The Application of Large Language Models in Vertical Domains: learning assistance

Dehou Lin*, Ping Li

Wenzhou Polytechnic, Wenzhou 325006, China
* Corresponding author

Abstract: Although the performance of many large language models (LLMs) is excellent in the general field, with the popularity of LLMs, more and more people find that the performance of LLMs in vertical domains is not satisfactory. Therefore, this paper decided to study the application of LLM in a vertical domains and use LLM to assist volunteer learning. This research plan uses the general LLM in combination with learning methodology, the knowledge enhancement of Skills and Retrieval-augmented Generation (RAG) along with structured guidance of the teaching process, to build a learning assistance AI. This study recruited 10 volunteers to use learning assistance AI to learn new textbooks, and most of them reported that they had improved through the learning of this system, could quickly get started, and quickly gained a preliminary understanding of new knowledge. Therefore, this study concludes that learning assistance AI is effective.

Keywords: Large Language Models, vertical domains, learning assistance AI.

1. Introduction

With the released of GPT-3.5, the LLM behind ChatGPT [1, 2], more and more people have begun to pay attention to and use LLMs. GPT-3.5 is based on the Transformer architecture [3] and trained on massive text data, which endows it with excellent semantic understanding. It can accurately understand user's intent and generate appropriate responses, allowing users to easily find answers to their questions. LLMs can generate high-quality, natural, and fluent text covering various styles and themes, such as news reports, academic papers, novels, and stories. This features provides strong support for content creation and copywriting. Thereafter, the company OpenAI released GPT-4 [4], which exceeded GPT-3.5 in performance across the board. In addition to OpenAI's GPT series, other companies and institutions have also launched their own LLMs, such as Pengcheng Lab's "Pengcheng Pangu" [5], Anthropic's Claude [6], and Alibaba's Qwen3 [7].

Although the performance of many LLMs is excellent in general domains, however, many people find that their performance in vertical Domains is disappointing. In general, general-purpose LLMs have the following shortcomings：

(1) Strong data dependence. If the training data is biased, inaccurate, or incomplete, LLMs will output wrong answers.

(2) Timeliness of training data. Outdated training data may cause LLMs to produce incorrect outputs.

(3) Hallucinations. Because LLMs make predictions based on statistical probability rather than causal reasoning, it is difficult to guarantee the accuracy of their outputs.

(4) Poor performance in vertical domains. Because LLMs are trained primarily on general knowledge, even for very simple problems within a vertical domain, their outputs may be unsatisfactory.

In order to address the above issues, some optimization approaches have been proposed for LLMs, such as fine-tuning [8, 9] and adding web search capabilities, which are a bit complicated and expensive. LLMs equipped with websearch may lead to false or misleading information due to the abundance of misinformation on the web.To conclude this work, we want to investigate if the learning assistance AI system affects volunteers' learning performance.

It is worth mentioning that in this paper we will use Skills and RAG to solve all of the above shortcomings. Using Skills to make the learning more effective, the system will record volunteers' learning and save them to Markdown. The system is intended for college education, enabling college volunteers to better use AI for learning assistance.

We asked volunteers to utilize the learning assistance AI system and conducted an exploratory user study which investigated the people's subjective learning performance when the learning assistance AI system is used for learning the new knowledge. We focused on the users perception in contrast to strict controlled experiments based on causal relationships, which explored the people' subjective experience from the learner's perspective. After 20 days of learning the learning assistance AI system, we interviewed the volunteers to ask if the tool has been a role in their learning in terms of knowledge understanding, learning performance, learning confidence and intention of using it.

In this study, we want to investigate if the learning assistance AI system affects volunteers' learning performance. After usage, volunteers believe that the system significantly improved their learning performance and show the practical application of learning assistance AI to improve the learning efficiency and give insights to the design of educational AI.

2. Related Work

Research on learning with the assistance of LLMs has begun to grow rapidly with the advancement of AI technology. We believe it can be mainly divided into three directions: using LLMs directly in education; vertical LLMs and RAG; and intelligent tutoring systems (ITS). These directions are not completely mutually exclusive, there is some overlap among them.

