogue and innovation in data-driven educational research. In particular, this editorial introduces the journal's objectives and scope, outlines representative techniques and methodological approaches employed in EDM, and highlights key trends, challenges, and opportunities that define the current landscape of the field.

Keywords: educational data mining, personalized learning, early warning systems, learning analytics, artificial intelligence in education, teaching optimization, student performance prediction, intelligent tutoring systems.

Submitted: 27 September 2025
Accepted: 29 September 2025
Published: 10 October 2025

Vol. 1, **No.** 1, 2025.
10.62762/TEDM.2025.646805

***Corresponding author:**
✉ Zongwen Fan
zongwen.fan@hqu.edu.cn

1 Introduction

Educational Data Mining (EDM) has increasingly evolved into a critical interdisciplinary field situated at the intersection of education, artificial intelligence, computer science, statistics, psychology, and the learning sciences [1]. It involves the application of data mining techniques and algorithms to large-scale, education-related datasets with the objective of uncovering latent patterns, emerging trends, and actionable insights. Such data encompass a wide range of dimensions, including student learning behaviors, academic performance, course enrollment, instructional practices, and institutional resource allocation. The overarching aims of EDM are to enhance educational quality, optimize teaching and learning processes, enrich the student learning experience, and promote equity and fairness across educational systems.

With the rapid proliferation of online learning platforms, intelligent tutoring systems, and digital classrooms, educational data have expanded dramatically in both volume and modality [2]. These data now encompass not only structured records such as grades, test scores, and activity logs, but also unstructured inputs including textual discussions, spoken interactions, and even physiological or teaching videos. While such multi-modal data offer unprecedented opportunities to generate deeper

Citation
Fan, Z. (2025). Inaugural Editorial for the ICCK Transactions on Educational Data Mining. *ICCK Transactions on Educational Data Mining*, 1(1), 1–5.

insights into teaching and learning, their effective integration, analysis, and interpretation continue to pose significant methodological and practical challenges.

The launch of the *ICCK Transactions on Educational Data Mining* (*TEDM*) represents a timely and necessary initiative to establish a rigorous platform dedicated to the dissemination of cutting-edge EDM research. This inaugural editorial introduces the journal's objectives, scope, EMD techniques, situates its mission within the broader landscape of contemporary EDM research, and reflects on the emerging trends and enduring challenges that are shaping the future trajectory of this rapidly evolving field.

2 Objectives of the journal

The main objectives of *TEDM* are as follows:

(1) To advance theoretical foundations and methodological developments in EDM, with particular emphasis on explainability and interpretability.

(2) To promote innovation in algorithms and models that enable data-driven, adaptive, and personalized learning.

(3) To bridge academic research with real-world educational practice, thereby supporting evidence-based decision-making and instructional improvement.

(4) To enhance fairness, accountability, and transparency in the application of artificial intelligence and data analytics in education.

(5) To facilitate global collaboration by fostering interdisciplinary and cross-cultural dialogue within the EDM community.

3 Scope of the journal

The scope of the journal includes but is not limited to:

(1) **Student modeling and performance prediction.** Student modeling and performance prediction constitute a central pillar of intelligent and data-driven education. By developing comprehensive models of learners, researchers can capture not only their cognitive mastery of knowledge but also their behavioral and emotional dynamics. By leveraging multi-source learning data (e.g., online quiz results, time-on-task, and engagement traces), these models enable a deeper understanding of student progress, difficulties, and optimal learning paths. When integrated with predictive analytics, student models evolve into powerful tools for forecasting learning trajectories, identifying at-risk students, and guiding personalized teaching interventions. This transforms education into a proactive and adaptive system, moving beyond one-size-fits-all instruction toward precision learning that empowers both educators and learners.

(2) **Personalized learning and adaptive tutoring systems.** Personalized learning and adaptive tutoring systems represent a transformative shift in modern education, wherein instructional strategies are dynamically tailored to the unique needs, abilities, and progress of individual students. These systems utilize diverse learning data—including quiz responses, behavioral engagement, and affective states—to adapt content, pacing, and difficulty levels in real time. Recognizing that learners differ in prior knowledge, motivation, learning style, and cognitive development, such systems ensure that each student is appropriately challenged, supported during difficulties, and guided along a learning path that maximizes both growth and confidence.

(3) **Multi-modal and cross-modal learning analytics.** Multi-modal and cross-modal learning analytics expand traditional educational analytics beyond single-source data such as test scores or log files. By integrating heterogeneous data streams—including clickstream interactions, eye-tracking, physiological signals, speech, gestures, and facial expressions—researchers can construct holistic representations of students' cognitive, behavioral, and emotional states. These approaches capture the complexity of human learning and provide richer insights into how motivation, engagement, and affect interact with knowledge acquisition, ultimately supporting more comprehensive and human-centered learning analytics.

