Scientific Open Access Publishing

Education Insights

Article

Generative AI-Driven Practical Teaching Reform of Artificial Intelligence Application Courses

Yebo Gu [1,†], Hanqing Wang [1,†] and Fenhua Bai [1,*]

[1] Kunming University of Science and Technology, Kunming, China
* Correspondence: Fenhua Bai, Kunming University of Science and Technology, Kunming, China
† These authors contributed equally to this work

Abstract: With the rapid proliferation of generative artificial intelligence (AI) in higher education, AI-related practical courses generally face problems such as student over-reliance on AI tools, diminished depth of algorithmic understanding, and insufficient code reproduction capability. These issues frequently lead to critical learning gaps, notably the "understanding--generation disconnection" and the "understanding--reproduction decoupling." To address these contemporary challenges and enhance students' comprehensive practical abilities in real project scenarios, this study utilizes the Artificial Intelligence Application course at a southwestern engineering university as a primary vehicle. We construct an integrated teaching mechanism comprising "specification--generation--comparison--verification--evaluation" and carry out a comprehensive teaching reform through two rigorous rounds of classroom action research. Based on 400 valid questionnaires and detailed course test data collected from 120 students, SPSS 17.0 is utilized to analyze scale reliability, validity, and score differences. The empirical results show that in the second round of teaching, students' performance in theoretical understanding, hands-on reproduction, and project presentation significantly improved compared with the first round. Furthermore, prompt strategies and fact-checking behaviors are significantly positively correlated with overall learning outcomes. Ultimately, the study demonstrates that systematically introducing generative AI usage norms, prompt training, and fact-checking mechanisms can effectively mitigate students' substitutive reliance on AI tools, thereby promoting deep learning and sustainable skill acquisition in practical engineering courses.

Keywords: generative ai; teaching reform; higher education; deep learning; learning outcomes

Received: 28 March 2026
Revised: 10 May 2026
Accepted: 22 May 2026
Published: 26 May 2026

1. Introduction

Against the backdrop of the coordinated advancement of "education power, digital China, and talent power" and the continuous implementation of digital education initiatives, generative AI has rapidly integrated into the learning of AI application courses in higher education, becoming a critical variable driving teaching reform and training model reconstruction. Compared with traditional programming assistance tools, natural-language human-computer interaction and intelligent generation capabilities not only output code and comments but also support algorithm clarification, error localization, and performance optimization through conversational interaction. These advanced functionalities enable generative AI to serve as both teaching assistants for instructors and learning partners for students, fostering a more interactive and dynamic educational environment [1, 2]. However, technological "empowerment" does not automatically equate to "efficiency gains." Without aligning classroom processes, assignment formats, and quality governance mechanisms, such convenience may inadvertently lead to challenges such as reduced depth of understanding and compromised academic integrity. For instance, students may rely excessively on AI-generated solutions without fully grasping the underlying concepts, which can hinder their ability to apply knowledge independently. Furthermore, the absence of robust mechanisms to monitor and evaluate

AI-assisted learning outcomes may exacerbate these issues, necessitating a comprehensive approach to integrate generative AI effectively into educational frameworks.

Observations from AI and computing courses reveal three common learning deviations after introducing generative AI: understanding-generation disconnection, where gaps emerge in translating ideas into executable programs; understanding-reproduction decoupling, characterized by poor cross-device reproduction due to a lack of holistic understanding; and generation dependence-transfer deficiency, where students develop a substitutive reliance on ready-made solutions, hindering their ability to transfer knowledge across contexts. Additionally, generative AI can fabricate false citations and reasoning, leading to "pseudo-understanding" and increasing the risk of academic misconduct when outputs are not rigorously verified. These challenges underscore the importance of designing AI application-oriented practical courses that enhance learning outcomes without compromising integrity or depth of competence. To address these issues, this paper examines an AI major practical course in southwestern China, proposing an integrated "specification-generation-comparison-verification-evaluation" teaching mechanism. This framework aims to systematically guide students through the process of understanding, generating, and critically evaluating AI-assisted outputs, thereby fostering deeper learning and reducing reliance on automated solutions. By analyzing teaching data and observations, the study identifies actionable pathways for quality governance, ensuring that generative AI serves as a tool for empowerment rather than a crutch that diminishes academic rigor. The findings contribute to the broader goal of cultivating compound talents capable of leveraging AI technologies effectively while maintaining high standards of academic integrity.