2.1. Using LLMs Directly in Education

When GPT-3.5 released, research on learning with the assistance of LLMs mainly focused on automatic question answering, replacing manual grading of essays, and applications in programming education [10]. Dong et al. pointed out that LLMs have great potential in educational scenarios such as medical educational content generation, English learning assistance, and academic research support. However, LLMs suffer from problems such as lack of common sense, outdated knowledge, and output of incorrect information, along with concerns about ethics and fairness [11]. Benyamin Tabarsi et al. argued that LLMs can improve volunteers' learning efficiency and engagement, but directly providing answers to questions may inhibit the development of their metacognitive abilities and problem-solving skills [12]. Cibu et al. analyzed other papers and pointed out that LLMs can improve learning efficiency, support simulation training in professional fields such as medical education, automatic evaluation and feedback of academic research, and provide mental health support, but also raised concerns about LLMs privacy protection and education inequality [13].

LLMs commonly exhibit problems such as hallucination and outdated knowledge in educational contexts. To solve these problems, many researchers how to adapt LLMs to vertical domains.

2.2. Vertical LLMs and RAG

Generally, when LLMs are used in a vertical domain, the main approaches are fine-tuning a large model to produce a vertical domain LLM, or employing RAG techniques.

Unggi Lee et al. constructed a new model, Llama-Polya by fine-tuning Meta's Llama-3.1-8B model. The reasoning process of this model is more balanced; it reduces premature answers and enhances teaching coherence and metacognitive prompting. However, its performance remains insufficient in terms of personalization and mathematical rigor [14]. Sharma et al. fine-tuned the Mistral-7B model using QLoRA on their own ConvoLearn dataset. Its performance matches that of a strong closed-source benchmark model. However, the model has limitations regarding simulated data, real-world settings, and the relationship between evaluation signals and direct validation [15]. Although fine-tuning an LLM is a good method, its cost is high and unaffordable for ordinary users. For example, fine-tuning a Llama 70B model requires thousands of dollars and many expensive GPU cards. Therefore, many scholars have proposed RAG technology.

The RAG technique can mainly solve the problems of knowledge timeliness, hallucination and lack of professional knowledge. It dynamically obtains information through an external knowledge base, and acquires real-time or private knowledge without adjusting the LLM. Zongxi Li et al. argue that augmenting LLMs with external knowledge retrieval can effectively enhance factual accuracy and enable real-time knowledge updates, thereby addressing LLMs hallucinations and the problem of static knowledge obsolescence [16]. Norhayati Yahaya et al. indicate that RAG technology does indeed improve the performance of educational AI systems. [17] Nevertheless, RAG technology also has limitations, such as increased system complexity, higher token consumption, and poor performance in complex reasoning, among others.

2.3. Intelligent Tutoring Systems

The design principle of Intelligent Tutoring Systems serves as a theoretical framework for learning assisted AI. David Jones et al. proposed IntelliCode, a multi-agent LLM system that tracks learning by a version controlled profile with six coordinated agents, and learns about learning, misuse, review, and engagement, aiming to be like a human teacher. It was first tested with simulated learners, not in real schools, and did not "thoroughly" tackle ethical concerns like privacy protection or educational inequality [18]. Yuxin Liu et al proposed a self-learning AgentTutor based on user feedback that considers volunteers' level and style as long as they listen to it, which learns personalized adaptive learning by multi-turn and dynamic meta-adaptation [19]. These systems solve problems with learning assistance AI such as short memory, rigid overstrategies, and reflection, but are more expensive, unable to update knowledge in real time, and suffer from over-the-hallucinations.

2.4. Research Gap and Positioning of This Study

Although a large number of studies have explored the application of LLMs in educational assistance, the following problems still exist:

(1) Insufficient depth in the integration of learning methodologies with LLMs technologies;

(2) LLMs system that allow arbitrary addition of textbooks and can be directly used in the learning process are limited;

(3) Scarcity of research addressing the dimension of skills.

This study proposes a vertical-domain LLM system called learning assistance AI. The system combines our learning methodology, RAG and Skills to give accurate learning help that leads users through each step. The effectiveness of the system is verified through user experience interviews.

