Journel of AI and Machine Learning
Volume 1, Issue 1

A Systematic Review of State-of-the-Art Explainable AI in Education

[1]Jeerawan Jeenno, [2]Oraphan Thongkam
[1,2]Department Of Physics, Faculty Of Education, Buriram Rajabhat University, Thailand

Abstract

Explainable AI (XAI) establishes artificial intelligence (AI) models which combine high-performance capabilities with enhanced transparency for school administrative choices. This systematic literature review aims to study past research about XAI use in education to connect theoretical understanding with real-world practice while advancing ethical AI practices. The current systematic literature review data was gathered from 15 peer-reviewed articles published in renowned databases, including Web of Science, Scopus, Springer, Elsevier, etc. The search was focused on studies determining the application of XAI in education, offering insights from advancements and their implications. The review examines XAI tools alongside their feature sets and operational boundaries and their fundamental needs within educational contexts. The focus of AI and ML researchers on enhancing XAI tools exists, but there are some differences between their targeted audiences and expected results. Interpretable Machine Learning (IML) or XAI produces explanations about prediction outputs while generating customized remedies through tutoring sessions. Adaptive learning systems depend on XAI to develop students' cognitive abilities for analysis and problem resolution. The intrinsic techniques of XAI in educational data science enable researchers to forecast underrepresented and underperforming student profiles and online learner success rates alongside poor course completion prospect identification for academically struggling students. XAI can help see the learned features and evaluate the bias needed for suspicion about unfair results. XAI can help see the learned features and assess the necessary bias for suspicion about unfair results.

Keywords: Artificial Intelligence (AI); AI in Education (AIED); Explainable AI (XAI); Personalized Learning; Intelligent Tutoring Systems (ITS)

Introduction

The educational sector has experienced substantial growth in utilizing Artificial Intelligence (AI) technology, enhancing personalized teaching methods. AI in education (AIED) systems employ AI to manufacture interfaces that improve educational experiences through natural language processing, speech interfaces, chatbot interfaces, and video data analysis tools (Pedro et al., 2019). Educational platforms monitor learner data by tracking their teacher interface engagements. The GDPR now establishes social data protection standards that demand both a comprehension of AI platforms and a validation of their operational authenticity from learners (Memarian & Doleck, 2023). The educational needs of Explainable AI (XAI) require systems to empower students to take control of their learning process. XAI builds interfaces that enable learners to receive support during their metacognitive processes. Despite their flaws, AI systems introduce information and facilitate student behavior which benefits the system rather than learning objectives (Khosravi et al., 2022).

The education sector benefits from AI implementation for policy development while achieving optimal management systems and delivering improved outcomes through enhanced learning process automation and improved educational opportunities, resulting in superior educational content delivery (Liu et al., 2024). XAI requires consideration of ethical consequences because educational institutions must maintain access to transparent methods everyone can understand. School dropout happens when someone leaves their classes without returning. XAI functions as a tool to generate transparency in AI models,

Jeerawan Jeenno
Department Of Physics, Faculty Of Education, Buriram Rajabhat University, Thailand

Received: 05-Jan-2024 Revised: 27-March-2025 Accepted: 08-March-2025

thus enabling their transformation from uninterpretable black boxes into semitransparent grey boxes. XAI establishes AI models which combine high-performance capabilities with enhanced transparency for school administrative choices (Adadi & Berrada, 2018). Training datasets retain prejudices that make these biases endure. Because of this, XAI models allow educators to confidently employ AI solutions using school data (Melo et al., 2022). Platforms powered by AI often replicate existing discriminatory patterns, making education and infrastructure stabilizing essential for developing nations. The researchers support improved definition systems of fairness, clearer model visibility, and preprocessing training data to achieve greater trust and responsibility in AI implementations. Studies must maintain permanent interaction between researchers and practitioners to move forward (Khosravi et al., 2022).

The educational sector is experiencing quick changes through machine learning (ML) methods, which create new predictions and enhance student success capabilities. These models' black-box structure creates interpretability issues, generating doubts about their educational deployment (Liu et al., 2024). Transition programs now require interpretable ML models because researchers dedicate increased resources toward developing and testing these predictive models for student adaptability factors. The construct of student adaptability contains multiple elements between socioeconomic factors and personal learning traits, technological abilities, and the surrounding educational context (Manna & Sett, 2024; Nnadi et al., 2024).