2. Research Design

2.1. Research Subjects and Setting

This study is conducted within the Artificial Intelligence (AI) major at an engineering university located in southwestern China. The curriculum for this major is structured around three key components: "AI foundation," "core technology," and "engineering applications." Currently, the program enrolls approximately 400 undergraduate students, with AI-focused classes offered across all academic years, from freshman to senior. The student demographic is notable for its high proportion of male students, reflecting broader trends in engineering disciplines. Additionally, students in this program exhibit strong practical needs, as they are frequently engaged in hands-on projects and make extensive use of generative AI tools to support their learning and research activities. These tools are integrated into various aspects of their academic and project work, enhancing their ability to solve complex problems and develop innovative solutions [2, 3].

The Artificial Intelligence Application course serves as a critical component of the professional practical teaching framework within the AI major. This course is specifically designed for students who have successfully completed foundational prerequisite courses, such as programming and machine learning. It provides a comprehensive, hands-on learning experience, requiring students to engage in the full lifecycle of AI project development [4]. This includes tasks such as data preprocessing, model construction, deployment, and verification, all of which are conducted within a standardized development environment configured by the university. The course is offered 2 to 3 times per semester, with each class accommodating approximately 50 to 60 students. The instructional format is predominantly task-driven, emphasizing active learning through group projects. These projects encourage collaboration and problem-solving, as students work together to address real-world challenges. By participating in this course, students gain practical experience and develop the technical and teamwork skills necessary for future careers in AI and related fields.

2.2. Research Methods and Overall Approach

This study employs a dual approach that integrates classroom action research with teaching experiments to explore effective methods for enhancing the Artificial Intelligence

Application course. Operating under a unified course syllabus and clearly defined teaching objectives, the research is structured into two distinct rounds of teaching implementation. In the initial round, the focus is on embedding generative AI usage norms, prompt engineering techniques, and foundational fact-checking practices into the existing teaching framework [5, 6]. This phase aims to identify common challenges and establish practical strategies for integrating these elements into the curriculum. In the subsequent round, the teaching procedures and evaluation methods are refined to address the insights gained from the first phase. This includes optimizing classroom workflows, enhancing assessment methods to emphasize process-based evidence, and incorporating explanatory evaluations. The ultimate goal is to develop a stable and effective implementation plan that aligns with the course objectives and supports student learning outcomes.

For the purposes of data collection and analysis, this study adopts a mixed-methods approach that combines both quantitative and qualitative methodologies. Quantitative data is gathered through standardized final exams and hands-on tests conducted during both rounds of teaching implementation. These assessments evaluate student performance across three critical dimensions: understanding, reproduction, and articulation. By comparing the results from the two rounds, the study seeks to identify measurable improvements in these areas. Qualitative data is collected through classroom observation records, project assignments, and reflective learning journals, which provide deeper insights into changes in student learning behaviors following the integration of generative AI tools [7]. Additionally, questionnaire responses and course performance data are analyzed to characterize students' usage patterns of large language models and their perceptions of these tools. This comprehensive dataset enables a nuanced analysis of the mechanisms underlying the teaching intervention, offering valuable insights into how generative AI can be effectively incorporated into educational practices to enhance learning outcomes.

3. Teaching Reform Implementation Measures

This study addresses critical challenges in teaching, focusing on the foundational aspects of understanding, reproduction, and articulation. It identifies prevalent issues such as the disconnect between understanding and generation, the decoupling of understanding and reproduction, and the dependency on generation that hinders effective transfer of knowledge. To tackle these issues, an integrated practical teaching framework has been developed, encompassing the stages of specification, generation, comparison, verification, and evaluation. This framework aims to create a cohesive and systematic approach to teaching reform, ensuring that students not only grasp theoretical concepts but also effectively apply them in practical scenarios. By emphasizing iterative processes such as comparison and verification, the framework fosters deeper comprehension and critical thinking [8]. Furthermore, the evaluation phase ensures continuous improvement and alignment with educational objectives, enhancing overall teaching efficacy.

3.1. Usage Norms and Verification Mechanisms

Tiered usage norms are established at the outset of the course to delineate clear boundaries for the application of generative AI tools. These norms categorize usage into three distinct levels: "encouraged use," which promotes active engagement with AI tools for creative and analytical tasks; "permissible reference," which allows students to consult AI-generated outputs as supplementary resources; and "prohibited substitution," which strictly forbids the replacement of original student work with AI-generated content. This structured approach ensures that students develop a nuanced understanding of the appropriate contexts for leveraging AI technologies while maintaining academic integrity and fostering critical thinking skills.

