ARTIFICIAL INTELLIGENCE IN MATHEMATICS EDUCATION AND ITS COMPUTATIONAL APPLICATIONS

I. D. Pawade
Assistant professor of Mathematics, Shivramji Moghe Arts, Commerce and Science College Kelapur (Pandharkawada) Dist. Yavatmal (Maharashtra)
ipawade37@gmail.com

Abstract
In this paper we examine the current role of artificial intelligence (AI) in mathematics education and explores recent studies, system approaches and computational applications that are transforming the teaching, learning, evaluations and research of mathematics. Also, we bring together results from systematic reviews and research studies on Intelligent Tutoring Systems (ITSs); adaptative learning platforms, automated grading, symbolic computation, original research on theorem - proving hybrids and the growing use of large language models (LLMs) as teaching tools. We discuss learning improvements, design models, technical methods and real word challenges such as scalability, fairness, transparency, data privacy and academic honesty. Finally, we suggest research paths and practical guidelines for teachers, developers and policymakers who wish to integrate AI into mathematics education.
***Keywords:** Artificial intelligence, mathematics education, intelligent tutoring system, automated theorem proving, neuro-symbolic AI, pedagogy.*

1. Introduction

Artificial intelligence (AI) originated as an idea in ancient times but became a formal field in the mid-20th century with the arrival of digital computers. The term "Artificial Intelligence" was introduced in 1956 at the Dartmouth Conference, marking the start of systematic AI research. Since then, AI has evolved from rule-based systems to advanced methods like machine learning, deep learning and natural language processing. Today, AI plays a major role in mathematics, helping with reasoning, learning, problem solving and decision making often performing tasks as well as or better than humans.

Mathematics education continues to face persistent challenges such as large classroom sizes, diverse students' backgrounds, limited opportunities for individualized feedback and the difficulties in evaluating higher order reasoning skills like proof construction. Artificial intelligence holds the potential to customize learning experiences, offer real time feedback, streamline routine assessments and advanced mathematical reasoning and discovery. Since 2020 the field has grown quickly because of better machine learning models, improved symbolic tools and release of powerful LLMs, resulting in more research studies and system use in schools, colleges and research contexts. Studies and reviews from 2023 to 2025 show a fast increase in research papers and practical applications related to mathematics education. This paper brings together recent research and technical progress, focusing on computational tools that directly help teaching and learning. These include Intelligent Tutoring Systems (ITS) and adaptive learning problems, automated grading (even for proofs), symbolic computation and computer algebra tools and the growing use of LLMs and AI assisted theorem proving. We explore both the educational results and system design aspects followed by a discussion of their ethical, practical and research implications.

2. Definition and Background

- **Artificial Intelligence in Education:** Any computational method that models, supports or arguments the teaching-learning process. In mathematics this ranges from rule-based tutors to deep learning models used for natural language feedback.
- **Intelligent Tutoring System (ITS):** These system track students understanding and deliver personalized instruction and feedback often using cognitive or Bayesian models. Intelligent Tutoring systems (ITS) are most well-established AI applications in mathematics education research.
- **Symbolic Computation and Theorem Proving:** Automated theorem provers and computer algebra system that manipulate mathematical expressions and generates proofs such as solution verification and open new possibilities for tasks like hint generation.
- **Automated Grading:** AI is used to evaluate student work from multiple choice questions to open ended algebra problems and formal proofs. Recent methods combine large language models (LLMs) with symbolic verification.

3. Literature Review

3.1 Systematic Reviews and Meta Analysis:

Several reviews and studies from 2015 to 2025 show that research in this area has grown quickly. The main topics include intelligent tutoring systems, adaptative learning, formative assessment, auto grading and intelligent feedback using natural language processing and experiments that combine machine learning with symbolic tools. This study shows generally positive but mixed results for student learning. Some ITS and adaptive learning systems lead to better test scores and quicker problem solving but impact levels and research quality vary widely. Reviews recommended conducting more rigorous randomized controlled (RCTs), longer term studies and providing clearer details about how the systems are designed.

3.2 ITS Adapative Learning Findings:

Intelligent Tutoring Systems (ITSs) for mathematics from basic arithmetic to college level calculus consistently show positive results when they are well designed and curriculum standards. Their success depends on students' models, guided hints and mastery-based learning approaches. However, many existing studies are limited in size or duration, and reviews highlight the need for more replication in real classroom environments.

