AI based Personalized Learning in Higher Education

Authors: Ali Salihi (South East European University, North Macedonia), Dayana Michelle Mora Torres (Universidad Politécnica de Valencia, Spain), Dennis Schies (Heilbronn University, Germany), Gellért Nagy (Babeș-Bolyai University, Romania), Marc Nauendorf (Heilbronn University, Germany)

Introduction

What is personalized learning?

Personalized learning is an educational approach that aims to tailor instruction, content, and learning pace to the unique needs, preferences, and goals of each individual student. Instead of applying a uniform method for all learners, personalized learning recognizes the diversity in students' prior knowledge, cognitive styles, motivations, and learning contexts. This approach often involves flexible pathways, adaptive assessments, and customized feedback, allowing learners to progress at their own speed and in ways that best suit how they learn. Personalized learning enhances engagement, fosters a deeper understanding of material, and supports more meaningful educational experiences by placing the student at the center of the learning process.

How could AI be used to help in personalized learning?

Artificial Intelligence (AI) plays a crucial role in enabling personalized learning by automating and enhancing the processes of data collection, analysis, and adaptation. AI can monitor learner behavior in real-time—tracking patterns in performance, engagement, and interaction—to build dynamic student profiles. Based on these profiles, AI systems can adjust the difficulty level of tasks, recommend personalized learning resources, and provide timely, individualized feedback. Machine learning algorithms such as Bayesian Knowledge Tracing (BKT) or Deep Knowledge Tracing (DKT) can predict a student's mastery of concepts and guide adaptive sequencing of content. Furthermore, AI-powered natural language processing can analyze student-written responses or discussion posts to assess comprehension, sentiment, and engagement. These capabilities allow AI to support a more responsive, efficient, and scalable form of personalized education.

Why is AI in personalized learning important?

AI is important in personalized learning because it makes individualized education not only possible, but scalable. In traditional classrooms, it is often impractical for a single teacher to continuously monitor and adjust instruction for each student. AI fills this gap by automating real-time analysis and customization, ensuring that every learner receives support tailored to their specific needs. This leads to more equitable learning opportunities, as AI can identify struggling students early, provide additional resources, and adjust content delivery without delays. Additionally, AI can uncover hidden patterns in learning behavior and offer insights that help educators refine their teaching strategies. Ultimately, AI enhances the precision, efficiency, and inclusiveness of personalized learning, making it a powerful tool in modern education.

Research Question

The goal of this report is to answer the question:

If the integration of AI based personalized learning systems for higher education is feasible with the current technologies?

The Evolution of AI in Education (AIED)

Before the rise of large language models (LLMs), personalization in educational technology was primarily grounded in **constructivist** and **cognitivist learning theories**, which emphasize learning as an active, contextualized process of knowledge construction (Piaget, 1970; Vygotsky, 1978). This foundation led to early AIED systems such as **Intelligent Tutoring Systems (ITS)** and **Adaptive Hypermedia Systems**, which relied on **rule-based engines** and **Knowledge Tracing (KT)** techniques to model student understanding over time (Corbett & Anderson, 1995).

A prominent technique, **Bayesian Knowledge Tracing (BKT)**, estimated a learner's probability of mastering a concept based on observable performance, updating beliefs through probabilistic inference (Baker et al., 2015). BKT was particularly effective in **structured, procedural domains** such as mathematics, spelling or learning programming syntax, where concepts are discrete and learning can be represented in well-defined steps (Van de Sande, 2013). However, BKT **struggled in less-structured domains** like essay writing or problem-solving, where student behavior is harder to model using binary knowledge states (Holmes et al., 2019).

To address BKT's limitations, researchers introduced **Dynamic Bayesian Networks (DBNs)**, which generalized BKT by modeling multiple latent variables and allowing more flexible concept interdependencies (VanLehn, 2006; Pardos & Heffernan, 2010). DBNs could incorporate contextual factors (e.g., time, fatigue) and model uncertainty more richly, making them more adaptable in realistic classroom scenarios.