A comparative analysis methodology is implemented to enhance students' ability to critically evaluate outputs from generative AI models. This involves utilizing at least two distinct large language models (LLMs) or different versions of the same model for key academic tasks. Students are required to analyze the outputs in terms of accuracy, comprehensiveness, and practical applicability. To document this process, assignments include a "comparison description" subsection, where students record observed discrepancies, provide reasoning for their evaluations, and reflect on the implications of these differences. This practice cultivates a critical mindset and equips students with the skills to identify anomalies and inconsistencies in AI-generated content, thereby fostering a deeper understanding of the limitations and potential biases inherent in these technologies.

To address the issue of generative hallucinations, rigorous fact-checking and reproducibility protocols are established as integral components of the coursework. A "fact-checking checklist" is introduced to guide students in verifying the accuracy of AI-generated outputs, ensuring that all claims, formulas, and conclusions are substantiated with credible sources. Additionally, a "minimum runnable example" system is implemented, requiring students to conduct small-scale reproduction experiments or boundary tests to validate key findings. These experiments are designed to assess the reliability and applicability of the generated content in practical scenarios. Students must include detailed descriptions of their verification processes in their reports, which serve as a critical criterion for evaluation. This approach not only mitigates the risks associated with erroneous AI outputs but also instills a robust scientific rigor in students, preparing them for advanced academic and professional challenges.

3.2. Prompt Training and Task Progression

Prompt training and optimization are initiated through concise "prompt micro-activities" conducted prior to each unit task. These activities serve as a preparatory phase where the instructor demonstrates essential techniques such as task decomposition, role setting, input constraints, and output formatting. By guiding students through the process of designing and iterating prompts, the instructor fosters a deeper understanding of the mechanics behind effective prompt creation. Students are encouraged to attach "prompt drafts and optimization notes" to their assignments, which not only promotes active engagement in problem formulation but also addresses the common challenge of bridging the gap between understanding and generation. This approach ensures that students develop a systematic methodology for prompt refinement, ultimately enhancing their ability to tackle complex tasks with greater precision and creativity.

Task progression is structured as a three-tier framework encompassing "minimal task → feature extension → comprehensive project." This progression is designed to gradually build students' competencies, starting with foundational skills and culminating in advanced project-based applications. In the initial stage, students focus on implementing partial functions and completing code segments, which lays the groundwork for understanding core concepts. The middle stage emphasizes modification and optimization, where students are guided to refine their work based on provided examples, fostering critical thinking and problem-solving abilities. The final stage involves collaborative group projects that span the entire workflow, from problem definition to solution deployment. This comprehensive approach facilitates the transition from merely knowing how to use tools to mastering the ability to transfer knowledge and skills across diverse contexts. By engaging in progressively complex tasks, students not only enhance their technical proficiency but also develop teamwork and project management capabilities, preparing them for real-world applications in academic and professional settings.

3.3. Formative Evaluation and Evidence Archiving

Formative and peer evaluation integrate process evidence, such as learning logs, code version records, dialogue segments from language models, and detailed iteration descriptions, into the assessment framework. These elements enable a comprehensive understanding of the learning process and provide a basis for evaluating student progress [9–11]. Intra-group peer evaluation and in-class presentations further encourage students to critically assess the quality of their work and offer constructive suggestions for improvement. The archiving of learning evidence packages is a crucial component of this system. Each student or group is required to submit a detailed package that includes complete code, runtime logs, prompt optimization records, fact-checking checklists, project reports, and reflective analyses. These submissions serve dual purposes: they provide instructors with a robust dataset for final grading and contribute to the development of authentic case studies and datasets that can be utilized for future course enhancements. This iterative process fosters a reviewable and traceable teaching methodology, ensuring continuous improvement and refinement of educational practices. By systematically collecting and analyzing these evidence packages, educators can establish a closed-loop system that supports both immediate instructional goals and long-term pedagogical advancements.