3. Method

3.1. Learning methodology

For learning from a textbook, AI serves merely as an auxiliary tool. Students must still integrate effective learning methods to complement AI. This study proposes an learning assistance AI system that adopts the SQ3R method for intensive reading, based on our experience and the inherent characteristics of AI. In reviewing, the Feynman technique, learning by teaching, to let participants understand the knowledge points. Review intervals are scheduled with reference to the Ebbinghaus forgetting curve, rather than strictly following its timing. Furthermore, in order to help subjects distinguish between ideas, we alternate programming with mathematics.

SQ3R is a learning technique of Francis P. Robinson [20]. It consists of five steps: Survey, Question, Read, Recite and Review. In this study, we made slight changes to SQ3R. These five parts of the proposed method could not be used in the original paper:

(1) Survey: Quickly scan the content and extract the key points.

(2) Question: Set key questions in order to set goals for the future reading.

(3) Read: Read intensively with the previous questions, then answer the Question step questions, and compare the Question Step questions with our system questions.

(4) Recite: Rephrase the core content in one's own words, and compare it with the content and performance generated by the AI.

(5) Review: Review the learned content based on the principle of the Ebbinghaus forgetting curve.

In fact, there is also an upgraded version of SQ3R, i.e., PR4R, which adds a step called Reflect compared to the former. This paper integrates the component named Reflect into the Review step, allowing participants to reflect while reviewing the content.

The core idea of Feynman Feynman technique is to explain concepts to others through popular language, so as to force oneself to think deeply, deepen the understanding of knowledge, and achieve the effect of reflection. This system sets a LLM as a white space, and then volunteers teach AI what they have learned, let the LLM evaluate and score, and volunteers check whether they have blind spots according to the LLM feedback.

The Ebbinghaus forgetting curve describes a pattern of memory discovered by German psychologist Herman Ebbinghaus through experiments. Immediately after information is initially learned, forgetting occurs rapidly. As time passes, the rate of forgetting gradually slows down [21]. However, this forgetting curve is more suitable for humanities learning. We improved it by incorporating LLMs. The details are as follows:After each chapter is studied, a review is conducted and scored. Before studying Chapter 2, Chapter 1 is reviewed and its score is recorded. When reaching Chapter 3, the score of the second review of Chapter 1 is compared with the score of the review of Chapter 2: the chapter with the lower score is selected for review; in case of a tie, Chapter 2 is reviewed. Then the system picks the recent review score for each chapter, selects the three with the lowest scores, and randomly chooses one of them to review.

3.2. SKILLs

Skills that are also called Agent Skills are skill pack or reusable capability modules for LLMs. Their goal is to pack vertical domain knowledge and workflows into a unit.A typical Agent Skill may be a skill.md file describing the task and usage and three folders: scripts/, references/, assets/.

In this study, we use two skills called pdf-split and research-summarizer. pdf-spl splits a PDF file, and breaks it down, by a few hundred pages, into several smaller files of a few dozen pages. Research-summerizer reads, analyzes, and studies a PDF.

Textbook PDFs typically contain several hundred pages and research-summarizer has no chunks.If the PDFs are directly read, then this content will saturate the context window.Therefore, we first split the PDF in chapters and keep each chunk within several dozen pages. After chunking, research-survey is used to analyze the PDF with SQ3R at prompt instructions but only used three steps, Survey, Question, and Recite.After analyzing each chunk, we save these results to Markdown file.

3.3. RAG

RAG can usually be divided in two parts: offline indexing and online inference. In offline, knowledge bases are chunked, vectorized and indexed. In online, knowledge fragments about the query are retrieved as context to guide the LLM to find an answer [22].

Many AI coding assistant support reading files, but they are generally inefficient. For example, Claude Code reads files using Read and Grep tools. It first uses Grep to search for keywords in the codebase, then reads the relevant files with Read. For files that are hundreds of pages long, using its built-in tools is very hard. Therefore, this system chose to employ RAG technology to process textbooks, making it easy for volunteers to retrieve the original text efficiently.

