JURNAL PENDIDIKAN MATEMATIKA (JPM)
Vol. 11, No. 2, 2025, pp 141–156 e-ISSN: 2656–4564
https://doi.org/10.33474/jpm.v11i2.24319
online at: https://riset.unisma.ac.id/index.php/jpm/article/view/24319

BRIDGING PEDAGOGY AND AI: A SYSTEMATIC REVIEW OF DEEP LEARNING IN MATHEMATICS EDUCATION

Valeria Yekti Kwasaning Gusti*, Fitria Amastini
Universitas Terbuka Banten, Indonesia
Email: *valeria.gusti@ecampus.ut.ac.id*

ABSTRACT
Deep learning is interpreted differently in educational and computational contexts. The purpose of this study is to systematically examine how deep learning is conceptualized and applied in mathematics education research, and to evaluate the extent to which these applications align with the Indonesian Ministry of Education and Culture's (Kemendikbud) definition of deep learning. A systematic literature review (SLR) was conducted using PRISMA guidelines. Searches in ScienceDirect, Scopus, Springer Link, and ProQuest produced 1,881 records containing the term "deep learning" published between 2015 and 2025. After duplicate removal and screening, 16 peer-reviewed journal articles explicitly addressing deep learning in mathematics education were included. Two main interpretations were found. Eleven studies framed deep learning pedagogically, focusing on conceptual understanding, problem-solving, collaborative learning, and real-world application. Five studies adopted a computational framing, using deep neural networks and other machine learning techniques for prediction, error analysis, adaptive instruction, and automated feedback. While the majority aligned with Kemendikbud's pedagogical emphasis, some studies treated deep learning purely as a technical method, without explicit links to student-centered outcomes. The review highlights a conceptual gap between pedagogical and computational uses of deep learning in mathematics education. Bridging this gap requires interdisciplinary collaboration between educators and technology developers to ensure technological applications support meaningful learning. The findings provide a reference for aligning global research on deep learning with national education policy, ensuring relevance for curriculum design and classroom practice.

Keywords: Deep learning, Mathematics Education, Technology Developer, Artificial Intelligence

*** Corresponding Author Email**: valeria.gusti@ecampus.ut.ac.id

Submission	Revised	Accepted	Published
May 30, 2025	July 12, 2025	July 18, 2025	August 2, 2025

How to cite (in APA style):
Gusti, V. Y. K., & Amastini, F. (2025). Bridging pedagogy and AI: a systematic review of deep learning in mathematics education. *Jurnal Pendidikan Matematika (JPM)*, *11*(2), 141–156. https://doi.org/10.33474/jpm.v11i2.24319

INTRODUCTION

The rapid transformation of global education in the last decade has been shaped by two intersecting forces: the urgent call for competencies that prepare learners for uncertain futures, and the accelerating influence of advanced digital technologies in teaching and learning. Education systems are now expected to cultivate not only knowledge acquisition but also creativity, collaboration, adaptability, and critical thinking (Darling-Hammond et al., 2019; Fullan et al., 2018). These expectations have intensified interest in pedagogical frameworks that promote deep, transferable learning outcomes across disciplines.

In educational discourse, deep learning refers to the sustained development of higher-order thinking, conceptual mastery, and the application of knowledge to real-world challenges (Fullan et al., 2017). In Indonesia, this vision is formally embedded in the *Kurikulum Merdeka*, which positions deep learning as a pathway to foster student agency, meaningful engagement, and authentic problem-solving (Indonesian Education Standar, Curriculum, and Assesment Agency [BSKAP], 2024). This perspective emphasizes collaborative learning, creativity, and character formation as integral components of competency-based education.

However, in computer science and artificial intelligence (AI), deep learning carries a markedly different meaning, describing a subset of machine learning that leverages multi-layered neural networks for pattern recognition, prediction, and adaptive automation (Goodfellow et al., 2016). In mathematics education research, these computational approaches have been applied to student modeling, personalized learning environments, and intelligent tutoring systems (Alnasyan et al., 2025; Strock et al., 2025; Villegas-Ch et al., 2025). AI-driven deep learning can analyze learning patterns, detect misconceptions, and adapt content to individual needs, thereby offering tools that potentially complement pedagogical deep learning (Celik et al., 2022).

The coexistence of these two conceptualizations, pedagogical deep learning and computational deep learning, presents both opportunities and challenges. While they originate from different epistemological traditions, their convergence has the potential to enrich mathematics education by integrating robust cognitive development with data-informed instructional design (Alam & Mohanty, 2023; Celik et al., 2022). Yet, alignment between these interpretations is not guaranteed, and without a shared conceptual framework, there is a risk of fragmented practice and miscommunication between educators, policymakers, and technologists (Celik et al., 2022).

Given this landscape, a systematic review of the literature is necessary because existing research on deep learning in mathematics education remains fragmented between pedagogical and computational perspectives, with limited studies comparing or integrating the two. This study therefore examines recent peer-reviewed research to clarify how deep learning is understood and applied in mathematics education, identifying key interpretations, implementation strategies, benefits, and limitations. By mapping these overlaps and divergences, the study addresses this conceptual gap and contributes to a more coherent and actionable discourse that bridges educational policy goals with emerging technological practices.