A wide range of research examines the "Fairness, Accountability, Transparency, and Ethics" (FATE) challenge as it appears in AIED (Memarian & Doleck, 2023). Kizilcec and Lee (2022) present a framework for understanding where the concept of algorithmic "fairness" in education starts with measurement before moving to model learning and concluding with action (Kizilcec & Lee, 2022). Baker and Hawn (2022) present a machine learning-based AIED taxonomy that enables an understanding of algorithmic bias origins alongside relevant mitigation approaches (Baker & Hawn, 2022).

Educators must explain matters differently based on who receives the information during instructional activities. Educational explanations are essential for teachers' accountability to students and parents, government, and student assessment. According to academic research, feedback is crucial because teachers need to assess student work while offering correction ideas and monitoring prompts in addition to emotional feedback. Student feedback allows learners to advance their learning while building domain knowledge, both self-regulated competencies and self-esteem and creating positive motivation patterns (Kar et al., 2024). Teaching specialists evaluate educational strategies, student learning designs, and support structures. Educational teaching scholarship depends on teacher feedback since reflecting on these assessments helps teachers generate their educational curricular pedagogy knowledge (Liu et al., 2024). Data provides strategic profiles and operational performance snapshots through business intelligence applications for management and insights development. Organizations utilize data to enhance decision-supporting functions and judgement-building processes in academic settings (Khosravi et al., 2022).

The education sector's fast-paced AI implementation has generated transparency issues in AI models that create difficulties for instructional staff, academic staff, and students. The field of XAI research lacks dedicated studies which focus on educational settings. Research today typically neglects to consider different stakeholder requirements across students and teachers alongside parents and policymakers. XAI models struggle to address societal biases because current models navigate through datasets that reflect biased norms. Research lacks systematic studies which evaluate XAI's potential to correct these prejudices when generating fair outcomes. Current implementations of XAI show limited success in improving both feedback mechanisms and automatic decision systems. This systematic literature review aims to study past research about XAI use in education to connect theoretical understanding with real-world practice while advancing ethical AI practices.

Method

Search Strategy

The current systematic literature review data was gathered from peer-reviewed articles published in renowned databases, including Web of Science, Scopus, Springer, Elsevier, etc. The search was focused on studies determining the application of XAI in education, offering insights from advancements and their implications. The search was gathered between 2019 and 2024 to ensure the inclusion of recent developments using the specific keywords listed in Table 1 below. Boolean operators, including AND, OR and NOT, enhanced the search strategy and collected data from relevant studies.

Table 1. Data Selection Strategy (Source: Author)

Years	Search Engines	Keywords
2019- 2024	✓ Google Scholar ✓ Scopus ✓ Web of Science ✓ IEEE ✓ Springer ✓ MDPI ✓ Elsevier ✓ Wiley	✓ Artificial Intelligence (AI) ✓ Machine Learning (ML) ✓ Explainable Artificial Intelligence (XAI) ✓ AI in Education (AIED) ✓ Interpretability in AI ✓ Transparent AI Models ✓ Ethical AI in Education ✓ Personalized Learning ✓ AI-Powered Feedback Systems ✓ Fairness in Educational AI ✓ Educational Decision-Making with AI

Inclusion and Exclusion Criteria

The articles were filtered based on the predefined criteria for inclusion and exclusion shown in Table 2 below to ensure quality and relevance.

Table 2. Inclusion and Exclusion Criteria (Source: Author)

Inclusion Criteria	Exclusion Criteria
Studies included were those published in full-text format and English language.	Studies not published in peer-reviewed journals were excluded.
Studies published between 2019 and 2024 were included.	Studies published before 2019 were excluded.
Studies exploring the implications of XAI in education were included.	Studies unrelated to XAI education were excluded.
Quantitative, qualitative, and mixed-methods studies were included.	None were excluded based on their methodology.

Data Extraction

Thematic analysis was used to extract data and analyze AI applications, techniques, ethical considerations and challenges. Key models of XAI in education were discussed, including personalized learning, administrative decision-making, and feedback systems. Thematic coding allowed for an in-depth exploration of the recent advancements in XAI, followed by the ethical implications, benefits, and challenges pertinent to the models.