Originally the system wanted to build RAG from scratch by hand. But as the technology evolves, other RAG tools like Dify, MaxKB, FastGPT, AnythingLLM etc, have come to our attention. As RAG is aimed to to make it easy for volunteers to read the original text and does not have much impact on the learning process, we choose a tool that can quickly be installed. And finally AnythingLLm was chosen as our RAG tool. Use goes as follows:

Create a workspace in AnythingLLM, connect to an LLM, import the PDF into your workspace, separate the images and text from the PDF, divide the extracted content into smaller chunks of about 500 characters, and convert the text chunks into vectors and return them to a vector database.

3.4. System Pipeline

This experiment take only two learning subjects: programming and mathematics. The textbook PDFs that we chose were university-level programming and mathematics., we split the textbook chapters. The chunks were saved into two folders named after subjects. We read each chunk through research-summarizer with only the Survey, Question and Recite steps of SQ3R and save the analyses as Markdown with each individual folder.At the same time, we run AnythingLLM on both textbook PDFs and store the results into a vector database. It was easy for volunteers to quickly read the original text.

Volunteers enter the system to begin learning. In the first session, the system will create a Markdown file called Volunteer's name for the volunteer, and then provides the intensive reading content of Chapter 1 for the volunteer to study. Volunteers read Survey and Question sections, and at the same time the system returns PDF of Chapter to read the PDF with the questions to mind.Volunteer answers the questions generated in Question, and they will compare their answers with the system output to decide if they are reasonable. Once learning is complete the Feynman technique is in place: volunteers express their knowledge about themselves in terms of words they know and compare it to Recite. Then volunteers' information will be stored in a Markdown file with fixed fields, including serial number, subject, chapter, date, score, short comment, and whether review, allowing the LLM to conduct rapid analysis. After a ten-minute break, volunteers move on to the next subject after taking the same steps. After the second subject, they visit the first one and add content to the Markdown file and then to the second, adding content, it is worth noting that only the review part is given a score. The Markdown files contain all their learning and review activities. In chapter 2, volunteers review the content of chapter 1. Note that all review takes place before learning. This review mainly uses the Feynman Technique: volunteers explain the key knowledge from the previous chapter in their own words to the system. Then the system provides evaluation to deepen the volunteer's understanding, and then records a quick comment and a score in Markdown file. During the third chapter, the review scores from the first two chapters' scores are compared, content with low scores is reviewed, and if the scores are the same, the second chapter is reviewed. From the fourth chapter onward, from the most recent review of each section one of the three lowest-scoring entries is randomly selected for review. After the review is completed, the corresponding scores in the Markdown file are updated. The specific process is illustrated in Figure 1.

4. Discussion and Implications

4.1. Discussion and Implications

For each of these 10 volunteers, we performed a 20-day learning session using the system to learn programming and advanced mathematics. None of them had studied programming or advanced mathematics prior to this learning session. Throughout the course of the learning process, we followed the participants' learning experience and interviewed them in detail after the lesson, asking them to report their own perception of the system during the real learning process. Based on their reports, we investigated how the system affects real learning.Since each volunteer has different learning abilities and backgrounds, we did not provide examination-based assessment. Our main findings are the following:

4.1.1. Knowledge Comprehension

Most volunteers felt that the system helped them to get better understanding of what they had learnt. The LLM could explain more complex concepts in more accessible language, and the LLM's output was more consistent with the text. Volunteers had learned about this difficulty because books often are too abstract to explain, and this system was able to generate much easier-to-understand content. Even if someone could not understand the output, they could reason with the system in a way that rewiring explanations of knowledge points, which would make the output more consistent for their knowledge level and comprehension.

4.1.2. Learning Efficiency

Seven volunteers believed the system had improved their learning ability.They felt the system's learning methodology was very good, and the LLM enhanced the effectiveness of this methodology. The Survey generated by the system allowed them to grasp the knowledge structure of each chapter and their learning goals. The questions generated were mostly about major content of that chapter and the answers were good. The Feynman technique that explained what they had learned in a way to the system was also very helpful. Usually, after learning some new knowledge they would forget that stuff within a few days, but Ebbinghaus forgetting curve made them know more. Most of them also thought they were introverted. They said that explaining to the machine had less psychological burden than explaining to other people because they had no worry that they would make mistakes or feel embarrassed. This actually enabled them to use the method more effectively.