METHOD

A comprehensive systematic literature review (SLR) was conducted to address the research questions. SLR is a structured method for collecting and synthesizing relevant research on a specific topic using pre-determined eligibility criteria (Page et al., 2021). This study examined peer-reviewed journal publications published between January 2015 and February 2025 to capture the most recent decade of educational research and contemporary relevance. This timeframe also coincides with key national policy initiatives emphasizing digital transformation and pedagogical innovation, making it a suitable window for mapping current trends and identifying emerging research directions. The review was guided by three research questions: (1) How is deep learning conceptualized in mathematics education research? (2) How is it applied across pedagogical and computational contexts? and (3) To what extent do these applications align with the Indonesian Ministry of Education and Culture's (Kemendikbud) definition of deep learning? The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework was employed to guide the review process, as it provides a standardized methodology with checklist-based guidelines that enhance quality assurance and replicability (Page et al., 2021).

The PRISMA process consists of four sequential stages identification, screening, eligibility, and inclusion. During the identification stage, a total of 1,881 records were retrieved from four academic databases: ProQuest (*n* = 745), Scopus (*n* = 532), ScienceDirect (*n* = 354), and SpringerLink (*n* = 250. Each step was applied systematically to ensure the rigor and transparency of the review. This approach was chosen to enable a robust synthesis of current scholarship and to map the evolving discourse deep learning in mathematics education, with specific attention to both pedagogical and computational interpretations. Figure 1 presents the PRISMA flow diagram as adapted for this study.

Table 1. Synonyms and Alternative Terms for Main Search Terms

Deep Learning	**Mathematics Education**
Deep neural networks	Mathematics education
Neural networks	Math education
Machine learning	Mathematics
Computational intelligence	STEM education
Artificial intelligence	Math
Learning analytics	Maths

Table 2. Inclusion and Exclusion Criteria

Inclusion criteria	**Exclusion criteria**
Published between 2015 and 2025	Published before 2015
Indexed, peer-reviewed journal articles	Non-indexed journals, book chapters, conference proceedings, theses, prefaces, opinion pieces
Written in English	Non-English publications
Explicit application of deep learning within mathematics or STEM education	Studies where "deep learning" is unrelated to education or mathematics
Empirical, conceptual, or review studies with a clear definition or application of deep learning	Studies lacking clear definitions or applications of deep learning in education

Systematic Review Process

Identification

The search was conducted in ScienceDirect, Scopus, Springer Link, and ProQuest. Two main search terms were developed from the core research focus: deep learning and mathematics education. Synonyms and alternative terms were compiled (Table 1) to capture a wide range of relevant studies. These terms were combined using Boolean operators in title, abstract, and keyword searches. Across the four databases, the initial search produced 1,881 records. All results were exported and imported into Zotero to manage the screening process and remove duplicates.

Screening

The PRISMA principles (Page et al., 2021) guided the screening process. Duplicate removal reduced the dataset from 1,881 to 1,675 unique articles. Titles and abstracts were independently screened by two reviewers to ensure consistency and minimize bias. Any discrepancies in inclusion decisions were discussed until consensus was reached. Studies were excluded if they did not focus on mathematics education or used the term "deep learning" solely in a computational sense without an educational link. No exclusions were made based on geographic region.

Eligibility

In the eligibility stage, full-text articles were examined for alignment with the study's objectives. This involved reviewing each study's title, abstract, methodology, results, and discussion to ensure that deep learning was explicitly addressed within a mathematics or STEM education context. Studies that did not clearly explain deep learning's role in educational practice were excluded.

Inclusion

Following the eligibility phase, 16 studies met all inclusion criteria and were retained for the final analysis. These studies formed the basis for identifying patterns in the interpretation and application of deep learning, as well as for comparing global perspectives with the Indonesian Ministry of Education and Culture's pedagogical framework.

RESULT

Selection Summary

An initial pool of records containing the term "deep learning" was identified across ScienceDirect, Scopus, Springer Link, and ProQuest, n = 1,881. After duplicate removal (n removed = 1,675), 206 unique records remained. Following title/abstract and full-text screening for relevance to mathematics education and explicit treatment of deep learning, 16 studies were retained for analysis.

Table 3. Summary of Included Studies (2015–2025) (Part 1)

No.	Authors & Year	Methodology/ Research Design	Educational Context	Focus of Deep Learning	Main Findings
1	(Fauskanger & Bjuland, 2018)	Theoretical review	Computer Science / AI	Computational	Established foundational architectures for deep neural networks; basis for DL models later used in education.
2	(Qu et al., 2021)	Scholarly monograph	AI & Education Research	Computational	Defined core deep learning methods and principles for supervised and unsupervised learning, forming conceptual foundation for educational DL applications.
3	(Demir & Karaboğa, 2021)	Analytical commentary	Machine Learning/ Ethics	Computational	Critiqued opacity of DL ("black-box") and advocated interpretable models; key relevance for educational transparency.

