LLMs in Education

Large Language Models (LLMs) have become one of the fundamental elements of the recent transformation in educational technologies. Leveraging their Transformer-based architectures, these models demonstrate high accuracy in natural language processing, effectively interpret contextual relationships, and generate explanations with clear pedagogical relevance. Unlike traditional knowledge-transfer-based teaching methods, LLM-based systems can instantly adapt to the student's learning pace, needs, and potential misunderstandings. These adaptive features make LLMs stand out as both a dynamic support tool in learning processes and a scalable solution in instructional design.The inclusion of LLMs in educational environments has fundamentally transformed the forms of interaction between students and teaching staff. The capacity to provide pedagogically framed explanations to students' questions and identify conceptual errors transforms these models from simple "response generators" into active teaching components. Various studies in the literature report that the personalized feedback processes offered by LLMs significantly strengthen students' motivation to learn and their self-regulation skills [1], [2].

From the perspective of teachers, LLMs are increasingly being used to alleviate the heavy workload in educational processes. Time-consuming tasks such as preparing course content, simplifying texts, generating exam questions, and conducting rubric-based assessments can be carried out more quickly and consistently thanks to LLMs. Institutional assessment platforms can use these models to identify conceptual errors in student writing and provide structured feedback [3]. These automation processes allow teachers to devote their time to higher-level pedagogical planning.However, the literature also highlights some potential drawbacks of using LLMs. Models sometimes produce context-disconnected or erroneous explanations, which can lead to students acquiring incorrect information. Furthermore, it is stated that excessive use may weaken students' independent problem-solving skills and lead to superficial learning. Indeed, as shown in a study published in the journal Applied Sciences [4], high dependence on LLMs is negatively correlated with students' final performance. Therefore, it is crucial to use LLMs in a balanced and conscious manner from a pedagogical perspective.The integration of LLM-based solutions into education raises many ethical considerations. Training these models on large and heterogeneous datasets naturally increases risks such as privacy concerns, bias generation, and misinformation creation. Clearly defining and institutionally monitoring the processes of collecting, processing, and storing student information are fundamental requirements for ensuring the safe use of these technologies [2].

Findings in the literature reveal that LLM-based systems offer an effective structure that significantly supports teaching processes. Facilitating students' conceptual learning, reducing teachers' operational burden, and providing personalized feedback make these technologies a valuable element for contemporary education models. However, for LLMs to be used efficiently, it is critically important to blend technological capabilities with pedagogical principles and maintain a human-centered approach in the learning process.

1) Example Projects and Applications in General Education

LLM-based applications have evolved into a broad ecosystem that offers flexible, scalable, and personalized solutions tailored to different learner profiles in general education. These systems not only transfer knowledge but also provide integrated learning environments that guide, structure, and adapt to the learning process according to individual needs. Along with the digitalization process, the primary goal of LLM-based platforms is to reduce teacher workload, increase learning speed, and create sustainable teaching models suitable for different learning levels. In this context, this section details how institutional, commercial, and academic systems are positioned in educational processes.

1.1) Khanmigo (Khan Academy)

Khan Academy's GPT-4-based Khanmigo tool is one of the prominent examples in general education. The system provides an interactive learning process that encourages step-by-step thinking rather than simply providing the solution to the student. When a student takes the wrong approach, the model guides them to the correct conceptual structure with leading questions and hints. This approach not only supports independent reasoning but also reduces cognitive load, thereby contributing to more meaningful and sustained learning [2][5].

1.2) Duolingo Max

Duolingo Max represents one of the most successful applications of LLMs in language learning. The platform's AI-powered dialogue system enables learners to interact in real time using natural language. When a student uses an incorrect expression, the model not only provides corrections but also explains the grammatical and semantic aspects of the error in detail, facilitating a deeper learning process. Such explanatory feedback strengthens lasting comprehension in language learning, while the role-playing scenarios offered by the system allow users to experience real communication situations in a safe environment and support the development of conversational fluency [6].

1.3) Google LearnLM

Google LearnLM is designed as an extensive system consisting of large language models customized for educational purposes. The model can generate tailored explanations based on the student's age level, learning objectives, and subject area. Particularly effective in STEM subjects, LearnLM simplifies complex content to an appropriate cognitive level, identifies potential misunderstandings, and suggests learning paths to make the student's progress more efficient. Unlike traditional online content, it demonstrates a more advanced cognitive adaptation capability [7].