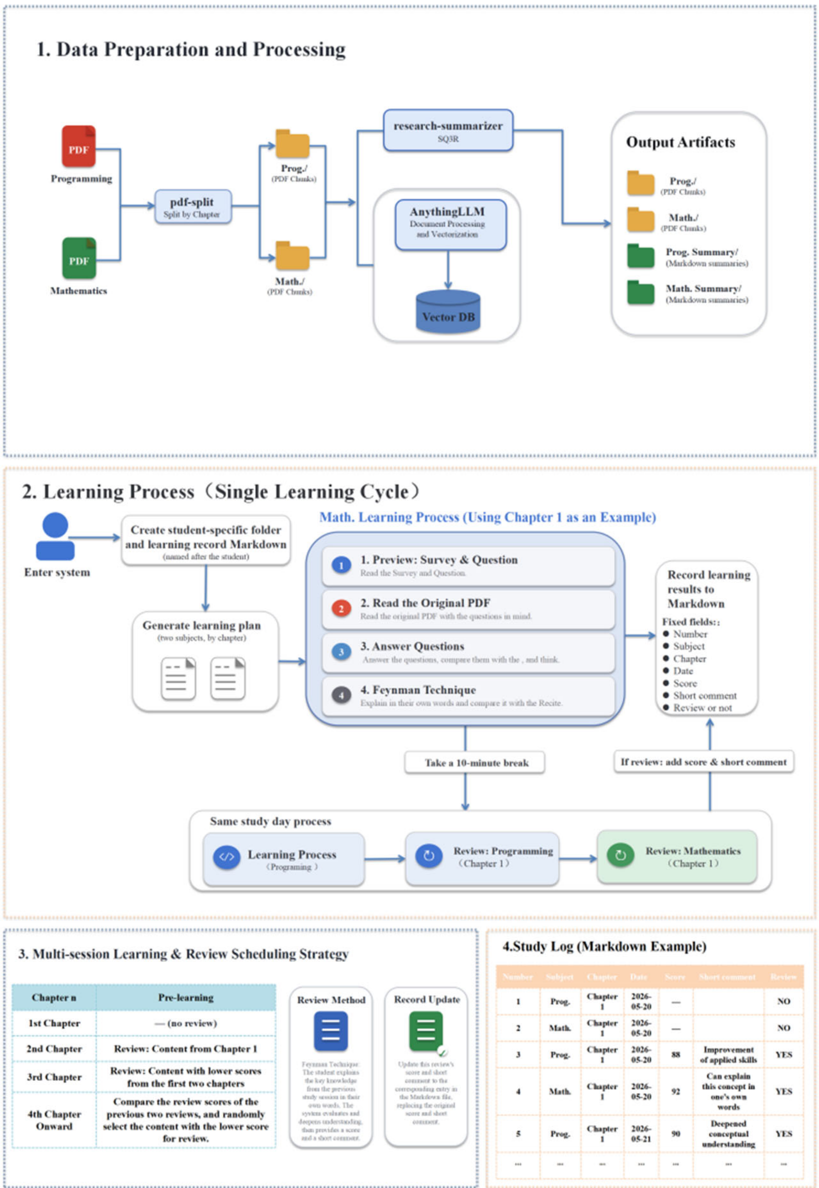

Figure 1. System pipeline

Of the last three volunteers, one didn't really trust what the LLM generated and spent a lot of time checking it. The other two already had their own learning methodology and just weren't comfortable with how this system worked.

4.1.3. Increased Learning Confidence

Of the seven volunteers who believed the system improved their learning efficiency, six said that after using the system, they felt more confident to learn more complex knowledge. They believed that the system's learning methods and the content generated by the LLM greatly helped them learn new knowledge. One of them was worried that relying too much on the system for a long time might stop people from thinking for themselves.

4.1.4. Continuance Intention

Six volunteers thought that having the system could be a help in learning, and were very willing to continue to use it. Two volunteers hoped that the system could adapt to their own learning methodology; if this were possible, they would be willing to continue using it. One volunteer felt that if he could quickly check the original output of the LLM, then he would also be very willing and willing to keep using it, and one other volunteer felt this would inhibit him from learning.

4.2. Research Limitations

(1) Sample and Time Limitations: We only recruited 10 volunteers and took 20 days to learn, and could not see if their long-term learning habits or progress would develop.