4. Questionnaire Survey, Course Performance, and Effectiveness Analysis

4.1. Questionnaire Survey

A total of 412 questionnaires were distributed, and all were successfully collected, demonstrating a high level of participation and engagement among respondents. After a thorough review process, 12 responses were excluded due to incompleteness or evident patterns that suggested non-serious answers, resulting in 400 valid questionnaires. This yields a valid response rate of 97.09%, which is statistically significant and indicative of reliable data collection practices. The sample size is sufficiently large to provide a comprehensive representation of AI major students, ensuring that the findings can be generalized to broader academic contexts. Furthermore, the data collected offers valuable insights into the current usage patterns of large language models (LLMs) among students, highlighting trends and preferences that can inform future curriculum development and teaching strategies. The robust dataset also serves as a solid foundation for subsequent statistical analyses, enabling the validation of innovative teaching mechanisms and methodologies aimed at enhancing educational outcomes in AI-related fields.

4.2. Reliability and Validity Analysis of the Questionnaire

SPSS 17.0 was utilized to assess the reliability and validity of the questionnaire scale, ensuring robust statistical analysis. Reliability was evaluated using Cronbach's α, a widely accepted measure of internal consistency. The total scale achieved an α value of 0.857, demonstrating strong reliability. Furthermore, each subscale exhibited α values exceeding 0.70, with dimensions such as LLM usage intensity, prompt strategies, and fact-checking behaviors surpassing 0.80. These results indicate excellent reliability across the scale and high consistency among individual items, reinforcing the credibility of the measurement tool used in this study. Such reliability is critical for ensuring that the questionnaire accurately captures the intended constructs without significant measurement error.

Validity analysis was conducted to ensure the questionnaire's appropriateness for factor analysis. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy yielded a value of 0.841, indicating that the data were highly suitable for factor analysis. Additionally, Bartlett's test of sphericity was significant ($P = 0.000 < 0.05$), further confirming the adequacy of the dataset for this statistical procedure. Principal component analysis was employed to extract underlying factors, resulting in seven common factors with eigenvalues greater than 1.000. These factors collectively explained 71.324% of the cumulative variance, demonstrating a strong representation of the data's underlying structure. The extracted factors provide valuable insights into the dimensions measured by the questionnaire, ensuring its validity for academic research purposes.

4.3. Course Tests

To evaluate the impact of the two rounds of teaching reform, a standardized course test was conducted at the conclusion of each round. This test comprised three distinct components: a theoretical examination, a hands-on practical test, and a comprehensive project report. The course team meticulously designed and graded both rounds of tests to ensure uniformity and comparability across all components. During the first round of testing, a total of 58 students participated, and their performance in the theoretical exam, practical test, and project report was analyzed to assess the effectiveness of the reform measures. In the second round, 62 students completed all components of the test, providing a robust and comparable dataset for further analysis. The inclusion of both quantitative and qualitative assessments allowed for a holistic evaluation of student learning outcomes and the overall efficacy of the teaching reforms [12–14]. By maintaining consistency in test design and grading criteria, the course team ensured that the results accurately reflected the impact of the reforms on student performance. This approach not only facilitated a detailed comparison between the two rounds but also provided valuable insights into areas requiring further improvement in future iterations of the course structure.

4.4. Reliability and Validity of the Course Tests

To evaluate the stability and discrimination of the course test instruments across two rounds of curriculum reform, a detailed analysis of the reliability and validity of final exam scores was conducted. Reliability was assessed using Cronbach's α, a statistical measure of internal consistency. The combined reliability coefficient for the theoretical paper and hands-on test was calculated at 0.823, with the theoretical paper achieving an α of 0.801 and the hands-on test scoring an α of 0.812. These values, all exceeding the threshold of 0.80, indicate strong internal consistency and a stable representation of student abilities across different test formats. This suggests that the instruments are well-designed to measure the intended competencies effectively and consistently over time [15, 16]. Furthermore, the high reliability underscores the robustness of the assessment tools in capturing diverse aspects of student performance.

Validity was examined through multiple dimensions. Content validity was ensured by aligning the exam strictly with the syllabus and covering core requirements, including algorithm understanding, code debugging, reproduction, and solution explanation. These components were rigorously reviewed by multiple instructors to ensure alignment with the "understanding, reproduction, articulation" structure. Construct validity was supported by difficulty coefficients ranging from 0.35 to 0.78 and discrimination indices between 0.32 and 0.68, which reflect an appropriate balance between challenge and discriminatory power. Additionally, the correlation coefficient between theoretical and hands-on scores was calculated at 0.61 ($P < 0.001$), demonstrating a strong synergy and complementarity between the two components. This correlation highlights the integrated nature of the assessment, where theoretical knowledge and practical skills reinforce each other to provide a comprehensive evaluation of student capabilities.