---

1. Advances in AI Technologies in the last decade (2015–2025)

The introduction of **Deep Knowledge Tracing (DKT)** by Piech et al. (2015) marked a paradigm shift. Using **Recurrent Neural Networks (RNNs)**, particularly **Long Short-Term Memory (LSTM)** architectures, DKT learned hidden patterns from student interaction sequences without requiring expert-defined models. This improved predictive accuracy but came with trade-offs: **lack of interpretability**, **instability under sparse data**, and **difficulty explaining recommendations to teachers or learners** (Khajah et al., 2016; Ghosh et al., 2020).

Between 2017 and 2025, **deep learning** techniques became central to AIED research. Tools like **LSTM-based DKT** enabled systems to **capture sequential dependencies** in learner behavior, helping predict future performance more effectively than earlier models (Piech et al., 2015; Ghosh et al., 2020). However, the black-box nature of these systems posed challenges for **transparency**, **explainability**, and **stakeholder trust** (Tuomi, 2022).

Meanwhile, **Reinforcement Learning (RL) and Reinforcement Learning by Human Feedbak (RLHF)** emerged as a complementary approaches for educational decision-making, allowing agents to learn personalized content delivery policies based on reward feedback (Doroudi & Brunskill, 2019). Still, RL's reliance on large interaction data and computational overhead limited its applicability in classrooms with small cohorts or ethical constraints (Roll & Wylie, 2016; Holmes et al., 2019).

Additionally, hybrid models combining **Bayesian inference** and **neural networks**—sometimes called "interpretable deep KT"—are being developed to bridge the gap between **accuracy** and **interpretability** (Minn et al., 2022).

2. Leap Towards LLMs

Since 2022, **Large Language Models** such as GPT have introduced new paradigms for personalization. Unlike previous models, LLMs can **interpret, generate, and scaffold text-based interactions** at scale. This breakthrough is supported by emerging empirical studies that suggest LLM-based tutors can enhance student engagement and learning outcomes, particularly in small-scale implementations and in domains that involve creative or open-ended tasks (Nature Humanities & Social Sciences Communications, 2024). A key example is **Khanmigo**, launched by Khan Academy in partnership with OpenAI. It provides **Socratic-style dialogues** tailored to a student's question history. While promising, research on its long-term efficacy is still emerging (; Hwang et al., 2024).

However, LLM's personalization is mostly **interactional** rather than based on long-term student modeling (Zawacki-Richter et al., 2024; Holmes & Tuomi, 2022), which stiffles the process towards the profiling of a student's baseline of knowledge. Another concern around LLMs as personalized learning tools are AI hallucinations, where the system generates plausible-sounding but incorrect information (Ji et al., 2023). In educational settings, such hallucinations can mislead learners, provide inaccurate feedback, or reinforce misconceptions, highlighting the need for **robust validation mechanisms** and **human oversight**.

In such instances, **closed-source platforms like NotebookLM may represent a more controlled and potentially safer alternative for AI in Education**, as they limit exposure to uncontrolled model behavior and hallucinations. However, it is important to note that **NotebookLM is currently in beta** and lacks rigorous, peer-reviewed evaluation in educational contexts. Therefore, **its use should be approached cautiously and accompanied by ongoing empirical investigation and classroom experimentation**.

Current State of the Art

Tools Existing

A range of advanced AI-powered tools are now available for personalized learning, each with unique strengths and approaches:

- **Squirrel AI**
  - Uses a nano-level knowledge graph to break subjects into tens of thousands of fine-grained concepts, enabling precise identification of learning gaps.
  - Features an Intelligent Adaptive Learning System (IALS) that continuously updates each student's learning path based on ongoing assessments and practice results.
  - Has demonstrated large-scale impact, serving millions of students and achieving measurable improvements in mastery and accuracy.
- **Khanmigo**
  - Leverages GPT-4 to guide students using the Socratic method, prompting critical thinking rather than providing direct answers.
  - Adapts hints and scaffolding to student responses and aligns examples with student interests for deeper engagement.
  - Integrates seamlessly with existing classroom workflows and supports both students and teachers.
- **Other Notable Tools**
  - Adaptive learning platforms (e.g., DreamBox, Knewton) use real-time data to adjust content difficulty and learning pathways.
  - Intelligent Tutoring Systems (ITS) provide individualized feedback and targeted resources.