3.3 Autograding and Assessments

Recent studies show that modern AI models can automatically grade structured answers and even open-ended proofs quite accurately when paired with symbolic checking and specialized training data. These systems can give students step-by-step feedback and lessen teachers grading load. However, their accuracy decreases when faced with unfamiliar or unique solution methods, so careful validation against human grading is still crucial.

3.4 Theorem Proving and Symbolic Hybrids

Research integration that and deep learning including LLMs with automated theorem provers has progressed rapidly. Machine learning helps guide proofs researchers, suggest lemmas and turn informal reasoning into formal steps. This approach shows great potential for both mathematical research and education such as generating intermediate steps or verifying students' work but fully automating complex proof discovery is still an unsolved challenge.

4. Computational Applications

In this section we explain common computational applications, how they describe and how they usually work in teaching and learning.

4.1 Intelligent Tutoring System (ITS)

Architecture components

- Domain model (skills, problem structure).
- Students model (knowledge tracing).
- Pedagogical module (sequencing, hint policy).
- Interface (worked examples, step entry, visual manipulatives).

Methods

- Knowledge tracing: classic Bayesian Knowledge Tracing (BKT), Deep Knowledge Tracing or hybrid methods.
- Reinforcement learning for hint and problem selection policies.
- Constraint based modeling for error detection in algebraic manipulations.

Use cases

- Step-by -step algebra problem solving with hints and error diagnosis.
- Scaffolded geometry proofs using interactive diagrams.
- Mastery-based sequencing for arithmetic fluency.

Evidence and limits

Studies show that Intelligent Tutoring System (ITS) can improve learning when they match the curriculum and are used often, but challenges like teacher training and classroom integration still exist.

4.2 Autograding and Feedback Systems

Architectures: Some systems use rule-based methods to read student answers, simplify them and compare them with correct ones using a computer algebra system (CAS). Others use machine learning or language models to understand written answers, categorize them and give feedback. A mixed approach uses ML to interpret responses and symbolic tools to check correctness.

Pedagogical benefits: They provide quick feedback, make grading easier for many students and offer personalized learning hints.

Risk: They may misunderstand new ways of students reasoning depends on patterns or make grading errors, so teachers still need to monitor results.

4.3 Symbolic Computation and CAS integration

Role of CAS: Computer algebra system performs exact algebraic simplifications, symbolic differentiation or integration and solve equation. They also act as a reliable reference for automatic grading and checking student answers.

Hybrid ML Symbolic approaches: Machine learning suggests solution steps or transformations while the CAS checks if they are correct or provides counter examples. This makes the system more accurate and reliable in checking mathematical correctness.

4.4 Theorem Proving and LLM assisted Proof Support

Automated and interactive theorem provers like Isabell and Coq, now enhanced with machine learning, helps in formal mathematics by translating natural proofs into formal scripts and suggesting proofs steps. In education they check the logic of student proofs and offer hints. Though ML has improved proofs guidance, complete automation for complex problems is still developing.

4.5 Large Language Models (LLMs) and Generative Tools

Large Language Models (LLMs) can explain concepts, give step-by-step solutions and act as conversational tutors. They Power apps that solve photographed problems though their accuracy varies for complex topics. Used wisely, LLMs can support learning by offering hints and multiple solution paths but they should be combined with verification tools to prevent mathematical errors and misuse.

5. Measured Impacts on Learning

Research and reviews indicate that while results vary, well-designed AI tools generally show positive short-term effects on learning. Stated advantages include improved test performance, quicker skill mastery, higher student engagement through interactive feedback and reduced grading workload for teachers enabling them to focus on more meaningful tasks. However, experts highlight the need for stronger research design, long-term studies, assessments of deeper reasoning skills and analysis of equity outcomes.

6. Challenges and Risks

6.1 Validity and Robustness

AI systems, especially those using machine learning may not work well when students use unfamiliar methods or tricky inputs. Using symbolic checks helps reduce these mistakes but can't fully remove the risk.

6.2 Pedagogical Transparency

Teachers and students should know why an AI gives a certain hint. If the reasoning behands an LLMs response isn't clear or verified. It can reduce confidence and trust in the system.

6.3 Fairness and Access

AI systems trained on biased data can be unfair to underrepresented students and schools with limited digital resources may not get the same benefits as well-equipped ones.