(4) **Early warning and intervention systems.** Early warning and intervention systems have become essential for proactive student support. By continuously monitoring academic performance, behavioral patterns, and engagement signals, these systems identify at-risk learners in real time and flag potential issues such as declining

motivation, knowledge gaps, or the risk of dropout. Unlike traditional post-hoc evaluations that respond only after poor outcomes occur, early warning systems enable timely interventions, fostering a culture of prevention rather than remediation.

(5) **Course and curriculum evaluation.** Course and curriculum evaluation is critical for ensuring the quality, relevance, and effectiveness of educational programs amid rapid technological change and evolving learner needs. While traditional evaluation relied heavily on student feedback surveys and outcome-based assessments, educational data mining now enables more nuanced analyses through engagement data, performance trajectories, and multi-modal indicators of learning behaviors. These approaches allow researchers and educators to better understand not only what students learn but also how they learn, thereby informing improvements in instructional design, assessment strategies, and curriculum development.

(6) **Resource allocation and educational policy analytics.** Resource allocation and policy analytics are increasingly indispensable for ensuring equitable and efficient education systems. Traditional allocation methods—whether of funding, infrastructure, or teaching staff—often rely on broad demographic indicators, risking the neglect of context-specific needs. EDM enables fine-grained analyses of how resources influence learning outcomes across different schools, courses, and learner groups. By leveraging performance data, engagement metrics, and institutional effectiveness indicators, resource allocation can be optimized to target areas of greatest need, reduce inequalities, and promote sustainable educational development.

(7) **Fuzzy, probabilistic, and explainable models.** Fuzzy, probabilistic, and explainable models constitute an important frontier in EDM, where accuracy must be balanced with interpretability, fairness, and trust. Fuzzy models capture the inherent uncertainty and gradual nature of human learning by representing degrees of mastery rather than binary outcomes. Probabilistic models provide a principled framework for reasoning under uncertainty, accounting for incomplete or noisy data while estimating future performance outcomes. Explainable models further ensure transparency in algorithmic decision-making, offering insights that are interpretable to educators, learners, and stakeholders. Together, these approaches foster reliable, flexible, and trustworthy educational analytics.

(8) **Ethical, transparent, and equitable applications of EDM.** As EDM increasingly informs teaching, learning, and policy decisions, ethical considerations have become paramount. The collection, analysis, and use of student data entail significant responsibilities, including the protection of privacy, mitigation of bias, and assurance that interventions benefit all learners rather than privileging certain groups. Transparent practices—where algorithms, data sources, and decision-making processes are understandable and accountable—are essential for building trust in data-driven educational systems and for ensuring that technological advancements contribute to more equitable educational opportunities.

4 EDM techniques

EDM employs a diverse spectrum of computational methods to extract actionable insights from complex educational data [3], thereby supporting personalized, adaptive, and equitable learning. These methods can be broadly categorized into feature engineering, predictive modeling, unsupervised learning, sequential analysis, data balancing, and explainable modeling.

Feature selection and representation learning provide the foundation for EDM [4]. Techniques such as mutual information, recursive feature elimination, LASSO, Principal Component Analysis (PCA), autoencoders, and multi-modal feature fusion are employed to identify the most informative cognitive, behavioral, and contextual attributes. Embedding methods further capture temporal and relational dependencies inherent in sequential learning activities.

Predictive modeling constitutes a core component of EDM, leveraging supervised learning algorithms [5] including decision trees, support vector machines, logistic regression, and ensemble approaches such as Random Forest, XGBoost, and CatBoost. Neural networks further advance prediction in areas such as student performance, mastery estimation, engagement monitoring, and dropout risk. Deep learning architectures–including recurrent neural networks, long short-term memory networks, graph neural networks, and attention-based transformers–enable the modeling of temporal, relational, and multi-modal learning processes [6]. In addition, probabilistic models [7] such as Bayesian knowledge tracing and hidden Markov models, along with fuzzy inference systems, provide principled approaches for handling uncertainty and generating interpretable predictions.

Unsupervised learning techniques reveal hidden structures in educational data [8]. Clustering methods, including k-means, hierarchical clustering, density-based clustering, and self-organizing maps, facilitate the discovery of latent patterns in learning behaviors, engagement dynamics, and knowledge acquisition trajectories, enabling the grouping of students with similar needs and the design of targeted interventions. Reinforcement learning is increasingly applied in adaptive tutoring and personalized learning contexts, where agents optimize instructional actions or learning paths through iterative interactions with student models and feedback on learning outcomes.

Sequential and relational analysis techniques further enrich EDM [9]. Sequential pattern mining and process mining capture temporal dynamics in learning, modeling pathways of skill acquisition and dependencies among knowledge components. Association rule mining identifies co-occurrences between learning activities, assessment items, and performance outcomes, supporting curriculum design and adaptive recommendation systems. Dimensionality reduction and visualization approaches, including t-SNE, uniform manifold approximation and projection, and PCA-based projections, facilitate the exploration and interpretation of high-dimensional educational datasets.