(2) Lack of Objective Metrics: Because volunteers vary in learning ability and prior knowledge, we used the learners' subjective knowledge.While volunteers' self-reports could be used as evidence, it was not possible to verify that their self-report is consistent with the actual learning results.

(3) No control group: This study could not rule out the influence of factors such as placebo effects or natural growth over time, because it is not a causal experimental design.

(4) LLM Performance: LLMs may vary in performance, and volunteers' experiences and learning outcomes might change depending on the LLM used. This study used the DeepSeek-4-Pro model for cost reasons.

4.3. Practical Implications

4.3.1. Recommendations for learning assistance AI System Designers

Based on the above, the following suggestions are offered for the design of AI-powered learning assistants:

(1) Everyone has their own learning methodology. Designers should provide a modular learning methodology that users can use and combine on their own.

(2) For the user to verify the outputs from the LLM, one or more other LLMs can be used to analyze the outputs of the system LLM, or the outputs from several LLMs can be given directly to the user so that they can judge for themselves.

5. Conclusions and Future Work

5.1. Conclusions

Based on the process tracking and interviews taken from 10 volunteers during a 20-day learning learning session, this study concludes:

5.1.1. Overall Positive Reception

Most volunteers indicated that the system helped them deepen their understanding of knowledge and improve their learning efficiency, and most volunteers indicated that they would continue to use the system. The LLM can transform complex and abstract concepts into accessible language, and the conversational mechanisms adapt dynamically to the level of understanding of the individual user, making it easy to acquire knowledge. The system combines surveys and hint descriptions generated using the research summary generator to help users quickly build a knowledge framework and focus on the key content of the generated issues. While the Feynman Technique deepens the user's understanding of knowledge points, the system combines memory retention principles to help the user effectively preserve knowledge in the long-term.

5.1.2. Individual Differences and Potential Risks

If a learning method does not match the user, it may lead to reduced learning efficiency. Therefore, while focusing on LLMs, we also need to conduct more in-depth research into traditional learning methods. One volunteer expressed concern that long-term use of learning assistance AI system might inhibit independent thinking. This concern is an issue that future research needs to take seriously.

5.1.3. Research Positioning and Generalization Boundaries

This study analyzes users' subjective perceptions rather than testing causal relationships in an experimental design. Due to limitations such as sample size and learning period, caution is needed when generalizing the conclusions. The findings of this study are mainly applicable to understanding users' subjective experiences when first using an learning assistance AI syetem, and should not be used to directly infer long-term effects or causal relationships.

In summary, the learning assistance AI system received positive recognition from the majority of volunteers in terms of improving knowledge comprehension and learning efficiency, particularly in providing learning communication support for introverted learners. Therefore, this study concludes that learning assistance AI systems have a positive impact on enhancing learning outcomes.

5.2. Future Work

Based on the above, future work should be advanced in the following directions：

(1) System Optimization: Optimize the system based on volunteer interview feedback, for example, by adding modular learning method components.

(2) Incorporating Quantitative Research: Conduct objective knowledge tests and quantitative assessments of learning ability before and after volunteers use the system, in order to explore the quantitative relationship between subjective perceptions and actual usage outcomes.

(3) Observing the Impact on Learning Ability from Long-Term Use: Recruit volunteers to use the system over an extended period to observe the various effects of long-term use, such as whether prolonged reliance on learning assistance AI systems enhances or inhibits users' learning abilities.

6. Project Fund

This work was supported by the General Research Project of Wenzhou Polytechnic (Grant No. WZY2025025).