- Predictive analytics platforms analyze student data to forecast outcomes and recommend interventions.

Best Practices in Using AI in Personalized Learning

Based on the documents, the following best practices are essential for effective AI-driven personalized learning:

- **Comprehensive Data Collection**
  - Gather data from multiple sources: academic performance (grades, test scores), behavioral data (attendance, engagement), demographics, psychometric data, and extracurricular activities.
  - Build detailed learner profiles to inform adaptive pathways and predictive models25.
- **Advanced Analytics and Machine Learning**
  - Use algorithms such as decision trees, random forests, neural networks, and deep learning architectures (e.g., CNNs, attention-based models) to analyze student data and predict outcomes.
  - Implement ensemble methods (e.g., XGBoost, LightGBM) for robust predictions in diverse learning environments.
- **Real-Time Adaptation and Feedback**
  - Provide immediate, actionable feedback to students based on real-time performance and engagement data.
  - Adjust content difficulty and learning pathways dynamically to maintain optimal challenge and engagement.
- **Teacher-AI Collaboration**
  - Foster partnerships between educators and AI systems.
  - Train teachers to interpret AI-generated insights and maintain professional judgment over recommendations4.
- **Ethical Implementation**
  - Ensure data privacy and security by adhering to regulations (e.g., GDPR, FERPA) and implementing encryption, anonymization, and secure storage.
  - Monitor for algorithmic bias and use class balancing and explainable AI (XAI) techniques to promote fairness and transparency.
- **Continuous Monitoring and Improvement**
  - Regularly evaluate system effectiveness using metrics such as learning outcomes, engagement, retention, and skill mastery.
  - Incorporate feedback from users to refine AI tools and adapt to evolving educational needs.
- **Phased Implementation and Change Management**
  - Start with pilot programs focused on high-impact areas (e.g., early warning systems for at-risk students).
  - Gradually expand as organizational readiness and capabilities develop.
  - Involve stakeholders (teachers, administrators, students) in planning and feedback loops.

Summary Table

Tool/Platform	Key Features	Best Practice Area
Squirrel AI	Nano-level knowledge decomposition, real-time adaptation, proven scale	Advanced personalization, data-driven adaptation
Khanmigo	Socratic method, GPT-4-based, teacher support, seamless integration	Teacher-AI collaboration, critical thinking
Adaptive Learning Platforms	Real-time feedback, dynamic content adjustment	Real-time adaptation, learner profiling
Intelligent Tutoring Systems	Individualized feedback, targeted resources	Personalized support, intervention
Predictive Analytics	Outcome forecasting, intervention recommendations	Early identification, resource allocation

Conclusion

The current state of the art in AI-powered personalized learning is characterized by advanced tools like Squirrel AI and Khanmigo, which demonstrate both technical innovation and practical integration. Best practices emphasize robust data collection, sophisticated analytics, real-time adaptation, teacher-AI collaboration, ethical safeguards, and continuous improvement. These approaches ensure that AI enhances educational outcomes while respecting privacy, fairness, and human agency.

Key Aspects of personalized learning using AI

Student profiling

Why is student profiling necessary in personalized learning?

Student profiling is essential in personalized learning because it allows educators and learning systems to tailor content, pacing, and methods to the unique needs of each learner. Traditional education models often apply a one-size-fits-all approach, which can leave many students disengaged or unsupported. A well-defined student profile provides detailed insight into an individual's background, learning preferences, knowledge level, and motivation. This understanding enables adaptive systems—especially those driven by artificial intelligence—to dynamically adjust learning pathways, provide timely feedback, and select the most effective instructional strategies. As a result, student profiling significantly enhances learning efficiency, engagement, and academic performance.