PRISMA Framework

Figure 1 below depicts the selection of relevant studies using specific keywords, search strategy, and exclusion and inclusion criteria to gather particular data for the research aim. The PRISMA framework was applied to the selection of these studies.

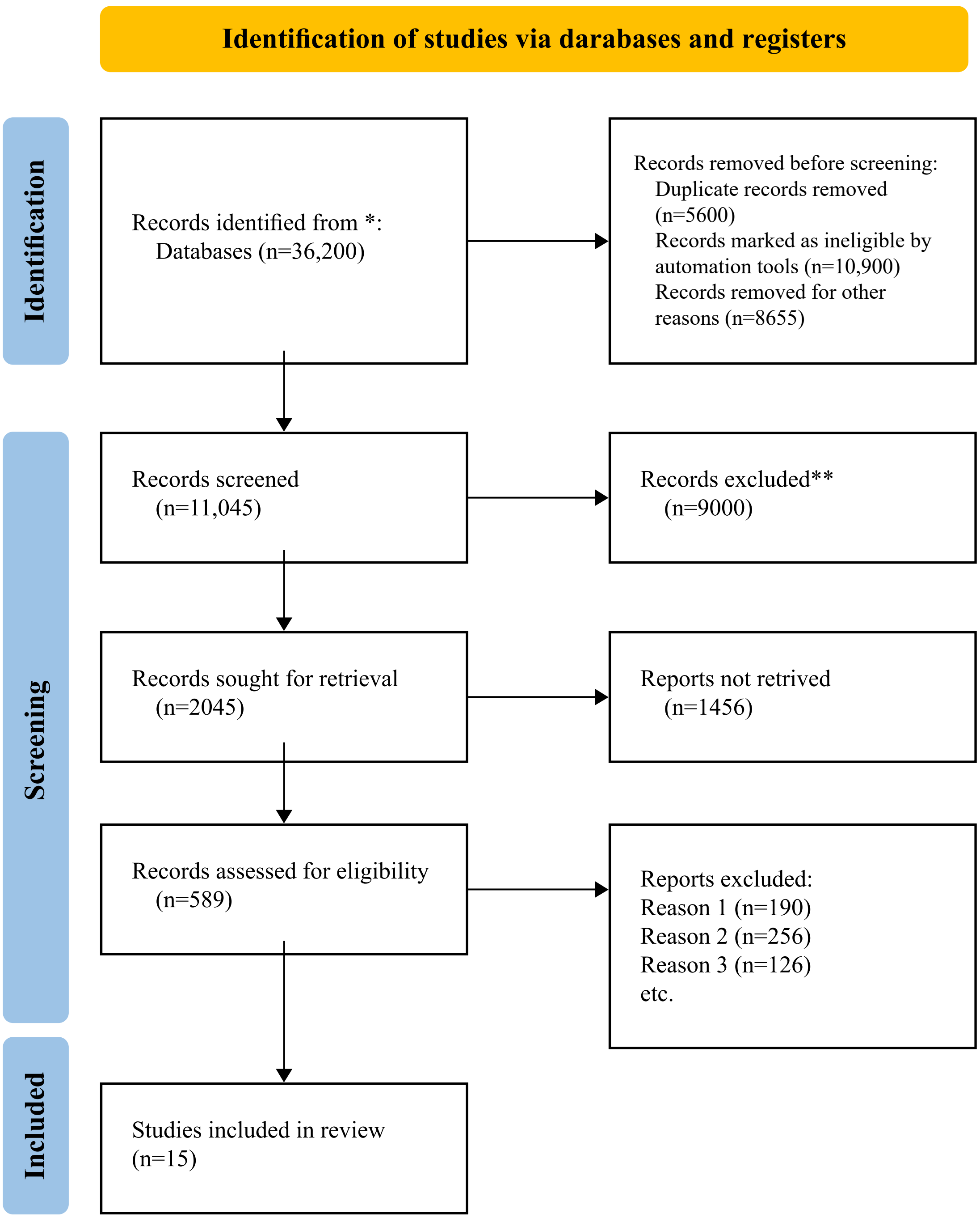

Figure 1: PRISMA Framework (Source: Author)

Results

Table 3 below presents the key findings of the selected tables for SLR.