1.4) OpenStax Tutor

OpenStax Tutor provides a framework that analyzes data from completed activities and identifies individual learning gaps. The system identifies areas where students struggle and creates personalized learning plans. This feature offers a significant advantage, especially in crowded classrooms where teaching tailored to individual needs is not possible. In addition, real-time assessment outputs provide teachers with a strong data foundation for monitoring the overall learning status of the class [8].

1.5) Automated Scoring Systems (ETS & Pearson)

Another category frequently used in education is automated assessment and feedback systems. Developed by institutions such as ETS and Pearson, these tools analyze student texts according to rubrics and can make assessments based on criteria such as content structure, logical coherence, and conceptual accuracy. These models reduce teachers' workload by increasing the speed and consistency of assessment, particularly in writing-focused courses and large-scale exams [3]. In addition, the model provides structured feedback on areas for improvement rather than just scoring [9].

1.6) PQG Systems (AST + Concept Graphs)

Programming Question Generation (PQG) systems used in programming education leverage Large Language Models (LLMs) to generate coding questions appropriate for the student's level. In such systems, models produce curriculum-aligned content using technical components such as Abstract Syntax Trees (AST) and Local Knowledge Graphs (LKG) [10]. This significantly reduces the time teachers spend preparing new questions while also enabling the generation of more exercises targeting topics that students find challenging.

1.7) Multi-agent LLM Systems

Finally, multi-agent LLM systems are more advanced structures based on task sharing among several artificial intelligence agents. In these models, for example, one agent reviews the student's solution and identifies errors, while another agent generates explanatory feedback; yet another updates the learning plan based on the student's progress. Research shows that these multi-agent structures can provide more consistent and reliable educational outputs compared to the responses generated by a single model [11].

The examples at hand demonstrate that LLM-based technologies offer a wide range of functionality, flexibility, and adaptability in educational settings. These applications create scalability in teaching processes while providing students with a more personalized learning experience; they reduce teachers' workload while enhancing pedagogical quality. However, the long-term sustainable evaluation of this technological potential depends on a purpose-driven and controlled integration process that is aligned with pedagogical design principles.

2) Effectiveness or Impact of LLMs in Education

Research examining the impact of LLM-based technologies in education focuses on multi-layered areas such as learning outcomes, cognitive processes, motivation, social interaction, assessment accuracy, and contributions to instructional design. Studies conducted in this vein comprehensively reveal both the advantages and limitations of LLM-supported learning experiences.

2.1) Learning Performance Impact

Findings related to learning performance indicate that LLM-supported explanations lead to meaningful development, particularly in STEM and programming fields. LLMs can consistently deliver the step-by-step guidance students require when working through complex concepts. Experimental analyses of programming education show that students can distinguish error types more clearly, interpret the logical structure of code more quickly, and experience significant increases in learning speed. Findings from the Intelligent Deep Learning Tutoring System show that LLM-style explanatory feedback enhances students' problem-solving processes and substantially supports conceptual development [4].

2.2) Affective (Motivational & Emotional) Effects

The affective dimension is one of the prominent areas of LLM-supported learning studies. Reduced anxiety among students, a more relaxed approach to the trial-and-error process, and a willingness to make mistakes are among the strengths of LLM-based feedback mechanisms. A study has shown that LLM-supported paired programming increases both intrinsic motivation and the perception of benefits in the learning process among students [12]. The non-judgmental nature of LLMs helps students take more risks and view mistakes as a natural part of learning.

2.3) Social Interaction & Collaboration Effects

When evaluated in terms of social interaction, LLM-supported collaborative environments have been reported to enhance students' awareness of collaboration. Although the sense of social presence in human-human interaction cannot be fully achieved, some studies show that pair programming with LLM can match human-paired programming in performance outputs and contribute to more systematic error correction [12]. Students can ask more questions and think longer about conceptual issues when working with LLM.

2.4) Assessment and Evaluation Impact

Another noteworthy area is the impact of LLM-based systems on assessment and evaluation processes. Assessment models developed by ETS and Pearson can provide highly consistent feedback in terms of content integrity, structure, consistency, and rubric alignment [3]. In large-scale examinations, such systems substantially reduce grading time, lessen the burden on instructors for text-analysis-heavy tasks, and help decrease variability across human evaluators.

2.5) Negative Effects & Over-reliance Issues

Applied research also shows that LLM use does not always produce positive results. In a study published in Applied Sciences, Roldán-Álvarez and Mesa (2024) found that over-reliance on LLMs is negatively correlated with student achievement [4]. Based on Spearman correlation results, students who relied heavily on the model demonstrated reduced conceptual depth and a tendency toward more superficial learning. This highlights the need for balanced use in LLM-based learning processes.