Addressing class imbalance is essential for ensuring fairness and robustness in EDM applications [10]. Techniques such as over-sampling, under-sampling, adaptive resampling, and cost-sensitive learning mitigate bias and improve prediction quality for underrepresented student groups.

Finally, explainable and interpretable modeling is critical for ensuring that EDM outcomes are actionable for educators and transparent for learners [11]. Approaches such as rule-based models, fuzzy rules, SHapley Additive exPlanations and local interpretable model-agnostic explanations, and attention heatmaps provide insight into the decision-making processes of complex models. Advanced methods, including graph-based modeling, multi-task learning, and deep reinforcement learning, further extend the frontier of EDM by enabling real-time adaptive interventions and the simultaneous prediction of multiple learning outcomes.

5 Conclusion

In this editorial, we have underscored the importance of EDM in addressing the growing need for a specialized platform dedicated to the dissemination of high-quality research in this domain. The launch of *ICCK Transactions on Educational Data Mining* represents a timely response to this need, providing a venue that promotes theoretical advancement, methodological innovation, and practical transformation across the educational landscape. We envision the journal as a catalyst for scholarly dialogue, interdisciplinary collaboration, and data-informed practices that shape the future of teaching and learning.

On behalf of the editorial board, I extend a warm invitation to researchers, practitioners, and policymakers worldwide to contribute their work, share their insights, and collaborate in establishing this journal as a premier venue for advancing the frontiers of EDM and fostering global progress in data-driven education.

Data Availability Statement

Not applicable.

Funding

This work was supported without any funding.

Conflicts of Interest

The author declares no conflicts of interest.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] Khine, M. S. (2024). *Artificial Intelligence in Education*. Singapore: Springer.

[2] Liu, M., & Yu, D. (2023). Towards intelligent E-learning systems. *Education and Information Technologies, 28*(7), 7845-7876. [CrossRef]

[3] Amriza, R. N. S., Chou, T. C., & Ratnasari, W. (2025). Beyond the Classroom: Understanding the Evolution of Educational Data Mining With Key Route Main Path Analysis. *Computer Applications in Engineering Education, 33*(2), e70010. [CrossRef]

[4] Batool, S., Rashid, J., Nisar, M. W., Kim, J., Kwon, H. Y., & Hussain, A. (2023). Educational data mining to predict students' academic performance: A survey study. *Education and Information Technologies*, *28*(1), 905-971. [CrossRef]

[5] Feng, G., & Fan, M. (2024). Research on learning behavior patterns from the perspective of educational data mining: Evaluation, prediction and visualization. *Expert Systems with Applications*, *237*, 121555. [CrossRef]

[6] Bayoudh, K. (2024). A survey of multimodal hybrid deep learning for computer vision: Architectures, applications, trends, and challenges. *Information Fusion*, *105*, 102217. [CrossRef]

[7] Liu, F., Bu, C., Zhang, H., Wu, L., Yu, K., & Hu, X. (2024). FDKT: Towards an interpretable deep knowledge tracing via fuzzy reasoning. *ACM Transactions on Information Systems*, *42*(5), 1-26. [CrossRef]

[8] Derevyanchuk, O., Hu, Z., Balovsyak, S., Holub, S., Kravchenko, H., & Sapsai, I. (2025). Complex of specialized methods of educational data mining for the training of vocational education teachers. *International Journal of Modern Education and Computer Science (IJMECS)*, *17*(1), 28-46.

[9] Feng, G., & Chen, H. (2025). Educational process mining: A study using a public educational data set from a machine learning repository. *Education and Information Technologies*, *30*(6), 8187-8214. [CrossRef]

[10] Farooq, U., Naseem, S., Mahmood, T., Li, J., Rehman, A., Saba, T., & Mustafa, L. (2024). Transforming educational insights: strategic integration of federated learning for enhanced prediction of student learning outcomes. *The Journal of Supercomputing*, *80*(11), 16334-16367. [CrossRef]

[11] Niu, T., Liu, T., Luo, Y. T., Pang, P. C. I., Huang, S., & Xiang, A. (2025). Decoding student cognitive abilities: a comparative study of explainable AI algorithms in educational data mining. *Scientific Reports*, *15*(1), 26862. [CrossRef]

Zongwen Fan received both his B.S. and M.Sc. degrees from Huaqiao University, China, in 2014 and 2017, respectively. He completed his Ph.D. degree from the University of Newcastle, Australia, in 2021. He is currently a lecturer with the College of Computer Science and Technology, Huaqiao University, China. His main research interests include machine learning, data mining, and fuzzy computing. (Email: zongwen.fan@hqu.edu.cn)