Table 3: SLR Table for Selected Studies (Source: Author)

S.No	Author & Year	Title	Methodology	Key Findings
1.	(Akanbi et al., 2024)	A Comparative Study of Explainable AI Techniques for Bias Mitigation and Trust in E-Learning Recommendation Systems	Comparative study of XAI techniques	Demonstrates improved fairness and trust in e-learning recommendation systems using XAI.
2.	(Arrieta et al., 2020)	Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI	Taxonomy and review of XAI applications	Proposes a framework for responsible AI focusing on interpretability and ethical implications.
3.	(Ayeni et al., 2024)	AI in education: A review of personalized learning and educational technology	Review of AI-driven educational tools	AI enhances personalized learning but requires addressing ethical and privacy challenges.
4.	(Cano & Leonard, 2019)	Interpretable multiview early warning system adapted to underrepresented student populations.	Development of multiview models for underrepresented student populations	Models improve dropout prediction and early interventions for at-risk students.
5.	(Cavalcanti et al., 2024)	Towards explainable prediction feedback messages using BERT	Implementation of BERT for feedback generation	Achieves improved feedback quality and interpretability in educational contexts.
6.	(Conati et al., 2021)	Toward personalized XAI: A case study in intelligent tutoring systems	Case study in intelligent tutoring systems	Demonstrates the effectiveness of XAI in tailoring educational content and enhancing student trust.
7.	(Conijn et al., 2023)	The effects of explanations in automated essay scoring systems on student trust and motivation	Examination of trust and motivation with XAI systems	Explains the impact of AI-generated grades and transparency on student trust and motivation.

S.No	Author & Year	Title	Methodology	Key Findings
8.	(Coussement et al., 2020)	Predicting student dropout in subscription-based online learning environments: The beneficial impact of the logit leaf model	Logit Leaf Model (LLM) for dropout prediction	Combines decision trees and logistic regression for effective dropout prediction in online environments.
9.	(Fiok et al., 2022)	Explainable artificial intelligence for education and training	Review of XAI tools and their applications	Highlights the need for task-specific explanation modes in educational settings.
10.	(Gardner et al., 2019)	Evaluating the fairness of predictive student models through slicing analysis	Slicing analysis of predictive models	Identifies biases in student success predictions and suggests fairness-improving strategies.
11.	(Hall et al., 2024)	Exploring Explainability and Transparency in Automated Essay Scoring Systems: A User-Centered Evaluation	User-centered evaluation of AI grading tools	Identifies key factors like trust, user interface, and transparency for effective automated grading tools.
12.	(Liang et al., 2024)	Towards the automated generation of readily applicable personalized feedback in education	Framework for automated feedback generation	Demonstrates high-quality personalized feedback using XAI-based predictive models.
13.	(Parkavi et al., 2024)	Enhancing personalized learning with explainable AI: A chaotic particle swarm optimization-based decision support system	Development and testing of a chaotic particle swarm optimization (C-PSO) model for educational decision-making	C-PSO model outperforms traditional algorithms in efficiency, solution quality, and personalized learning outcomes.
14.	(Ogata et al., 2024)	EXAIT: Educational explainable artificial intelligent tools for personalized learning	Development of the EXAIT system	Enhances self-regulated learning and metacognitive awareness through explainable AI tools.
15.	(Whig et al., 2024)	Unveiling the Black Box: Exploring Explainable AI in Education-Trends, Challenges, and Future Directions	Comprehensive review of XAI trends and challenges in education	Highlights the need for balancing transparency, interpretability, and ethical considerations in XAI for education.

Discussion

Applications of XAI in Education

- P*ersonalized Learning Systems*

Successful student-teacher partnerships need well-designed teaching-learning processes that combine educational technology with learner data and educational data mining (EDM) tasks, learning management systems (LMS) and knowledge discovery methods. The effectiveness of learning modalities ranging from formal to non-formal and lifelong learning depends directly on adopting Information and Communication Technology (ICT) in engineering education. A previous investigation introduced an assessment model for educational technology student performance that borrows techniques from explainable artificial methods, ML, and swarm intelligence. Chaotic Particle Swarm Optimization (C-PSO) model outperforms other algorithms in efficiency, effectiveness, and solution quality (Parkavi et al., 2024).