Table 4. (Continued) Summary of Included Studies (2015–2025) (Part 2)

No.	Authors & Year	Methodology/ Research Design	Educational Context	Focus of Deep Learning	Main Findings
4	(Zhang et al., 2021)	Systematic Review	Higher Education	Mixed	Found AI, including DL, used for feedback, prediction, and personalization; emphasized educator involvement gaps.
5	(Tian et al., 2022)	Systematic Review	Higher Education	Computational	Examined learning analytics leveraging DL to improve study success; showed promise in predictive modeling and adaptive learning.
6	(Liu, 2022)	Empirical study	University Mathematics	Pedagogical	Integrated ecological and mathematical learning within DL pedagogy; emphasized conceptual and reflective engagement.
7	(Tsiakmaki et al., 2020)	Systematic Review (2019–2023)	Open / Online Learning	Mixed	Identified DL's growing role in adaptive assessment and personalization in open learning environments.
8	(Jin & Suh, 2024)	Experimental design	Secondary Mathematics	Computational	Applied DL algorithms in blended teaching; improved student outcomes in ethnic minority regions.
9	(Nurdiana et al. 2024)	Quasi-experimental	Higher Education	Pedagogical	Demonstrated that flipped classrooms improve motivation and deep learning processes among college students.
10	(Strock et al., 2025)	Quantitative experimental	Child Mathematics Learning	Computational	Used DL neural networks to model learning disabilities; provided insights into cognitive mechanisms in math learning.
11	(Villegas-Ch et al., 2025)	Mixed-method evaluation	STEM Education	Computational	Developed adaptive DL-based tutoring systems; improved personalized feedback and engagement.

Table 5. (Continued) Summary of Included Studies (2015–2025) (Part 3)

No.	Authors & Year	Methodology/ Research Design	Educational Context	Focus of Deep Learning	Main Findings
12	(Alnasyan et al., 2025)	Computational modeling	Online Learning Environments	Computational	Introduced Kanformer model (Transformer-based DL) for predicting student performance; showed strong predictive accuracy.
13	(Pujawati et al., 2025)	Design-based study	Collaborative Science/Math Inquiry	Computational	Developed automated feedback scaffolds using DL and NLP; improved student idea convergence.
14	(Slamet et al., 2025)	Qualitative case study	Secondary Mathematics	Pedagogical	Reported integration of ICT and DL-based pedagogy; enhanced engagement and conceptual depth.
15	(Qiu & Ishak, 2025)	Conceptual / Theoretical	Transformative Learning	Pedagogical	Reinterpreted "deep learning" as transformation and meaning-making; linked to lifelong learning processes.
16	(Putri et al., 2025)	Mixed-method review	Educational Technology	Mixed	Explored convergence of AI (including DL), mobility, and pedagogy; emphasized hybrid human–machine learning environments.

Framing and Application of Deep Learning in Mathematics Education

Figure 1 presents the categorization of how deep learning (DL) was framed in the 16 included studies. Four studies interpreted DL solely as a computational approach, two studies focused entirely on pedagogical interpretations, five studies combined both computational and pedagogical perspectives, and five studies did not clearly define the term in their context.

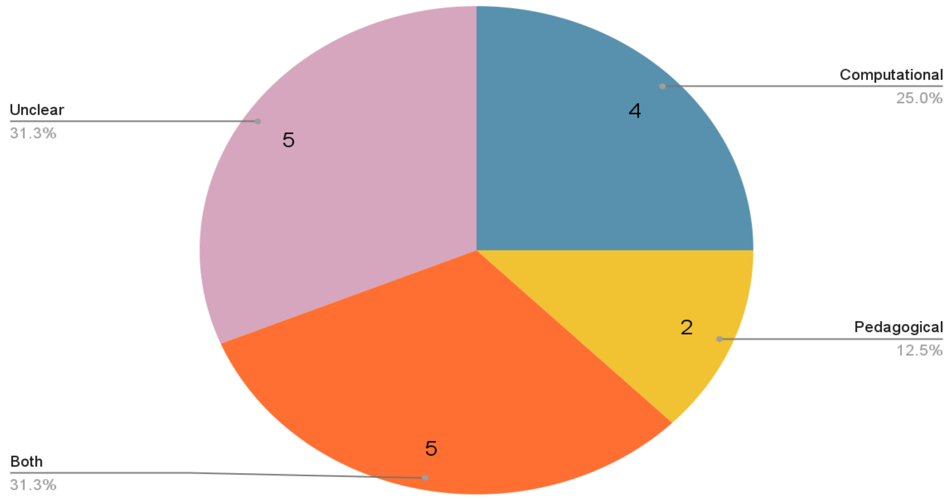

Figure 1. Framing of deep learning in included studies (n = 16)

Table 6. PRISMA Flow of Study Selection Process

Stage	Description	n
Identification	Records identified through database searching (ScienceDirect, Scopus, Springer Link, ProQuest)	1,881
	Duplicates removed	206
Screening	Records after duplicates removed	1,675
	Records excluded after title/abstract screening	1,659
Eligibility	Full-text articles assessed for eligibility	16
	Full-text articles excluded (with reasons)	0
Included	Studies included in qualitative synthesis	16

The combination of computational and pedagogical perspectives appeared in approximately one-third of the studies. This is significant in the context of the Indonesian Ministry of Education and Culture's definition, which situates deep learning within a pedagogical framework emphasizing higher-order thinking, collaboration, creativity, and real-world application (BSKAP, 2024). Similar distinctions between policy-oriented pedagogical definitions and technology-oriented research definitions have also been highlighted in the broader literature on deep learning and its applications in education (Darling-Hammond et al., 2019; Fullan et al., 2018).