What components does the student's profile have?

A comprehensive student profile consists of five major components:

1. **Knowledge Factors** – This includes the learner's existing base of knowledge, conceptual understanding, and subject-specific proficiency. It helps in identifying gaps and tailoring content difficulty accordingly.
2. **Cognitive Factors** – These refer to how the student processes information, including their learning style (e.g., visual, auditory), cognitive style (e.g., analytical vs. holistic), attention span, and working memory capacity.
3. **Motivational Factors** – These involve the learner's intrinsic or extrinsic motivations, goal orientation (e.g., mastery vs. performance), level of self-efficacy, and interest in the subject matter.
4. **Behavioral Traits** – This component includes patterns of engagement, task persistence, response to feedback, and preferences for pace or structure in learning.
5. **External and Contextual Factors** – These encompass environmental influences such as socio-economic background, cultural expectations, physical learning environment, access to technology, and time availability.

Each of these dimensions contributes to understanding how a student learns best and informs decisions about how to personalize their educational experience.

What do we have now? (What data do we already collect?)

Currently, educational institutions and platforms already collect a variety of data points, particularly in digital learning environments. These include:

- **Academic performance data**, such as grades, test scores, and completion rates.
- **Engagement metrics**, including time spent on tasks, frequency of logins, and interaction patterns.
- **Behavioral data**, such as clickstreams, submission timestamps, and activity heatmaps.
- **Learning activity outcomes**, including quiz performance, feedback requests, and forum participation.
- **Demographic data**, like age, gender, and sometimes location.

While these datasets offer useful insights into student behavior and performance, they tend to focus heavily on observable outcomes rather than underlying cognitive and motivational processes.

What is missing? (What data should we collect?)

To create a truly personalized learning experience, several important types of data are still underutilized or missing altogether:

- **Cognitive characteristics**, such as preferred learning modality, cognitive load tolerance, and processing speed.
- **Motivational indicators**, including self-reported interest, goal orientation, or effort levels.

- **Emotional and affective states**, like frustration, confusion, or boredom, which can influence learning success.
- **Contextual data**, such as current physical environment, access to quiet space, internet reliability, or familial support.
- **Learning goals and preferences**, which are often assumed but rarely explicitly collected or updated over time.

These types of data are critical for developing a holistic view of the learner and making the personalization process more precise and responsive.

How can we collect that data? (What methods are available?)

There are several methods—both technological and pedagogical—that can be used to collect the missing elements of a student's profile:

1. **Self-report questionnaires and surveys** – These can capture subjective data on motivation, learning preferences, emotions, and goals. While simple to administer, they may lack accuracy if students are not reflective or honest.
2. **Implicit data collection via learning analytics** – AI systems can infer cognitive traits and preferences through user behavior patterns, such as how learners navigate content, how quickly they respond, or how they engage with feedback.
3. **Affective computing tools** – Facial expression analysis, voice tone recognition, or wearable sensors can detect emotional states during learning activities.
4. **Diagnostic assessments and adaptive tests** – These can identify knowledge gaps, conceptual strengths, or working memory limitations.
5. **Teacher observations and input** – Educators often have valuable insights that can be incorporated into the profile, especially for traits not easily measurable by machines.

Combining multiple data sources (a multimodal approach) is often the most reliable strategy to develop accurate and dynamic student profiles. As technologies mature, these data collection methods will continue to evolve and improve in precision and ethical sensitivity.

Limitations of using AI in personalized learning

AI-driven personalized learning offers significant potential for tailored education, but its feasibility is challenged by various limitations.

Ethical Concerns and Societal Implications

Ethical challenges are crucial for responsible AI adoption in education.

Data Privacy and Security

AI systems collect extensive sensitive student data, raising concerns about security, transparency, and misuse. Without robust safeguards, data is vulnerable to breaches or exploitation, undermining trust and potentially leading to discrimination. Compliance with regulations like FERPA and GDPR is essential. A trust deficit can hinder engagement and innovation.