Adaptive learning systems depend on XAI to develop students' cognitive abilities for analysis and problem resolution. Through its demonstrable nature, XAI solutions break the mystery surrounding AI decisions so students can see the reasoning behind AI-driven directions. The process facilitates learner observation of their thinking patterns while developing awareness about their cognitive processes. The artificial intelligence framework XAI enables RiPPLE to provide individualized activity suggestions which match user knowledge levels and present the system's processing rationale. Both immediate educational needs and the development of autonomous discerning thinkers emerge from XAI-equipped adaptive learning environments (Ogata et al., 2024).

Educational content generated through AI adapts teaching materials to each student's requirements, building a unified setting. The method reaches outcomes through multiple educational channels by utilizing virtual reality systems, interactive digital tools, and online learning environments. Educational administration undergoes automation because AI leads school staff to devote attention to personalized teaching approaches. Additional work needs to focus on solving privacy-related problems, algorithmic biases, and digital access barriers. Educational progress relies on an alliance between teachers, policymakers and developers of technology to build moral principles and distribute digital educational materials fairly (Ayeni et al., 2024).

- ***Intelligent Tutoring Systems (ITS)***

AI-based decision support tools developed using ITS exist to help stakeholders at higher educational institutions. Standard dropout prediction studies implement ML approaches which effectively distinguish between university dropouts and graduates. Interpretable Machine Learning (IML) or XAI produces explanations about prediction outputs while generating customized remedies through tutoring sessions (Nagy & Molontay, 2024).

The ITS provides students with algorithms which help them tackle constraint satisfaction problems. The primary objective of ITS involves developing educational platforms that incorporate network modeling to analyze student profiles while offering tailored instruction (Conati et al., 2021). Delivering specific interventions through ITS systems makes education an important context for AI because their personalized results could affect students' long-term development. Research into XAI addresses diverse applications, including recommender systems and office-based assistants together with everyday intelligent systems (i.e. Google Suggest, iTunes Genius, etc.) (Akanbi et al., 2024; Arrieta et al., 2020; Conati et al., 2021). By improving meaningful explanations about ITS activities, instructors and learners will have better acceptance and effectiveness in ITS educational systems. Research confirms that Open Learner Models enhance student learning capacities and academic outcomes and enhance students' trust while using these methods (Hooshyar et al., 2019; Zhou et al., 2016).

- ***Educational Decision-Making***

The intrinsic techniques in educational data science enable researchers to forecast underrepresented and underperforming student profiles and online learner success rates alongside poor course completion prospect identification for academically struggling students (Cano & Leonard, 2019). Virtual learning dropout risk prediction utilizes the Logit Leaf Model (LLM), which fuses decision trees with logistic regression techniques (Coussement et al., 2020).

Researchers employ LIME to explain local predictions through two primary methods: LIME implementations and Shapley Additive explanation (SHAP). Researchers deploy post hoc model-agnostic interpretable tools LIME and SHAP for ML applications in educational contexts. LIME produces comprehensible explanations for each prediction by constructing a local linear regression model approximating black-box models. SHAP delivers results based on Shapley value theory, showing single prediction descriptions and feature-based model output modification (Molnar et al., 2020; Smith et al., 2022). It is further predicted that university students' average weighted grade points can be interpreted using applied SHAP and CatBoost regressor to globally estimate and analyze the model features (Mingyu et al., 2022).

- ***Assessment and Feedback Systems***

A past study developed a learning activity framework to create predictive features for personalized feedback systems in prescriptive learning analytics (PLA). Learning activities provide the framework for this study as researchers compare generated feedback with experienced teachers' written feedback from a large-scale university course. The produced feedback quality through the PLA platform excels while maintaining prediction functionality across all evaluation stages and after completion assessment by expert teachers (Liang et al., 2024).

Students' trust levels and motivational factors are examined through this research regarding educational applications of explainable AI. The research combines need-lodgement assessments with explainable AI systems. The research findings indicate that both full-text global explanations and accuracy statements do not impact participants' trust levels or motivational tendencies. The system-generated grade and its comparison to students' personal estimates play a substantial role in determining the impact of AI explanations (Conijn et al., 2023).