6.4 Academic Integrity

Generative AI makes it easy for students to produce homework answers. Teachers need to update assessments by using more in-class, project-based, or oral methods and apply tools to detect or verify AI use. Many schools and colleges are already adapting to these Changes.

6.5 Data Privacy and Security

AI systems that record detailed student activity must follow privacy laws and ensure strict rules for how data is used, stored and shared.

7. Design Principles and Best Practices

Effective use of AI in education depends on keeping teachers involved, combining machine learning with formal checks for accuracy, supporting AI tools with the curriculum, giving clear and helpful feedback and testing systems thoroughly across diverse learners. These practices ensure AI remains reliable, fair and supportive of real classroom learning.

8. Future Directions

8.1 Models for Mathematical Reasoning

Integrating symbolic methods with machine learning is more reliable models for proof checking, hint generation and intermediate step verification. Ongoing efforts to connect LLMs with symbolic tools are especially valuable for improving reliability and performance.

8.2 Scalable classroom Level Trials

There is a strong need for large, multi-site studies that track long-term learning outcomes, equity impacts and how well students apply skills to problem solving. Recent reviews (2024-2025) specifically emphasize the importance of conducting such research.

8.3 Autograding for Higher-order Skills

Improving AI based autograding to accurately evaluate proofs, modeling and reasoning tasks can strengthen its role in measuring advanced mathematical thinking. Early research on grading induction proofs shows encouraging results.

8.4 Responsible Use Policies and Tooling's

Develop standards and tooling for attribution, verification and privacy, and co-design policies with educators to guide responsible deployment.

8.5 AI as Research Partner in Mathematics

AI is increasingly becoming a research partner in mathematics, contributing to new discoveries and assisting with complex proofs. As these systems improve, they can help advanced students explore conjecture and problem-solving methods. The 2025 milestone where AI models achieved gold-level performance on difficult math competitions highlights rapid progress and educational potentials.

9. Recommendation for stakeholders

For Educators

- Test AI tools in small, controlled setting while tracking how students perform and respond.

- Use AI to reduce time spent on routine grading so teachers can focus on interactive, discussion-based learning.
- Update homework and assessments to emphasize reasoning, problem-solving and project work, especially when students have access to generative AI tools.

For Developers

- Pair machine learning with symbolic verification and record the source of every feedback item for transparency.
- Provide teacher dashboards that clearly explain how the AI makes decisions and allow teachers to review or override them when needed.

For Policymakers

Support infrastructure and teacher training, promote transparent evaluation and mandate student-data protections.

10. Conclusion

Artificial intelligence (AI) is bringing transformative changes to multiple aspects of mathematics education. It offers personalized learning through intelligent tutoring systems (ITS), scalable assessment via autograding, enhances problem-solving using symbolic methods, machine learning combinations and conversational assistance from large language models (LLMs). Recent studies highlight positive progress but also point out research gaps and ethical concerns. The most productive path forward is careful, human-Centre integration, hybrid systems that pair ML's flexibility with symbolic rigor, rigorous classroom evaluation and policies that prioritize equity, transparency and data protection. Continued collaboration among educators, mathematicians, ML researchers and policy makers is crucial to use AI responsibly and effectively.

References

1. Son, T. (2024). *Intelligent Tutoring Systems in Mathematics Education*. MDPI.
2. Garzón, J., et al. (2025). *Systematic Review of Artificial Intelligence in Education*. MDPI.
3. Zhao, C., et al. (2024). *Autograding Mathematical Induction Proofs with Natural Language Models*. arXiv.
4. Li, Z., et al. (2024). *Survey: Combining deep learning with theorem proving*. arXiv.
5. Buchberger, B. (2023). *Automated Programming, Symbolic Computation, and Machine Learning*. Springer.
6. Opesemowo, OAG. (2024). *A Systematic Review of AI in Mathematics Education*. EJMSTE.
7. Holman, K., Vasquez, T., Taub, M., et al. (2025). *Artificial Intelligence Interventions in Mathematics Education*. ERIC / EJ1481890.
8. Yıldız, SG. (2025). *Trends and insights of AI in mathematics education: A bibliometric analysis*. IEJME.
9. Wired. *Generative AI Transformed English Homework. Math Is Next.* (2024).
10. Reuters. *Google and OpenAI's AI models win milestone gold at global math competition.* (July 21, 2025).