PERSPECTIVE

Beyond Accuracy: Toward Interpretable, Multi-Objective, and Trustworthy Educational Data Mining Systems

Farshid Keivanian[1,2,3,4,*] **and Zongwen Fan**[5]

[1] School of Computer Science, Faculty of Engineering and Information Technology, University of Technology Sydney, Ultimo, NSW 2007, Australia
[2] Discipline of Information Technology, Australian Catholic University, North Sydney, NSW 2060, Australia
[3] Centre for Artificial Intelligence Research and Optimization (AIRO), Faculty of Design and Creative Technologies, Torrens University, Ultimo, NSW 2007, Australia
[4] Department of Information Technology, Sydney International School of Technology and Commerce, Sydney, NSW 2000, Australia
[5] College of Computer Science and Technology, Huaqiao University, Xiamen 361021, China

Abstract

Educational Data Mining (EDM) has achieved substantial gains in predictive performance, yet many existing approaches remain centered on single-objective optimization, most often accuracy. This does not adequately reflect the multi-dimensional nature of real-world educational decision-making, which requires balancing interpretability, fairness, robustness, efficiency, and timeliness. This perspective advocates a shift toward multi-objective, interpretable, and trustworthy EDM frameworks. We highlight the role of multi-objective optimization in modeling trade-offs through Pareto-optimal solutions and address the challenge of actionable decision-making through bargaining-based mechanisms, such as Nash bargaining, to select balanced and transparent outcomes. In addition, we discuss the value of fuzzy logic and adaptive methods for handling uncertainty and supporting interpretable reasoning in dynamic learning environments. Finally, we emphasize the importance of governance, accountability, and rigorous evaluation, and argue that emerging technologies should be assessed not only by performance gains but also by their practical and educational relevance. Overall, this perspective outlines a human-centered research agenda for the development of trustworthy, interpretable, and context-aware EDM systems.

Keywords: educational data mining, multi-objective optimization, interpretable AI, trustworthy AI, educational decision-making.

1 Introduction

Educational Data Mining (EDM) has developed rapidly over the past decade, driven by advances in machine learning and the growing availability of data from digital learning environments. Predictive models are now widely applied in tasks such as student

Submitted: 28 March 2026
Accepted: 08 April 2026
Published: 11 April 2026

Vol. 2, **No.** 2, 2026.
10.62762/TEDM.2026.988161

***Corresponding author:**
Farshid Keivanian
Farshid.Keivanian@uts.edu.au

Citation
Keivanian, F., & Fan, Z. (2026). Beyond Accuracy: Toward Interpretable, Multi-Objective, and Trustworthy Educational Data Mining Systems. *ICCK Transactions on Educational Data Mining*, 2(2), 52–55.

performance prediction, dropout risk detection, personalized learning, and early intervention. These developments have significantly improved the ability to identify at-risk learners and support data-driven educational decision-making.

Despite these advances, a fundamental limitation persists: most existing EDM approaches are designed to optimize a single objective—typically predictive accuracy. This narrow focus does not fully capture the complexity of real-world educational decision-making, where multiple, often competing, criteria must be considered simultaneously. In practice, models must not only be accurate but also interpretable, fair, robust, timely, and computationally efficient. For example, a highly accurate model may still be unsuitable if it introduces bias against certain student groups, lacks transparency for educators, or cannot operate efficiently in real-time systems. These challenges underscore the need for a shift from single-objective optimization toward multi-objective, human-centered frameworks that better reflect the realities of educational environments.

2 Multi-Objective Learning and Decision-Making in EDM

2.1 Limitations of Single-Objective Optimization

Traditional EDM models often prioritize accuracy as the primary performance metric. While this focus has led to improvements in predictive capability, it neglects other critical dimensions such as interpretability, fairness, and robustness. As a result, high-performing models may fail to meet practical and ethical requirements in educational applications.

2.2 Multi-Objective Optimization and Pareto Efficiency

Multi-objective optimization provides a principled framework for addressing these limitations by explicitly modeling trade-offs among competing objectives. Instead of producing a single optimal solution, these approaches generate a Pareto set of solutions, each representing a different balance among criteria such as accuracy, interpretability, and efficiency [1, 2]. However, Pareto-optimal solutions introduce a new challenge: they are not directly actionable. Educational systems typically require a single deployable model or decision rule, necessitating effective mechanisms for selecting among competing alternatives.

2.3 Bargaining-Based Decision Mechanisms

To bridge this gap, bargaining-based frameworks—rooted in game theory—offer a compelling solution. Nash bargaining [3], for instance, enables the selection of a compromise solution that balances improvements across multiple objectives while ensuring fairness and transparency [4]. Such approaches provide a principled and auditable basis for decision-making in multi-objective EDM systems and support the deployment of models that align more closely with educational priorities.