According to this research, the analysis of instructor feedback through automated content detection uses BERT as the transformer model. Based on Cohen's kappa measurement, this method achieves improved results over previous approaches by 35.71%. Explainable artificial intelligence generates predictive feature insights for each classifier component the study evaluates (Cavalcanti et al., 2024).

A research paper on the Packback Deep Dives, an essay writing and grading application that incorporates an AI component, analyzed AI explainability and transparency usability. The important factors that emerged for AES assessment are Feedback: Quality and Perception, AES Performance, Trust in AES Judgment, User Concerns, AES Speed, and User Interface. They have implications for designing interventions and feedback and grading tools that are transparent and easily understandable (Hall et al., 2024).

Benefits of XAI in Education

The rapid advancement of AI has generated many advantages across multiple domains, yet experts have raised concerns over control and decision transfer risks created by uninterpretable AI systems. ML systems that rely on XAI work to eliminate trust issues while enhancing human-machine alliance. A review examines XAI tools alongside their feature sets and operational boundaries and their fundamental needs within educational contexts. The focus of AI and ML researchers on enhancing XAI tools exists with some differences between their targeted audiences and expected results (Fiok et al., 2022). The XAI approach demonstrates how AI systems generate transparency with understandable mechanisms because they provide inner-process explanations to users. Providing explanations through XAI does not prove beneficial across all user groups. The research establishes when AI systems should provide explanations while recognizing task importance and user and contextual differences. A future personalized XAI system would let AI agents formulate appropriate explanation modes for certain situations (Conati et al., 2021).

The Educational eXplainable Artificial Intelligent Too (EXAIT) symbiotic learning system uses system explanations and self-explanations to boost student understanding of their learning state while enhancing their awareness of it. The system operates throughout Japanese secondary education mathematics and English classes but shows promise to extend its usage into different academic fields along with progressive education levels. Researchers study how effectively the system scaffolds student self-explanations through computer-generated guidance. The research evaluates possibilities for system growth as a knowledge-sharing platform and its capability to function as an intellectual common among learners (Ogata et al., 2024).

Higher education proves essential for a modern economy built on knowledge and technology, yet global enrollment retention rates and completion delays affect students most severely in STEM programs. Big data found in educational administration systems can drive educational research through their combination with AI-based methods already widely established across many scientific domains. AI systems enable automated grading of texts alongside pathway personalization and early identification of at-risk students attending massive open online courses (Nagy & Molontay, 2024).

Challenges and Ethical Dilemmas in Implementing XAI in Education

Predictive AI models show enhanced accuracy when used to estimate outcomes for students organized by demographic properties like gender. Predictive model unfairness measurement depends on The Absolute Between-ROC Area (ABROCA) metric. Gardner et al. (2019) measured how five AI-based programs displayed bias in their predictions about student success in massive open online courses (Gardner et al., 2019). The team of Hutt and Gardner, together with Duckworth and D'Mello (2019), showed bias in student data-based predictive models for determining on-time graduation metrics (Hutt et al., 2019). According to Sha et al. (2021), the algorithms utilized for determining discussion forum postings exhibited bias against particular student groups (Sha et al., 2021).

Another emerging requirement arises from the fact that AI is used in many applications of people's daily lives, and when poor decisions are made due to faulty AI algorithms, such choices may have dire impacts. The European Commission's High-Level Expert Group identifies seven essentials for trustworthy AI: accountability, supervision, reliability and security, openness, balance, information and data management, tolerance, rights for all, planet and people benefits, and responsibility. Current fairness concerns are technical and biased due to discriminative AI algorithms, which also open new ethical, policy, and legal questions. Several researchers are currently trying to define fairness and develop discrimination-aware data mining. XAI can help see the learned features and evaluate the bias needed for suspicion about unfair results. Liability is when an individual will be held liable once a loss happens due to AI decisions. More work must be done to foster a culture of self-organizing decentralization and an accountability structure (Saeed & Omlin, 2023; Whig et al., 2024).

Four challenges are linked with XAI: confidentiality, complexity, unreasonableness, algorithmic bias, injustice, and explanation of justice-related dimensions. As for reasons, confidentiality means algorithms make inequality. In terms of complexity, it is hard to convey the message, and as for unreasonableness, it is seen as a fit to call rational decisions. An algorithmic bias needs to be mitigated and reasoned for, and both explainability and interpretability are key. Since stakeholders, including educational administrators and legal officers, may not understand the XAI in-depth, transparency becomes crucial. One way of encouraging the takeoff of ML in education is by making automated governance systems more transparent (Farrow, 2023; Gadekallu et al., 2024).