Table 5 shows the application areas identified across the included studies. The most frequent application was intelligent tutoring, automated feedback, and assessment systems (n = 9), followed by visualization tools (n = 5). Prediction and analytics applications, while prominent in the AI field, appeared infrequently in mathematics education contexts (Celik et al., 2022).

Table 7. Applications of Deep Learning in Mathematics/STEM Education (Non-Exclusive Counts)

Application type	Count
Intelligent tutoring / automated feedback / assessment	9
Visualization tools	5
Other	5
Prediction / analytics	1
Automated assessment with student reflection	1
Blended learning	1
Problem-based or design-based learning	1

The predominance of intelligent tutoring and automated feedback systems indicates a global tren towards leveraging deep learning technologies for instructional automation and personalization (Villegas-Ch et al., 2025). However, relatively few studies explored deep learning applications that directly align with the Indonesian Ministry of Education and Culture’s emphasis on reflective, collaborative, and critical thinking activities (BSKAP, 2024).

Reported Benefits

Figure 2 summarizes the reported benefits across the 16 studies. Most studies (n = 10) highlighted improved student performance or engagement, while only five explicitly connected DL applications to pedagogical deep-learning outcomes such as conceptual understanding or collaboration. The emphasis on performance and engagement outcomes reflects a widespread evaluation practice in AI-enabled educational systems, which often privileges quantitative achievement metrics over qualitative learning processes (Celik et al., 2022).

This suggests a partial misalignment with the Indonesian Ministry of Education and Culture’s framework, which prioritizes learning processes, collaboration, and problem-solving over mere test scores (BSKAP, 2024).

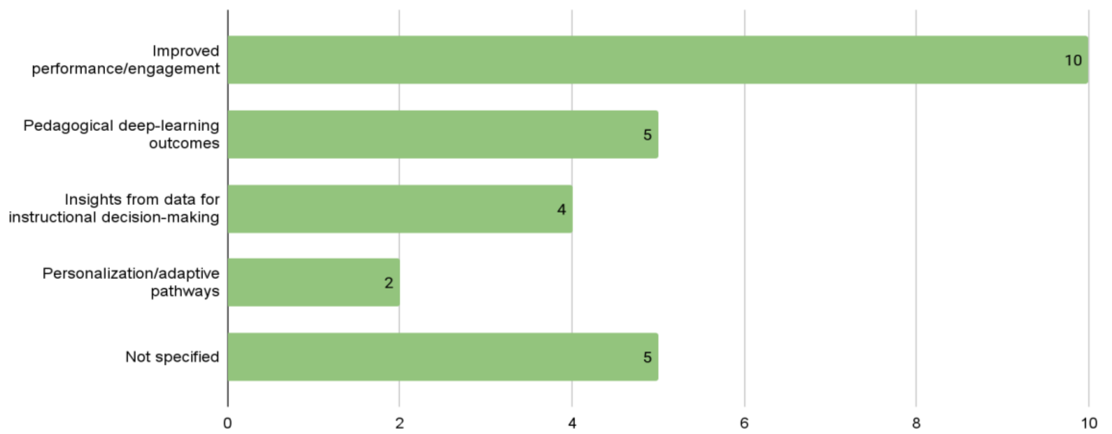

Figure 2. Reported Benefits of Deep Learning in Mathematics/STEM Education (Non-Exclusive Counts)

The emphasis on performance and engagement outcomes reflects a widespread evaluation practice in AI-enabled educational systems, which often privileges quantitative achievement metrics over qualitative learning processes (Celik et al., 2022). This suggests a partial misalignment with the Indonesian Ministry of Education and Culture's framework, which prioritizes learning processes, collaboration, and problem-solving over mere test scores (BSKAP, 2024).

Challenges and Limitations

Figure 3 presents the challenges reported in the included studies. The most frequent were teacher training needs (n = 7) and interpretability/trust concerns related to the "black-box" nature of DL systems (n = 4). Without sufficient professional development, teachers may struggle to integrate DL-based tools effectively (Celik et al., 2022; Villegas-Ch et al., 2025). Moreover, the opacity of algorithmic decision-making limits educators' trust, particularly in contexts where curricular alignment and accountability are critical (Celik et al., 2022).