Algorithmic Bias and Fairness

AI algorithms can reinforce or amplify biases from their training data, leading to inequitable treatment and discriminatory outcomes, especially for marginalized groups. This can result in unfair assessments or learning path recommendations, widening existing achievement gaps. Auditing for bias and careful dataset selection are crucial for fairness.

Equitable Access and the Digital Divide

AI in education can exacerbate existing digital disparities due to unequal access to technology and high implementation costs. This creates a two-tiered system where well-resourced institutions benefit more, widening socio-economic divides. Prioritizing equitable access is vital for inclusive learning.

Technical and Implementation Hurdles

Practical and infrastructural challenges impact AI's feasibility and scalability.

Data Management and Interoperability

Effective adaptive learning requires comprehensive learner data, but this data is often fragmented across disparate systems (data silos), limiting holistic analysis. Poor interoperability restricts scalability and the creation of unified student profiles, hindering truly adaptive support. Unified data standards are urgently needed.

High Implementation and Maintenance Costs

The substantial financial investment for implementing, maintaining, and updating AI systems, along with staff training, is a significant barrier for many institutions, especially those in underserved communities. These costs exacerbate the digital divide and limit equitable access to advanced AI tools.

Accuracy and Unpredictability of AI Output

AI systems are only as reliable as their training data; inaccurate or biased data leads to flawed outputs. Students frequently report needing to verify AI-provided information due to concerns about accuracy, unverified sources, or outdated data. Complex problems may also be beyond AI's current capabilities.

Faculty Adoption, Training, and Resistance

Successful AI integration depends on educators' willingness and capability to use these tools, often hindered by insufficient training, concerns about academic freedom, and a lack of trust. AI reshapes educators' roles, requiring significant professional development and a new understanding of their responsibilities.

Scalability Across Diverse Contexts

Adapting AI systems across diverse educational contexts (disciplines, languages, student populations) is resource-intensive and challenging for generalizability. Most research is in technologically advanced regions, limiting applicability elsewhere. Lack of data diversity can also lead to AI bias, particularly for vulnerable groups.

Recommendations

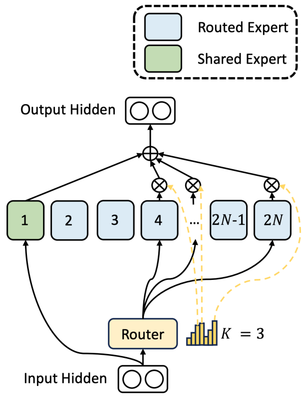
Future efforts must focus on:

- **Ethical Frameworks:** Developing robust data protection, transparency, and accountability mechanisms.
- **Technical Solutions:** Addressing data fragmentation and improving AI transparency (Explainable AI).
- **Implementation Strategies:** Prioritizing equitable access, comprehensive teacher training, and comparative studies across diverse contexts.

Regarding the technical limitations there are new ways of designing AI architectures that can help with

Advanced MoE Architectures: Fine-Grained Specialization and Shared Knowledge

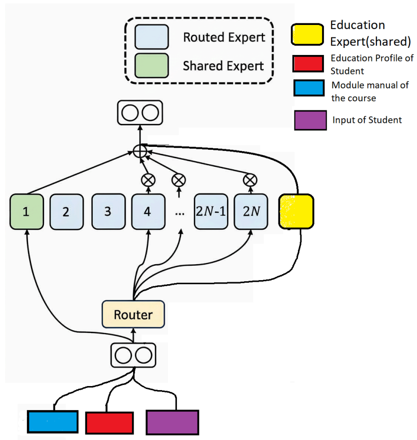
Building on the MoE foundation, newer models like DeepSeek have introduced **fine-grained expert segmentation**, where numerous smaller, highly specialized experts are utilized. This allows for even more precise resource allocation, as the router can select the most relevant small experts for a given task. The latest innovation integrates **shared experts**—general-purpose experts that are always active, providing foundational knowledge that complements the specialized insights from selected experts. This hybrid approach delivers both broad understanding and highly specialized responses, enhancing overall quality and efficiency.