References

[1] Ye, J., Chen, X., Xu, N., et al. (2023). A comprehensive capability analysis of GPT-3 and GPT-3.5 series models. arXiv preprint arXiv:2303.10420. https://doi.org/10.48550/arXiv.2303.10420

[2] Chen, L., Zaharia, M., & Zou, J. (2023). Analyzing ChatGPT’s behavior shifts over time. arXiv preprint arXiv:2307.09009. https://doi.org/10.48550/arXiv.2307.09009

[3] Vaswani, A., Shazeer, N., Parmar, N., et al. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30. https://doi.org/10.48550/arXiv.1706.03762

[4] Achiam, J., Adler, S., Agarwal, S., et al. (2023). GPT-4 technical report. arXiv preprint arXiv:2303.08774. https://doi.org/10.48550/arXiv.2303.08774

[5] Zeng, W., Ren, X., Su, T., et al. (2021). Pangu-α: Large-scale autoregressive pretrained Chinese language models with auto-parallel computation. arXiv preprint arXiv:2104.12369. https://doi.org/10.48550/arXiv.2104.12369

[6] Liu, J., Zhao, X., Shang, X., & Shen, Z. (2026). Dive into Claude code: the design space of today's and future ai agent systems. arXiv preprint arXiv:2604.14228. https://doi.org/10.48550/arXiv.2604.14228

[7] Yang, A., Li, A., Yang, B., et al. (2025). Qwen3 technical report. arXiv preprint arXiv:2505.09388. https://doi.org/10.48550/arXiv.2505.09388

[8] Lialin, V., Deshpande, V., Yao, X., & Rumshisky, A. (2023). Scaling down to scale up: a guide to parameter-efficient fine-tuning. arXiv preprint arXiv:2303.15647. https://doi.org/10.48550/arXiv.2303.15647

[9] Han, Z., Gao, C., Liu, J., Zhang, J., & Zhang, S. Q. (2024). Parameter-efficient fine-tuning for large models: a comprehensive survey. arXiv preprint arXiv:2403.14608. https://doi.org/10.48550/arXiv.2403.14608

[10] Kasneci, E., Seßler, K., Küchemann, S., et al. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274

[11] Dong, B., Bai, J., Xu, T., & Zhou, Y. (2024). Large language models in education: a systematic review. In *2024 IEEE International Conference on Artificial Intelligence and Education* (pp. 131–134). IEEE.

[12] Tabarsi, B., Yasir, T., Reichert, H., et al. (2024). Herald of advancement, disruption, or both: a systematic literature review on student-facing llm tools in undergraduate computing education. Authorea Preprints. https://doi.org/10.22541/au.17123456.1

[13] Cibu, B., Crăciun, L., Molănescu, A. G., & Cotfas, L. (2025). Exploring the educational applications of large language models: a systematic review and topic analysis. *Electronics*, 14(23), 4683. https://doi.org/10.3390/electronics14234683

[14] Lee, U., Jeong, Y., Lee, C., et al. (2026). Llama-polya: instruction tuning for large language model based on polya's problem-solving. arXiv preprint arXiv:2602.10597. https://doi.org/10.48550/arXiv.2602.10597

[15] Sharma, M., Pea, R., & Subramonyam, H. (2026). Convolearn: a learning sciences grounded dataset for fine-tuning dialogic ai tutors.

[16] Li, Z., Wang, Z, Wang, W., Hung, K., Xie, H., & Wang, F. L. (2025). Retrieval-augmented generation for educational application: a systematic survey. *Computers and Education: Artificial Intelligence*, 8, 100417. https://doi.org/10.1016/j.caeai.2025.100417

[17] Yahaya, N., Ngah, W., Jailani, W. A., et al. (2025). Ai-powered learning content generation through retrieval-augmented generation for improved accuracy and personalisation. *Jurnal Pendidikan Malaysia*, 50(2), 45–62.

[18] David, J., & Ghosh, S. (2026). Intellicode: a multi-agent llm tutoring system with centralized learner modeling. In *Proceedings of the 2026 International Conference on Educational Artificial Intelligence* (pp. 129–138). IEEE.

[19] Liu, Y., Song, Z., Lou, J., Wu, C., & Li, J. (2025). Agenttutor: empowering personalized learning with multi-turn interactive teaching in intelligent education systems. arXiv preprint arXiv:2601.04219. https://doi.org/10.48550/arXiv.2601.04219

[20] Robinson, F. P. (1946). *Effective study* (Rev. ed.). Harper & Brothers.

[21] Memory, O. K. C. (n.d.). *Memory: a contribution to experimental psychology*.

[22] Lewis, P., Perez, E., Piktus, A., et al. (2020). Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems*, 33, 9459–9474. https://doi.org/10.48550/arXiv.2005.11401