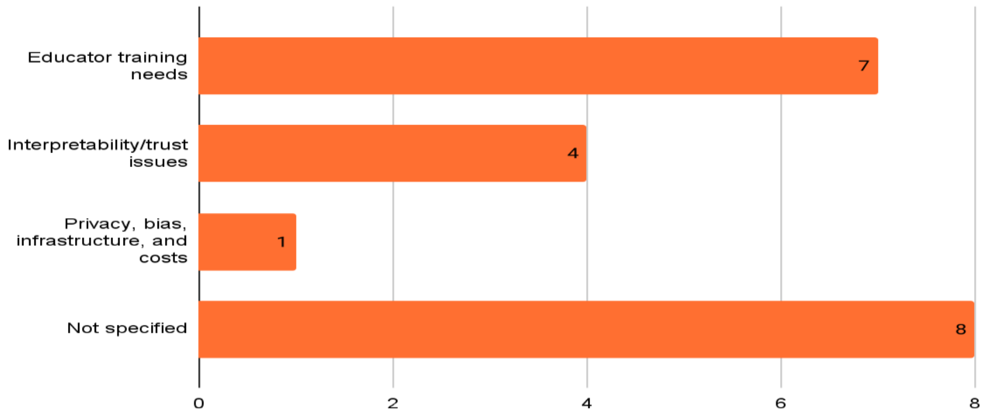

Figure 3. Reported Challenges in Implementing Deep Learning (Non-Exclusive Counts)
Alignment with Kemendikbud's Pedagogical Framing

Out of the 16 studies reviewed, seven demonstrated partial or full alignment with the Indonesian Ministry of Education and Culture's (Kemendikbud) definition of deep learning as outlined in the Kurikulum Merdeka (BSKAP, 2024). In these studies, deep learning was framed in ways that reflected higher-order thinking, creativity, collaboration, reflective practice, and the application of knowledge to real-world contexts. The remaining nine studies, while employing the term "deep learning," tended to focus on computational models or automated instructional processes with limited integration of these pedagogical principles. This distribution suggests that although there is growing recognition of deep learning in mathematics education, a substantial portion of the literature prioritizes technological efficiency over holistic, student-centered learning outcomes envisioned in national education policy.

The limited conceptual alignment underscores the need for greater interdisciplinary integration between computational approaches and pedagogical design principles grounded in national policy priorities (BSKAP, 2024; Fullan et al., 2018).

Table 6. Alignment with Kemendikbud's pedagogical framing

Alignment Category	Count
Aligned (pedagogical or both)	7
Not aligned	9

DISCUSSION

Conceptual Interpretations of Deep Learning

The term deep learning has undergone significant conceptual diversification across disciplinary contexts, resulting in distinct yet occasionally overlapping interpretations in mathematics education research. Within the corpus of 16 studies examined in this review, two primary framings were identified: (1) a pedagogical framing grounded in educational theory, and (2) a computational framing situated within the domain of computer science and artificial intelligence.

The pedagogical framing, evident in 11 of the reviewed studies, conceptualizes deep learning as a process of fostering sustained cognitive engagement, conceptual understanding, and the ability to apply knowledge in novel and authentic contexts. In this interpretation, emphasis is placed on higher-order cognitive processes such as analysis, synthesis, and evaluation, alongside affective and social dimensions including collaboration, creativity, and reflective thinking (BSKAP, 2024; Darling-Hammond et al., 2019). Pedagogical approaches associated with this framing frequently involve problem-based learning, inquiry-oriented instruction, and interdisciplinary integration, with mathematics serving as both a content area and a medium for developing broader competencies. This conceptualization demonstrates strong alignment with the Indonesian Ministry of Education and Culture's (Kurikulum Merdeka) definition of deep learning, which underscores higher-order thinking, collaborative problem-solving, and the integration of character-building within learning processes (BSKAP, 2024).

By contrast, the computational framing, identified in five studies, positions deep learning as a subset of machine learning techniques characterized by multi-layered neural network architectures capable of hierarchical representation learning (Goodfellow et al., 2016). Within this interpretation, the emphasis lies on algorithmic optimization, predictive modeling, and data-driven instructional adaptation. Applications in mathematics education include intelligent tutoring systems, automated assessment tools, and predictive analytics for identifying at-risk learners (Mohamed et al., 2022; Villegas-Ch et al., 2025).

However, unlike pedagogically oriented models that conceptualize deep learning as sustained cognitive engagement, computational implementations often prioritize accuracy, performance metrics, and scalability over epistemic or reflective dimensions of learning. This distinction echoes concerns raised in earlier analyses that the integration of AI into education tends to reproduce traditional assessment logics rather than transform learning processes (Celik et al., 2022). While these systems can improve personalization and efficiency, they risk instrumentalizing learning by emphasizing prediction rather than meaning-making. Compared with studies such as Pujawati et al. (2025) which interpret deep learning as a process of conceptual depth and reflective inquiry, computationally framed research treats "depth" as a technical property of network layers. This conceptual divergence reinforces the importance of distinguishing between *deep learning as an algorithmic mechanism* and *deep learning as an educational outcome*, a nuance often underexplored in the literature.

The coexistence of these two framings presents a conceptual divergence with potential implications for both research and policy. In pedagogically oriented discourse, deep learning denotes cognitive and metacognitive engagement directed towards durable, transferable knowledge. In contrast, in computational discourse, the term denotes algorithmic complexity and model performance. This divergence increases the risk of conceptual conflation, particularly when interdisciplinary collaboration occurs without explicit definitional alignment. Such misalignment may lead to discrepancies between policy aspirations, such as Kemendikbud's focus on holistic, student-centered deep learning, and technological implementations that prioritize predictive accuracy over pedagogical depth (BSKAP, 2024; Mohamed et al., 2022).

Clarifying the intended meaning of deep learning within mathematics education research is therefore essential. A shared conceptual framework could enhance coherence between pedagogical objectives and technological innovation, ensuring that the integration of computational approaches genuinely supports the cultivation of meaningful, future-relevant learning experiences.