3 Interpretable and Adaptive Modeling under Uncertainty

Educational data are inherently heterogeneous, uncertain, and dynamic. Learner behaviors, engagement patterns, and contextual factors evolve over time, making static and opaque models insufficient. Fuzzy logic offers a powerful mechanism for handling uncertainty and encoding human-understandable reasoning [5–7]. By representing knowledge in linguistic and rule-based forms, fuzzy systems enhance interpretability while maintaining flexibility. When combined with adaptive metaheuristic optimization techniques, these systems can dynamically adjust to changing educational contexts, improving both performance and usability. Such approaches are particularly valuable in scenarios requiring transparency and adaptability, enabling educators to understand, trust, and effectively utilize model outputs.

4 Trust, Governance, and Accountability in EDM

As EDM systems increasingly influence high-stakes educational decisions, issues of trust, accountability, and governance become critical. Model outputs can affect student assessment, intervention strategies, and long-term learning opportunities, raising important ethical considerations. To address these concerns, transparency, fairness, and traceability must be integrated as core design principles rather than optional features. This includes the ability to explain model decisions, audit system behavior, and ensure equitable outcomes across diverse student populations. Lessons from other high-stakes domains, such as healthcare, highlight the value of combining optimization with interpretability and auditability. Similar principles should guide the development and deployment of EDM systems to ensure responsible and trustworthy use.

5 Emerging Technologies and Practical Considerations

Emerging paradigms such as quantum machine learning (QML) are gaining attention in data-driven research. While early studies suggest potential advantages, current evidence indicates that QML remains largely exploratory, particularly for structured and temporal data common in education. Significant challenges remain in areas such as data representation, scalability, and evaluation rigor. Therefore, the adoption of such technologies in EDM should be approached with caution. New methods should be evaluated not only in terms of performance gains, but also in terms of robustness, interpretability, and practical applicability in real educational settings.

6 Conclusion and Future Directions

This perspective argues that the future of Educational Data Mining lies beyond accuracy-driven modeling. To address the complex and multi-faceted nature of educational decision-making, EDM must evolve toward multi-objective, interpretable, and trustworthy frameworks. Future research should prioritize the following:

1. The development of multi-objective learning models that explicitly capture trade-offs among key criteria;
2. The design of interpretable and auditable decision mechanisms, such as bargaining-based approaches;
3. The integration of adaptive, context-aware modeling techniques for dynamic learning environments;
4. The establishment of rigorous evaluation standards, including fairness, robustness, and real-world validation;
5. The cautious and principled adoption of emerging technologies, guided by educational relevance and feasibility.

Ultimately, the goal of EDM should not be limited to building more accurate predictive models; it should be to develop intelligent systems that are trustworthy, interpretable, and aligned with the real complexities of educational decision-making. Achieving this vision will require a shift from purely performance-driven approaches toward multi-criteria, human-centred, and context-aware frameworks that are designed with educators and learners, not only for them.

Data Availability Statement

Not applicable.

Funding

This work was supported without any funding.

Conflicts of Interest

The authors declare no conflicts of interest.

AI Use Statement

The authors declare that no generative AI was used in the preparation of this manuscript.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. A. M. T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE transactions on evolutionary computation, 6*(2), 182-197. [CrossRef]

[2] Ehrgott, M., Köksalan, M., Kadziński, M., & Deb, K. (2025). Fifty years of multi-objective optimization and decision-making: From mathematical programming to evolutionary computation. *European Journal of Operational Research*. [CrossRef]

[3] Nash, J. F. (1950). The bargaining problem. *Econometrica, 18*(2), 155–162. [CrossRef]

[4] Keivanian, F., Shafiabady, N., Noman, N., Fan, Z., & Mirjalili, S. (2026). A Fuzzy Multi-Objective Neuro-evolutionary Framework with Bargaining-Based Selection for Interpretable Body Fat Prediction. *Neurocomputing*, 132974. [CrossRef]

[5] Zadeh, L. A. (2015). Fuzzy logic—A personal perspective. *Fuzzy Sets and Systems, 281*, 4–20. [CrossRef]

[6] Keivanian, F., & Chiong, R. (2022). A novel hybrid fuzzy–metaheuristic approach for multimodal single and multi-objective optimization problems. *Expert Systems with Applications, 195*, 116199. [CrossRef]

[7] Keivanian, F., Hu, Y., & Shojae Chaeikar, S. (2026). Blockchain-enabled governance for Health IoT data access via interpretable multi-objective optimization and bargaining under privacy–latency–robustness trade-offs. *Electronics, 15*(4), 864. [CrossRef]

Farshid Keivanian is a researcher in artificial intelligence and optimization, affiliated with several Australian higher education institutions, including the University of Technology Sydney and Australian Catholic University. His research focuses on multi-objective learning, interpretable and trustworthy AI, fuzzy systems, and educational data mining. He has developed bargaining-based decision frameworks and adaptive fuzzy optimization methods for complex decision-making under uncertainty, with applications in education, healthcare, and engineering systems. His work has been published in leading journals, including Neurocomputing, Expert Systems with Applications, Journal of Computational Science, and Electronics, and emphasizes human-centered, transparent, and deployable AI systems. (Email: keivanian@gmail.com)

Zongwen Fan received the B.S. and M.Sc. degrees from Huaqiao University, China, in 2014 and 2017, respectively, and the 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 research interests include machine learning, data mining, and fuzzy computing. (Email: zongwen.fan@hqu.edu.cn)