The validity of the course tests was further substantiated through detailed analyses of their structural and discriminatory properties. Content validity was meticulously maintained by ensuring that the exam adhered strictly to the syllabus, covering essential topics such as algorithm understanding, code debugging, reproduction, and solution explanation. These elements were reviewed by multiple instructors to ensure that the test structure accurately reflected the "understanding, reproduction, articulation" framework. This rigorous review process guarantees that the exam content aligns with the intended learning outcomes and provides a comprehensive assessment of student competencies.

Construct validity was evaluated through statistical measures, including difficulty coefficients and discrimination indices. The difficulty coefficients, which ranged from 0.35 to 0.78, indicate that the test questions were appropriately challenging, neither excessively difficult nor overly simplistic. Discrimination indices, measured between 0.32 and 0.68, demonstrate the test's ability to differentiate effectively between students of varying skill levels [17, 18]. These metrics confirm that the test design achieves a balance between accessibility and the ability to identify differences in student performance.

The correlation between theoretical and hands-on scores, calculated at 0.61 ($P < 0.001$), further supports the validity of the assessment tools. This correlation underscores the complementary nature of the theoretical and practical components, highlighting their combined effectiveness in evaluating a broad spectrum of student abilities. By integrating these elements, the course tests provide a holistic measure of student performance, ensuring that both conceptual understanding and practical application are adequately assessed (As shown in Table 1).

Table 1. Reliability and Validity Indicators of the Questionnaire Scale and Course Tests

Tool Type	Indicator Category	Sample Size	Reliability Indicator	Validity/Discrimination Indicator
Questionnaire scale	Overall reliability	400	Total scale alpha = 0.857	-
Questionnaire scale	Subscale reliability	400	alpha range 0.702-0.842	-
Questionnaire scale	Structural validity (KMO)	400	-	KMO = 0.841, P = 0.000
Questionnaire scale	Factor structure and explanatory power	400	-	7 factors, cumulative variance 71.324%, loadings > 0.50
Classroom test	Overall reliability	120	Overall alpha = 0.823	-
Classroom test	Subpart reliability	120	Theory 0.801; computer-based test 0.812	-
Classroom test	Difficulty and discrimination	120	-	Difficulty 0.35-0.78; discrimination 0.32-0.68

5. Research Conclusions

Synthesizing the results of the two rounds of course tests, the generative-AI-embedded teaching reform demonstrates a significant positive impact on learning outcomes in the Artificial Intelligence Application course. The comparative analysis of student performance metrics reveals notable improvements across multiple dimensions. Specifically, in the second round of testing, the average theoretical exam score increased to 74.85, surpassing the first round's 68.42 by 6.43 points. Similarly, the average hands-on test score rose to 71.40, marking an 8.23-point improvement over the initial round's 63.17. Furthermore, the average project report score reached 79.12, exceeding the first round's 72.55 by 6.57 points. These advancements highlight substantial progress in students' algorithmic understanding, code reproduction, and solution articulation, effectively achieving the teaching objectives of "understanding, reproduction, articulation." Beyond numerical gains, these results underscore the transformative potential of integrating generative AI into educational frameworks, fostering deeper cognitive engagement and practical skill development. The findings suggest that such pedagogical innovations could serve as a model for broader applications in STEM education, particularly in courses requiring complex problem-solving and technical proficiency. Future research could explore the scalability of this approach across diverse academic disciplines, investigate long-term retention of skills, and assess the impact of generative AI on collaborative learning dynamics and interdisciplinary project outcomes.

6. Conclusion

This study investigates the integration of generative AI into higher education, focusing on its application within the *Artificial Intelligence Application* course at an engineering university in southwestern China. By centering on the three foundational baselines of "understanding, reproduction, and articulation," the research implements a comprehensive teaching mechanism that includes specification, generation, comparison, verification, and evaluation. Through the use of questionnaires, course tests, and observational methods, the study explores how large language models (LLMs) influence learning outcomes. The findings indicate that when clear usage boundaries are established, alongside prompt training, rigorous fact-checking, and formative evaluation, generative AI can effectively address challenges such as the disconnection between understanding and generation, the decoupling of understanding and reproduction, and the dependency on generation that hampers transferability. These improvements enhance students' abilities in algorithmic comprehension, code implementation, and articulation. However, the study acknowledges limitations, including the restricted research cycle, limited sample scope, and narrow range of indicators. Future research should aim to validate these mechanisms across diverse institutions and course types, employing advanced methodologies such as process mining and structural equation modeling to deepen insights. Additionally, expanding the scope to include interdisciplinary courses and exploring the integration of generative AI with other emerging technologies could provide broader implications for educational practices. This research offers a practical framework for designing and governing AI-focused courses, paving the way for more effective teaching strategies in the era of LLMs.