Pedagogical Applications in Mathematics Education

Within the reviewed literature, 11 of the 16 studies demonstrated a pedagogical orientation toward deep learning in mathematics education. In these studies, deep learning was operationalized as an instructional goal aimed at enabling students to develop robust conceptual understanding, transferable problem-solving skills, and the capacity to engage in reflective and collaborative inquiry. This approach aligns with constructivist and socio-constructivist theories of learning, which emphasize active student engagement, social interaction, and the integration of prior knowledge into new contexts (Fullan, M., Quinn & McEachen, 2017).

Common pedagogical applications identified include the design of problem-based learning (PBL) environments, where mathematics problems are situated within authentic, real-world scenarios. Such contexts encourage students to integrate multiple concepts, negotiate meaning collaboratively, and apply mathematical reasoning to complex challenges. For example, several studies reported the use of interdisciplinary STEM projects in which mathematics learning was embedded within engineering design tasks or scientific investigations, thereby fostering the application of mathematical ideas beyond the classroom setting (Goos et al. 2023).

Another frequently reported strategy was the implementation of collaborative learning structures, including small-group problem-solving sessions and peer teaching activities. These approaches were shown to promote not only deeper conceptual engagement but also communication skills and the ability to critique and refine mathematical reasoning in a group setting. In line with Kurikulum Merdeka's deep learning framework, such strategies explicitly target higher-order cognitive processes, creativity, and character development, indicating strong conceptual alignment with national education priorities (BSKAP, 2024).

Some studies also incorporated reflective practices as an integral component of the learning process, requiring students to articulate their problem-solving strategies, evaluate alternative approaches, and connect mathematical concepts to broader societal or technological issues. Reflection was framed as a metacognitive tool for consolidating learning, thereby enhancing transferability and long-term retention (Fullan et al., 2017).

Overall, the pedagogical applications identified in the reviewed literature illustrate how deep learning, when framed in educational terms, extends beyond procedural fluency and memorization. Instead, it positions mathematics as both a tool and a context for cultivating competencies essential for navigating complex, real-world problems. These findings reinforce the compatibility between international research perspectives and

Kurikulum Merdeka's emphasis on higher-order thinking, collaboration, and meaningful learning outcomes (BSKAP, 2024; Darling-Hammond et al., 2019).

Technological Approaches and Their Educational Potential

In contrast to pedagogical interpretations of deep learning, five of the sixteen reviewed studies adopted a primarily technological orientation, framing deep learning as a branch of machine learning, most commonly operationalized through deep neural networks (DNNs) and related architectures (Goodfellow et al., 2016). In these studies, deep learning was applied as a computational tool for enhancing educational processes in mathematics through automation, prediction, and adaptive personalization (Villegas-Ch et al., 2025).

One of the most common applications identified was automated performance prediction, where deep learning algorithms analyzed historical learner data to forecast future performance trajectories (Alnasyan et al., 2025). Such predictive models have been employed to identify at-risk students early, enabling targeted interventions. These systems rely on large datasets, often collected through online learning platforms or learning management systems, to detect patterns and trends that may not be immediately visible to educators.

Another frequently reported application was intelligent tutoring systems (ITS), where deep learning models dynamically adapted instructional pathways based on the learner's real-time performance (Villegas-Ch et al., 2025). This approach allowed for highly personalized learning experiences, adjusting difficulty levels, problem types, and feedback provision to match individual learner profiles. In mathematics education, such systems have been used to provide step-by-step guidance, hint generation, and error-specific feedback, thereby supplementing human instruction.

Several studies also employed automated feedback systems capable of assessing complex student responses, such as open-ended problem-solving tasks (Villegas-Ch et al., 2025). By leveraging natural language processing and pattern recognition capabilities, these models aimed to provide immediate feedback, maintain assessment consistency, and reduce the grading workload for educators.

Although these technological approaches offer considerable potential for scaling personalized learning and improving efficiency, the reviewed literature also highlighted certain limitations. Notably, the "black-box" nature of deep learning algorithms was cited as a challenge, as educators often lacked visibility into how the system generated predictions or recommendations (Celik et al., 2022). This opacity can undermine trust and hinder adoption in classroom settings. Additionally, infrastructure constraints, such as limited access to high-performance computing resources and reliable internet connectivity, were reported as barriers in lower-resource educational contexts.

Despite these challenges, the technological applications reviewed present significant opportunities for advancing mathematics education, particularly when integrated with pedagogically grounded strategies. By aligning computational innovations with established learning theories and national priorities, such as Kurikulum Merdeka's emphasis on higher-order thinking and meaningful engagement (BSKAP, 2024), these tools can serve as catalysts for deeper, more personalized, and contextually relevant learning experiences.

Integration Challenges and Alignment with Policy

While the reviewed studies demonstrate the potential of deep learning in mathematics education, several integration challenges were consistently identified. One prominent barrier concerns conceptual misalignment between computational interpretations of deep learning, as found in much of the machine learning literature, and the pedagogical framing emphasised in education policy documents, such as those issued by the Indonesian Ministry of Education, Culture, Research, and Technology (BSKAP, 2024). In the computational sense, deep learning refers primarily to multi-layered neural network architectures capable of feature extraction and pattern recognition from large datasets (Goodfellow et al., 2016). In contrast, the Kurikulum Merdeka framework conceptualises deep learning as a pedagogical construct encompassing higher-order thinking, collaboration, creativity, and problem-solving in authentic contexts.