2.6) Multi-Agent LLM Systems Impact

Recently developed multi-agent LLM systems offer a new approach aimed at improving consistency in learning processes and feedback quality. Having multiple LLM agents independently evaluate the same student solution and generate feedback by taking on different roles reduces error rates and strengthens pedagogical accuracy [11]. Various roles, such as expert, critical, and explanatory agents, support different stages of learning, making the system more robust.

2.7) Limitations, Risks, and Ethical Concerns

There are also studies pointing to limitations in the use of LLM. These models may occasionally misinterpret context and generate inaccurate or fabricated information (hallucinations), potentially leading to conceptual misunderstandings through pedagogically inconsistent explanations. Such situations may cause students to believe incorrect information to be true or reinforce their existing misconceptions. Furthermore, concerns related to data privacy, systemic biases, and ethical accountability are critically important, particularly in scenarios involving the storage and processing of student data [2] [11]. The increasing prevalence of LLM-based technologies in educational environments makes it imperative to establish institutional oversight mechanisms to maintain the reliability of these systems.

Overall, LLM-supported solutions offer multidimensional contributions to the educational ecosystem. Increased academic achievement, conceptual clarity, heightened motivation, consistency in assessment processes, and efficiency in instructional design are among the primary contributions. However, considering model errors, the risk of superficiality that excessive use may cause, and ethical issues, LLM integration must be carried out in a controlled, purpose-oriented manner and in line with pedagogical frameworks. When properly structured, these systems offer a powerful learning support mechanism that sustainably improves education quality.

REFERENCES

[1] [1] G. Fan, D. Liu, R. Zhang, and L. Pan, "The impact of AI-assisted pair programming on student motivation, programming anxiety, collaborative learning, and programming performance," International Journal of STEM Education, vol. 12, no. 16, pp. 1–17, 2025. https://doi.org/10.1186/s40594-025-00537-3

[2] S. Wang, T. Xu, H. Li, C. Zhang, J. Liang, J. Tang, P. S. Yu, and Q. Wen,"Large Language Models for Education: A Survey and Outlook,"*arXiv preprint* arXiv:2403.18105, 2024. https://arxiv.org/abs/2403.18105

[3] Z. Huang, Y. Li, and M. Chen, "Transformer-LSTM Models for Automatic Scoring and Feedback in English Writing Assessment," 2024. [Online]. Available: https://www.researchgate.net/publication/390938051_Transformer-LSTM_Models_for_Automatic_Scoring_and_Feedback_in_English_Writing_Assessment

[4] R. Roldán-Álvarez and J. Mesa, "Intelligent Deep Learning Tutoring System to Assist Instructors in Programming Courses,"*Applied Sciences*, vol. 14, no. 4115, pp. 1–24, 2024. https://www.mdpi.com/2076-3417/14/9/4115

[5] "Khanmigo," Khan Academy, Accessed: 2025-11-30. [Online]. Available: https://www.khanmigo.ai/

[6] "What is Duolingo Max?," Duolingo, Accessed: 2025-11-30. [Online]. Available: https://tr.duolingo.com/help/what-is-duolingo-max

[7] "LearnLM | Google Cloud," Google, Accessed: 2025-11-30. [Online]. Available: https://cloud.google.com/solutions/learnlm

[8] "OpenStax Assignable," OpenStax, Accessed: 2025-11-30. [Online]. Available: https://openstax.org/assignable

[9] "K-12 Large-Scale Assessments — Automated Scoring," Pearson Assessments, Accessed: 2025-11-30. [Online]. https://www.pearsonassessments.com/large-scale-assessments/k-12-large-scale-assessments/automated-scoring.html

[10] [2] C.-Y. Chung, I.-H. Hsiao, and Y.-L. Lin, "AI-assisted programming question generation: Constructing semantic networks of programming knowledge by local knowledge graph and abstract syntax tree," Journal of Research on Technology in Education, vol. 55, no. 1, pp. 94–110, 2023. https://doi.org/10.1080/15391523.2022.2123872

[11] Y. Sun, K. Matsuo, and T. Nakajima “LLM Agents for Education: Advances and Applications,”*arXiv preprint* arXiv:2503.11733, 2025. https://arxiv.org/abs/2503.11733

[12] D. Roldán-Álvarez and F. J. Mesa, "Intelligent Deep-Learning Tutoring System to Assist Instructors in Programming Courses," IEEE Transactions on Education, vol. 67, no. 1, pp. 153–161, Feb. 2024. https://doi.org/10.1109/TE.2023.3331055