References

1. A. Zeb, R. Ullah, and R. Karim, "Exploring the role of ChatGPT in higher education: opportunities, challenges and ethical considerations," *The International Journal of Information and Learning Technology*, vol. 41, no. 1, pp. 99–111, 2024.
2. S. Nikolic, S. Daniel, R. Haque, M. Belkina, G. M. Hassan, S. Grundy, ... and C. Sandison, "ChatGPT versus engineering education assessment: a multidisciplinary and multi-institutional benchmarking and analysis of this generative artificial intelligence tool to investigate assessment integrity," *European Journal of Engineering Education*, vol. 48, no. 4, pp. 559–614, 2023.
3. D. R. Cotton, P. A. Cotton, and J. R. Shipway, "Chatting and cheating: Ensuring academic integrity in the era of ChatGPT," *Innovations in Education and Teaching International*, vol. 61, no. 2, pp. 228–239, 2024.
4. E. Kasneci, K. Seßler, S. Küchemann, M. Bannert, D. Dementieva, F. Fischer, ... and G. Kasneci, "ChatGPT for good? On opportunities and challenges of large language models for education," *Learning and Individual Differences*, vol. 103, p. 102274, 2023.
5. C. K. Lo, "What is the impact of ChatGPT on education? A rapid review of the literature," *Education Sciences*, vol. 13, no. 4, p. 410, 2023.
6. C. K. Y. Chan, "A comprehensive AI policy education framework for university teaching and learning," *International Journal of Educational Technology in Higher Education*, vol. 20, no. 1, p. 38, 2023.
7. A. S. George, "The potential of generative AI to reform graduate education," *Partners Universal International Research Journal*, vol. 2, no. 4, pp. 36–50, 2023.
8. X. Le, "Assessment reform in higher education: An ethical approach to harness the power of generative artificial intelligence," *Education Research and Perspectives*, vol. 51, pp. 156–185, 2024.
9. M. Kumari and A. Raj, "Generative AI and teacher education: Prospects in NEP 2020," *Journal of Science Innovations and Nature of Earth*, vol. 5, no. 4, pp. 18–20, 2025.
10. L. Zhao, F. Yuan, and L. Miao, "Exploration of the reform of programming courses based on generative AI," in *2024 14th International Conference on Information Technology in Medicine and Education (ITME)*, pp. 869–873, Sept. 2024.
11. J. Gang, "Research on the impact of generative artificial intelligence on future education reform," in *2024 5th International Conference on Education, Knowledge and Information Management (ICEKIM 2024)*, pp. 481–489, Aug. 2024.
12. E. Langran, M. Searson, and J. Trumble, "Transforming teacher education in the age of generative AI," *Exploring New Horizons: Generative Artificial Intelligence and Teacher Education*, vol. 2, no. 13, 2024.
13. M. Nyaaba, "Transforming teacher education in developing countries: The role of generative AI in bridging theory and practice," *arXiv preprint arXiv:2411.10718*, 2024.

14. I. Shimizu, H. Kasai, K. Shikino, N. Araki, Z. Takahashi, M. Onodera, ... and E. Kawakami, "Developing medical education curriculum reform strategies to address the impact of generative AI: qualitative study," *JMIR Medical Education*, vol. 9, no. 1, p. e53466, 2023.
15. T. Su and S. Bei, "Research on the practical teaching reform of 'Deep Learning and Applications' course supported by generative AI technology," *Journal of Educational Theory and Practice*, vol. 2, no. 4, 2025.
16. Y. Ma, Y. Su, M. Li, Y. Zhang, W. Chai, A. Huang, and X. Zhao, "Preparing students for an AI-driven world: Generative AI and curriculum reform in higher education," *Frontiers of Digital Education*, vol. 2, no. 4, p. 30, 2025.
17. H. Yu and Y. Guo, "Generative artificial intelligence empowers educational reform: current status, issues, and prospects," in *Frontiers in Education*, vol. 8, p. 1183162, June 2023.
18. C. K. Y. Chan and L. H. Tsi, "Will generative AI replace teachers in higher education? A study of teacher and student perceptions," *Studies in Educational Evaluation*, vol. 83, p. 101395, 2024.