Architecture Additions: Shared Education Framework

This evolutionary path extends to specialized applications, such as a proposed **Shared Education** framework. This framework leverages the MoE architecture to create a personalized learning experience. It features separate input blocks for staff data entry and student profile creation. The core MoE dynamically adapts its processing based on these inputs, with a crucial **Education Expert** responsible for constructively processing student interaction data, student profiles, and task requirements.

At the heart of this educational framework is a **router** that individually determines which experts process each learner's history. This selection is informed by two key contextual sources: the **Education Profile of the Student** and the **Module Manual of the Course**. This dual input allows the router to tailor expert selection to individual learning patterns and needs.



The framework distinguishes between two types of experts:

- **Shared Experts:** These experts, such as those responsible for modeling **general learning behaviors** like forgetting curves, spaced repetition, or content review, are available to all learners and serve as a common cognitive infrastructure.
- **Routed Experts:** These experts are specifically assigned by the router to individual learners or groups to capture unique or context-dependent learning needs.

Conclusion

To Answer the question **If the integration of AI based personalized learning systems for higher education is feasible with the current technologies,** we can say that even though there are technologies that address some of the ethical and technical limitations the feasibility of AI-based personalized learning systems is still significantly constrained by ethical (data privacy, bias, access) and technical (data management, costs, accuracy, adoption, scalability) limitations. Current technologies require strategic and ethical considerations for beneficial implementation at scale. Ultimately, successful AI integration requires a balanced, human-centered approach that empowers educators and students, enhancing human connection, academic integrity, and equitable access.

Future Research

Looking ahead, the integration of synthetic educational media—such as AI-generated personalized summaries, flashcards, and explainer videos—represents a promising frontier for AIED. By leveraging multimodal generation models, these tools can dynamically adapt content format and complexity to suit individual learners' preferences, learning speeds, and cognitive load. While early platforms (e.g., Quizlet or Synthesia) show potential, there is a critical need for empirical validation regarding their pedagogical effectiveness, cognitive impact, and ethical use. Future research should also explore how to align these synthetic outputs with curriculum standards and teacher oversight to ensure quality, transparency, and learner trust.

Sources

Limitations

Aithor. (n.d.). *Ethical concerns of AI-powered personalized learning in education*. Retrieved from https://aithor.com/essay-examples/ethical-concerns-of-ai-powered-personalized-learning-in-education

Learning Guild. (n.d.). *The impact of learner data on adaptive AI-driven learning systems*. Retrieved from https://www.learningguild.com/articles/the-impact-of-learner-data-on-adaptive-ai-driven-learning-systems/

MDPI. (n.d.). *2813-4346/4/2/17*. Retrieved from https://www.mdpi.com/2813-4346/4/2/17

ResearchGate. (n.d.). *Personalized Learning and Beyond: Ethical Implications of AI in Modern Education*. Retrieved from https://www.researchgate.net/publication/391584998_Personalized_Learning_and_Beyond_Ethical_Implications_of_AI_ir

SpringerLink. (n.d.). *10.1186/s40561-023-00269-3*. Retrieved from https://link.springer.com/article/10.1186/s40561-023-00269-3

University of Illinois Urbana-Champaign, College of Education. (2024, October 24). *AI in schools: Pros and cons*. Retrieved from https://education.illinois.edu/about/news-events/news/article/2024/10/24/ai-in-schools--pros-and-cons

Scientific Publications on LLMs, Agents, and Knowledge Representation

Zhang, R., Zhang, Z., & Zhu, C. (2024). *Towards Cognitive-Level Instruction Understanding in Large Language Models*. arXiv. https://arxiv.org/abs/2504.08989
Cui, L., Liu, S., Jin, Z., Zhang, L., & Wang, H. (2024).
MindChain: Modeling Agent Chain of Thought via Mind Graph. arXiv. https://arxiv.org/abs/2504.14928

Wang, Y., Zhu, H., Tang, Y., Chen, J., Wu, Y., Gong, Y., ... & Zhang, R. (2024). *OpenAGI: When LLM Meets Domain Experts*. arXiv. https://arxiv.org/abs/2503.11486