From the sixteen studies analysed, only eleven demonstrated clear alignment with Kurikulum Merdeka's interpretation, primarily through an emphasis on fostering conceptual understanding, collaborative engagement, and real-world problem application in mathematics classrooms. The remaining five studies adopted a narrowly technical approach, focusing exclusively on algorithmic optimisation or automated analytics without explicit reference to student-centred learning principles. This divergence underscores the terminological ambiguity that arises when the same term is used differently across disciplines (Mohamed et al., 2022).

A second challenge relates to educator readiness and capacity. Studies reported that teachers often lacked the technical knowledge to implement deep learning-driven systems effectively (Celik et al., 2022). Even when technological solutions were available, insufficient professional development limited teachers' ability to interpret algorithmic outputs or adapt instructional strategies accordingly. This issue directly impacts alignment with curriculum reforms, which require teachers to integrate advanced learning technologies in ways that support curriculum goals (Development, 2021).

Infrastructure disparities further complicate integration, particularly in under-resourced schools where computing facilities and stable internet connectivity remain limited (Villegas-Ch et al., 2025). In such settings, the deployment of computationally intensive deep learning models may be impractical without targeted investment and policy support.

Finally, the literature points to a persistent lack of interdisciplinary collaboration between education policymakers, curriculum developers, and artificial intelligence researchers. Without cross-sector dialogue, there is a risk that AI-driven deep learning solutions will be developed without sufficient consideration of pedagogical objectives or local curricular frameworks (Celik et al., 2022). To address these challenges, coordinated alignment strategies are essential. For teachers, professional development programs should prioritize digital literacy and pedagogical integration of AI tools, enabling educators to interpret algorithmic outputs, adapt instruction responsively, and maintain pedagogical control. For policymakers, efforts must focus on establishing shared definitions of deep learning that bridge pedagogical and computational perspectives, ensuring policy coherence with classroom realities. This includes developing national standards and evaluation frameworks that balance innovation with equity and ethical safeguards. For technology developers, collaboration with educators during the design process is critical to ensure that learning analytics, adaptive systems, and intelligent tutoring models align with curricular goals and support higher-order cognitive engagement rather than solely data-driven optimization. By combining these efforts namely teacher empowerment, policy coherence, and design collaboration, the adoption of deep learning in mathematics education can evolve into a synergistic model that is technologically robust, pedagogically grounded, and contextually responsive to national education priorities.

CONCLUSION AND SUGGESTIONS

This systematic literature review contributes to the scholarly understanding of how *deep learning* is conceptualised and operationalised within mathematics education, synthesising sixteen peer-reviewed studies published between 2017 and 2025. The findings reveal two dominant interpretations: one computational, focusing on artificial intelligence and neural network applications, and one pedagogical, grounded in higher-order thinking, collaboration, creativity, and real-world problem-solving. The review advances theoretical clarity by juxtaposing these paradigms and examining their partial alignment with the Indonesian Ministry of Education and Culture's (Kemendikbud) conceptualisation of deep learning as an educational process. By identifying conceptual overlaps and divergences, this study provides an integrative framework that connects technological innovation with pedagogical depth, bridging a longstanding gap in interdisciplinary educational research.

Building on these insights, the study offers three key recommendations. First, for practice, teachers should receive sustained professional development to integrate AI-enabled tools meaningfully while maintaining focus on conceptual understanding and student agency. Second, for policy, governments and curriculum authorities such as Kemendikbud should establish unified definitions and implementation guidelines for deep learning that align technological adoption with pedagogical intent. Third, for research, future studies should employ mixed-method and longitudinal designs to explore how AI-driven systems can support reflective, collaborative, and context-sensitive learning in mathematics classrooms. These directions collectively aim to ensure that deep learning innovations remain pedagogically grounded, equitable, and aligned with national and global educational goals.

ACKNOWLEDGMENT

The author(s) would like to express sincere gratitude to the researchers and educators whose work informed this review. Appreciation is also extended to the Indonesian Ministry of Education and Culture (Kemendikbud) for providing publicly accessible policy documents that supported the policy alignment analysis in this study. Special thanks are given to colleagues for their constructive feedback, which significantly improved the clarity and depth of this manuscript.

REFERENCES

Alam, A., & Mohanty, A. (2023). Educational technology: Exploring the convergence of technology and pedagogy through mobility, interactivity, AI, and learning tools. Cogent Engineering, 10(2), 2283282. https://doi.org/10.1080/23311916.2023.2283282

Alnasyan, B., Basheri, M., Alassafi, M., & Alnasyan, K. (2025). Kanformer: An attention-enhanced deep learning model for predicting student performance in virtual learning environments. Social Network Analysis and Mining, 15, 25. https://doi.org/10.1007/s13278-025-01446-7

Badan Standar, Kurikulum, dan Asesmen Pendidikan [Indonesian Education Standar, Curriculum, and Assesment Agency] (2024). Kajian akademik Kurikulum Merdeka.