Conclusion

XAI establishes AI models which combine high-performance capabilities with enhanced transparency for school administrative choices. The education sector's fast-paced AI implementation has generated transparency issues in AI models that create difficulties for instructional staff, academic staff, and students. The field of XAI research lacks dedicated studies which focus on educational settings. Adaptive learning systems depend on XAI to develop students' cognitive abilities for analysis and problem resolution. Through its demonstrable nature, XAI solutions break the mystery surrounding AI decisions so students can see the reasoning behind AI-driven directions.

References

Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). IEEE access, 6, 52138-52160.

Akanbi, M. B., Okike, B., Owolabi, O., & Hammawa, M. B. (2024). A Comparative Study of Explainable AI Techniques for Bias Mitigation and Trust in E-Learning Recommendation Systems: Comparative Study of Explainable AI Techniques for Bias Mitigation and Trust in E-Learning. Journal of Institutional Research, Big Data Analytics and Innovation, 1(1).

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., & Benjamins, R. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information fusion, 58, 82-115.

Ayeni, O. O., Al Hamad, N. M., Chisom, O. N., Osawaru, B., & Adewusi, O. E. (2024). AI in education: A review of personalized learning and educational technology. GSC Advanced Research and Reviews, 18(2), 261-271.

Baker, R. S., & Hawn, A. (2022). Algorithmic bias in education. International Journal of Artificial Intelligence in Education, 1-41.

Cano, A., & Leonard, J. D. (2019). Interpretable multiview early warning system adapted to underrepresented student populations. IEEE Transactions on Learning Technologies, 12(2), 198-211.

Cavalcanti, A. P., Mello, R. F., Gašević, D., & Freitas, F. (2024). Towards explainable prediction feedback messages using BERT. International Journal of Artificial Intelligence in Education, 34(3), 1046-1071.

Conati, C., Barral, O., Putnam, V., & Rieger, L. (2021). Toward personalized XAI: A case study in intelligent tutoring systems. Artificial intelligence, 298, 103503.

Conijn, R., Kahr, P., & Snijders, C. C. (2023). The effects of explanations in automated essay scoring systems on student trust and motivation. Journal of Learning Analytics, 10(1), 37-53.

Coussement, K., Phan, M., De Caigny, A., Benoit, D. F., & Raes, A. (2020). Predicting student dropout in subscription-based online learning environments: The beneficial impact of the logit leaf model. Decision Support Systems, 135, 113325.

Farrow, R. (2023). The possibilities and limits of XAI in education: a socio-technical perspective. Learning, Media and Technology, 48(2), 266-279.

Fiok, K., Farahani, F. V., Karwowski, W., & Ahram, T. (2022). Explainable artificial intelligence for education and training. The Journal of Defense Modeling and Simulation, 19(2), 133-144.

Gadekallu, T. R., Maddikunta, P. K. R., Boopathy, P., Deepa, N., Chengoden, R., Victor, N., Wang, W., Wang, W., Zhu, Y., & Dev, K. (2024). Xai for industry 5.0-concepts, opportunities, challenges and future directions. IEEE Open Journal of the Communications Society.

Gardner, J., Brooks, C., & Baker, R. (2019). Evaluating the fairness of predictive student models through slicing analysis. Proceedings of the 9th international conference on learning analytics & knowledge,

Hall, E., Seyam, M., & Dunlap, D. (2024). Exploring Explainability and Transparency in Automated Essay Scoring Systems: A User-Centered Evaluation. International Conference on Human-Computer Interaction,

Hooshyar, D., Kori, K., Pedaste, M., & Bardone, E. (2019). The potential of open learner models to promote active thinking by enhancing self-regulated learning in online higher education learning environments. British Journal of Educational Technology, 50(5), 2365-2386.

Hutt, S., Gardner, M., Duckworth, A. L., & D'Mello, S. K. (2019). Evaluating fairness and generalizability in models predicting on-time graduation from college applications. International Educational Data Mining Society.