Zhang, C., Zhang, X., Shao, Y., Pan, X., Xu, K., Zha, H., & Liu, Y. (2024). *MoKai: Towards Modularized and Knowledge-aware Agent Framework*. *International Conference on Learning Representations (ICLR)*. https://openreview.net/forum?id=nTAC2NCQUO
Wang, Y., Tian, Y., Song, C., Yang, Z., & Zhang, J. (2024). Gradient Disentangled Global Reasoning for Few-Shot Semantic Segmentation.
IEEE Transactions on Neural Networks and Learning Systems. https://doi.org/10.1109/TNNLS.2024.3380342

Bayesian Knowledge Tracing & Extensions

Baker, R. S., Corbett, A. T., & Aleven, V. (2015). More accurate student modeling through contextual estimation of slip and guess probabilities in Bayesian Knowledge Tracing. In H. W. Woolf, E. Ameur, R. Nkambou, & S. Lajoie (Eds.), Intelligent Tutoring Systems (Vol. 5091, pp. 406–415). Springer. https://doi.org/10.1007/978-3-642-14363-2_41

van de Sande, B. (2013). Properties of the Bayesian knowledge tracing model. Journal of Educational Data Mining, 5(2), 1–10. https://doi.org/10.5281/zenodo.3554629

Pardos, Z. A., & Heffernan, N. T. (2010). Modeling individualization in a Bayesian networks implementation of knowledge tracing. User Modeling and User-Adapted Interaction, 20(4), 389–417. https://doi.org/10.1007/s11257-010-9079-1

Corbett, A. T., & Anderson, J. R. (1995). Knowledge tracing: Modeling the acquisition of procedural knowledge. User Modeling and User-Adapted Interaction, 4(4), 253–278. https://doi.org/10.1007/BF01099821

Deep Knowledge Tracing

Piech, C., Bassen, J., Huang, J., Ganguli, S., Sahami, M., Guibas, L., & Sohl-Dickstein, J. (2015). Deep Knowledge Tracing. Advances in Neural Information Processing Systems, 28, 505–513. Retrieved from http://papers.nips.cc/paper/5654-deep-knowledge-tracing

Khajah, M., Lindsey, R. V., & Mozer, M. C. (2016). How deep is knowledge tracing? In Proceedings of the 9th International Conference on Educational Data Mining, 94–101. Retrieved from https://arxiv.org/pdf/1604.02416.pdf

Ghosh, S., Heffernan, N., & Lan, A. (2020). Context-aware attentive knowledge tracing. Proceedings of the AAAI Conference on Artificial Intelligence, 34(05), 7854–7861. https://doi.org/10.1609/aaai.v34i05.6180

Interpretable / Hybrid Models

Minn, S., Yu, Z., Chen, Y., & Liu, M. (2022). Interpretable Deep Knowledge Tracing: Improving student performance prediction with attention and factorization. IEEE Transactions on Learning Technologies, 15(2), 162–174. https://doi.org/10.1109/TLT.2022.3140024

Limitations & Trust

Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial Intelligence in Education: Promises and Implications for Teaching and Learning. Center for Curriculum Redesign.

Tuomi, I. (2022). The algorithmic paradoxes of personalization in education. Learning, Media and Technology, 47(2), 160–175. https://doi.org/10.1080/17439884.2022.2047343

Reinforcement Learning in Education

Doroudi, S., & Brunskill, E. (2019). Does combining student models improve performance? An experimental approach. Journal of Educational Data Mining, 11(1), 1–31.

LLMs & Harms

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., ... Bang, Y. (2023). Survey of hallucination in natural language generation. ACM Computing Surveys, 55(12), 1–38. https://doi.org/10.1145/3571730

Emerging Empirical Evidence

Wang S. et al. (2024). “Large Language Models for Education: A Survey and Outlook.” arXiv.Org. April 1. https://arxiv.org/abs/2403.18105.