Celik, I., Dindar, M., Muukkonen, H., & Järvelä, S. (2022). The promises and challenges of artificial intelligence for teachers: A systematic review of research. TechTrends, 66, 616–630. https://doi.org/10.1007/s11528-022-00715-y

Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2019). Implications for educational practice of the science of learning and development. *Applied Developmental Science*, *24*(2), 97–140. https://doi.org/10.1080/10888691.2018.1537791

Demir, I., & Karaboğa, H. A. (2021). Modeling mathematics achievement with deep learning methods. Sigma Journal of Engineering and Natural Sciences, 39, 33–40. https://doi.org/10.14744/sigma.2021.00039

Organisation for Economic Co-operation and Development. (2021). 21st-century readers: Developing literacy skills in a digital world. OECD Publishing. https://doi.org/10.1787/a83d84cb-en

Fauskanger, J., & Bjuland, R. (2018). Deep learning as constructed in mathematics teachers' written discourses. International Electronic Journal of Mathematics Education, 13(3), 149–160. https://doi.org/10.12973/iejme/2705

Fullan, M., Quinn, J., & McEachen, J. (2017). Deep learning: Engage the world, change the world. Corwin.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.

Goos, M., Carreira, S., & Namukasa, I. K. (2023). Mathematics and interdisciplinary STEM education: Recent developments and future directions. ZDM–Mathematics Education, 55, 1199–1217. https://doi.org/10.1007/s11858-023-01533-z

Jin, H., & Suh, B. (2024). Research on a statistics education program utilizing deep learning predictions in high school mathematics. The Mathematical Education, 63(2), 209–231. https://doi.org/10.7468/mathedu.2024.63.2.209

Liu, R. (2022). Optimization model of mathematics instructional mode based on deep learning algorithm. Computational Intelligence and Neuroscience, 2022, 1817990. https://doi.org/10.1155/2022/1817990

Mohamed, M. Z. B., Hidayat, R., Suhaizi, N. N. B., Sabri, N. B. M., Mahmud, M. K. H. B., & Baharuddin, S. N. B. (2022). Artificial intelligence in mathematics education: A systematic literature review. International Electronic Journal of Mathematics Education, 17(3), em0694. https://doi.org/10.29333/iejme/12132

Nurdiana, A., Zulianti, H., Ciciria, D., Fitria, N., & Kirana, A. R. (2024). Practical applications of deep learning in mathematics to enhance student engagement and conceptual mastery. JTAM (Jurnal Teori dan Aplikasi Matematika, 8(4), 1361–1373. https://doi.org/10.31764/jtam.v8i4.33102

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. G., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, n71. https://doi.org/10.1136/bmj.n71

Pujawati, F., Azkia, M. N., & Susilawati, W. (2025). Exploration of the implementation of deep learning approach in teaching mathematics in secondary schools. Unnes Journal of Mathematics Education, 14(2), 98–105. https://doi.org/10.15294/ujme.v14i2.27374

Putri, D. K., Zahro, N., Widodo, W., & Suprapto, N. (2025). Primary teachers' perceptions of deep learning pedagogy in culture-integrated STEM education: A quantitative survey. Journal on Mathematics Education, 16(4), 1331–1346. https://doi.org/10.22342/jme.v16i4.pp1331-1346

Qiu, Y., & Ishak, N. A. (2025). AI-assisting technology and social support in enhancing deep learning and self-efficacy among primary school students in mathematics in China. International Journal of Learning, Teaching and Educational Research, 24(2), 21–37. https://doi.org/10.26803/ijlter.24.2.2

Slamet, D., Hendriana, B., & Supiat. (2025). Mengungkap perspektif siswa: Peran deep learning dalam visualisasi konsep dan pemecahan masalah matematika. International Journal of Progressive Mathematics Education, 5(1), 225–237. https://doi.org/10.22236/ijopme.v5i1.19310

Strock, A., Mistry, P. K., & Menon, V. (2025). Personalized deep neural networks reveal mechanisms of math learning disabilities in children. Science Advances, 11(23), eadq9990. https://doi.org/10.1126/sciadv.adq9990

Tian, X., Zhao, J., & Nguyen, K. T. (2022). Practical research on primary mathematics teaching based on deep learning. *Science Programming*, *2022*, 7899180.

https://doi.org/10.1155/2022/7899180
Tsiakmaki, M., Kostopoulos, G., Kotsiantis, S., & Ragos, O. (2020). Transfer learning from deep neural networks for predicting student performance. Applied Sciences, 10(6), 2145. https://doi.org/10.3390/app10062145
Villegas-Ch, W., Buenano-Fernandez, D., Navarro, A. M., & Mera-Navarrete, A. (2025). Adaptive intelligent tutoring systems for STEM education: Analysis of the learning impact and effectiveness of personalized feedback. Smart Learning Environments, 12, Article 1. https://doi.org/10.1186/s40561-025-00389-y
Zhang, Y., Liu, S., & Shang, X. (2021). An MRI study on effects of math education on brain development using multi-instance contrastive learning. Frontiers in Psychology, 12, 765754. https://doi.org/10.3389/fpsyg.2021.765754