Kar, S. P., Das, A. K., Chatterjee, R., & Mandal, J. K. (2024). Assessment of learning parameters for students' adaptability in online education using MLand explainable AI. Education and Information Technologies, 29(6), 7553-7568.

Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y.-S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable artificial intelligence in education. Computers and Education: Artificial Intelligence, 3, 100074.

Kizilcec, R. F., & Lee, H. (2022). Algorithmic fairness in education. In The ethics of artificial intelligence in education (pp. 174-202). Routledge.

Liang, Z., Sha, L., Tsai, Y.-S., Gašević, D., & Chen, G. (2024). Towards the automated generation of readily applicable personalized feedback in education. International Conference on Artificial Intelligence in Education,

Liu, Q., Pinto, J. D., & Paquette, L. (2024). Applications of explainable ai (xai) in education. In Trust and Inclusion in AI-mediated Education: Where Human Learning Meets Learning Machines (pp. 93-109). Springer.

Manna, S., & Sett, N. (2024). Need of AI in Modern Education: in the Eyes of Explainable AI (xAI). arXiv preprint arXiv:2408.00025.

Melo, E., Silva, I., Costa, D. G., Viegas, C. M., & Barros, T. M. (2022). On the use of explainable artificial intelligence to evaluate school dropout. Education Sciences, 12(12), 845.

Memarian, B., & Doleck, T. (2023). Fairness, Accountability, Transparency, and Ethics (FATE) in Artificial Intelligence (AI), and higher education: A systematic review. Computers and Education: Artificial Intelligence, 100152.

Mingyu, Z., Sutong, W., Yanzhang, W., & Dujuan, W. (2022). An interpretable prediction method for university student academic crisis warning. Complex & Intelligent Systems, 8(1), 323-336.

Molnar, C., König, G., Herbinger, J., Freiesleben, T., Dandl, S., Scholbeck, C. A., Casalicchio, G., Grosse-Wentrup, M., & Bischl, B. (2020). General pitfalls of model-agnostic interpretation methods for MLmodels. International Workshop on Extending Explainable AI Beyond Deep Models and Classifiers,

Nagy, M., & Molontay, R. (2024). Interpretable dropout prediction: towards XAI-based personalized intervention. International Journal of Artificial Intelligence in Education, 34(2), 274-300.

Nnadi, L. C., Watanobe, Y., Rahman, M. M., & John-Otumu, A. M. (2024). Prediction of Students' Adaptability Using Explainable AI in Educational MLModels. Applied Sciences, 14(12), 5141.

Ogata, H., Flanagan, B., Takami, K., Dai, Y., Nakamoto, R., & Takii, K. (2024). EXAIT: Educational eXplainable artificial intelligent tools for personalized learning. Research and Practice in Technology Enhanced Learning, 19.

Parkavi, R., Karthikeyan, P., & Abdullah, A. S. (2024). Enhancing personalized learning with explainable AI: A chaotic particle swarm optimization based decision support system. Applied Soft Computing, 156, 111451.

Pedro, F., Subosa, M., Rivas, A., & Valverde, P. (2019). Artificial intelligence in education: Challenges and opportunities for sustainable development.

Saeed, W., & Omlin, C. (2023). Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities. Knowledge-Based Systems, 263, 110273.

Sha, L., Rakovic, M., Whitelock-Wainwright, A., Carroll, D., Yew, V. M., Gasevic, D., & Chen, G. (2021). Assessing algorithmic fairness in automatic classifiers of educational forum posts. Artificial Intelligence in Education: 22nd International Conference, AIED 2021, Utrecht, The Netherlands, June 14–18, 2021, Proceedings, Part I 22,

Smith, B. I., Chimedza, C., & Bührmann, J. H. (2022). Individualized help for at-risk students using model-agnostic and counterfactual explanations. Education and Information Technologies, 1-20.

Whig, P., Ahamad, T., Mehndi, A., Alam, N., & Yathiraju, N. (2024). Unveiling the Black Box: Exploring Explainable AI in Education-Trends, Challenges, and Future Directions. In Explainable AI for Education: Recent Trends and Challenges (pp. 81-99). Springer.

Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., & Torralba, A. (2016). Learning deep features for discriminative localization. Proceedings of the IEEE conference on computer vision and pattern recognition,