Jiang, B. et al. (Eds.) (2025). Proceedings of the 33rd International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Teaching Machine Learning to Senior Primary Students: Evaluating a Course for AI Literacy in a Hong Kong Primary School

Siu Cheung KONG[a,b] *, Yin YANG[a] & So Sum CHOW[b]
[a]*Department of Mathematics and Information Technology, The Education University of Hong Kong, Hong Kong SAR*
[b]*Artificial Intelligence and Digital Competency Education Centre, The Education University of Hong Kong, Hong Kong SAR*
**sckong@eduhk.hk*

Abstract: This study examines the effectiveness of a machine-learning course designed to enhance artificial intelligence (AI) literacy among 145 sixth-grade students in a Hong Kong primary school. The six-hour course introduced foundational AI concepts, including supervised learning (artificial neural networks and K-nearest neighbours) and reinforcement learning, with hands-on activities featuring AlphAI robots. Quantitative analysis, supported by Wilcoxon signed-rank tests, revealed significant improvements in students' conceptual understanding of machine learning and deep learning. Qualitative feedback demonstrated the course's effectiveness, with hands-on activities using AlphAI robots proving effective in deepen comprehension and enabling practical application of machine learning methods. The findings show how the course demystify AI for young learners, encouraging critical engagement with it's "black box" nature. These insights emphasise the value of interactive, hands-on learning strategies in AI literacy education in preparing young students for an AI-driven future.

Keywords: AI literacy education; deep learning; machine learning; primary school, robotics

1. Introduction

The rapid rise of generative AI—a subset of machine learning capable of creating novel content—has reshaped how humans interact with technology. A prominent example is Generative Pre-trained Transformer (GPT), a class of large language models built on the Transformer architecture and attention mechanisms, enabling breakthroughs in natural language processing. Unlike traditional rule-based AI systems, generative AI models like GPT can use deep learning (multi-layered artificial neural networks) to produce human-like text, images, and other media, demonstrating machine learning's transformative potential. This progress allows AI to augment human expertise and solve complex problems across industries. Yet, despite its growing influence, the public often perceives machine learning—especially generative AI—as a mysterious "black box" (Tedre et al., 2021). The process, from user prompts to internal computations and final outputs, remains opaque to most people, leading to an unspoken misplaced fear for artificial intelligence and automation. This knowledge gap is especially concerning in education, where students should develop a foundational understanding of "how AI functions" to thrive in an increasingly AI-driven future (Royal Society, 2017).

2. Literature Review

2.1 Current Approaches of AI Literacy Education for Young Students

Recent advancements in AI literacy education recognise AI literacy as a vital skill for future generations, leading to the development of innovative programmes across primary and secondary education levels (Kong & Yang, 2023; Kong & Yang, 2024; Martin et al., 2023; Martin et al., 2024; Maspul, 2024).

Customised learning design focuses on developmental appropriateness and simplicity, ensuring young students can grasp complex AI concepts. Research indicates that interactive methods are effective in introducing machine learning concepts to younger audiences (Gibellini et al., 2023). Hands-on learning approaches have proven to be particularly effective across various age groups. Robot-assisted programmes are especially promising for primary learners (Hong & Kim, 2025).

2.2 Gaps in Current Research

Despite these advances, significant research gaps remain in the field of AI literacy education for young students. First, educational robots remain under-researched despite their potential to make AI concepts more transparent than conventional "black box" tools (Yue et al., 2022). The intuitive interface and visualisation capabilities of AlphAI learning robots with its software used in this study aim to address this gap, offering an innovative approach to demystifying AI processes for young learners (Absalon & Deneux, 2025; Yang & Kong, 2025).

Second, most AI literacy framework targets secondary and other students, with limited attention given to senior primary students (Kim et al., 2021). By focusing on sixth-grade students, this study seeks to address this oversight, providing a teaching practice that progresses from foundational artificial neural networks (ANN) to reinforcement learning while maintaining conceptual rigor. This focus on younger learners is crucial for developing early interest and understanding in AI, setting the groundwork for advanced study in later educational stages.

Despite AI's growing societal prominence, current educational practices often fail to provide effective instruction for developing AI literacy, especially among younger students. This gap highlights the urgent need for engaging curricula that can introduce core AI concepts while fostering hands-on experience. Building upon prior research that examined school teachers' perceptions of a machine learning course's effectiveness (Yang & Kong, 2025), the current study investigates the impact of the machine learning course—enhanced by AlphAI learning robots—on students' conceptual understanding and their learning experiences. Specifically, the study addresses the following research questions:

(1) To what extent did the machine learning course, incorporating AlphAI robots, enhance sixth-grade students' conceptual understanding of machine learning?
(2) How did students perceive the impact of the course on their conceptual understanding of machine learning?

3. Research Design

3.1 Course Structure and Implementation

The six-hour curriculum introduced sixth-grade students to fundamental machine learning concepts through structured sessions combining theoretical instruction and practical activities. Professionally trained teachers guided students through progressive learning activities using course worksheets and AlphAI robots to ensure comprehensive understanding (Yang & Kong, 2025).

3.1.1 Instructional Approach

The curriculum covered core machine learning and deep learning concepts, with particular emphasis on hands-on learning experience through robot training exercises to reinforce conceptual understanding while developing practical skills. A key instructional focus was clarifying both the relationship between AI, machine learning, and deep learning, and the distinct roles of fundamental concepts within this framework, as visualised in Figure 1.

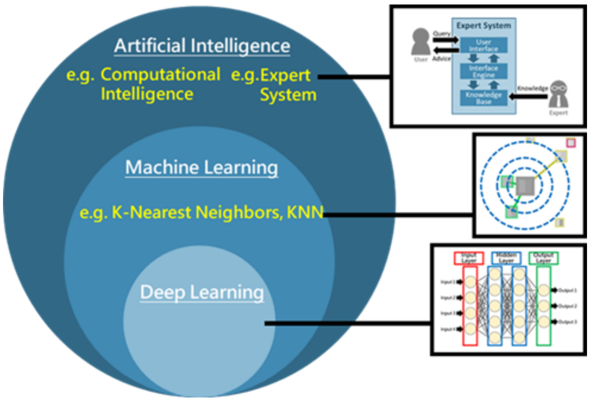
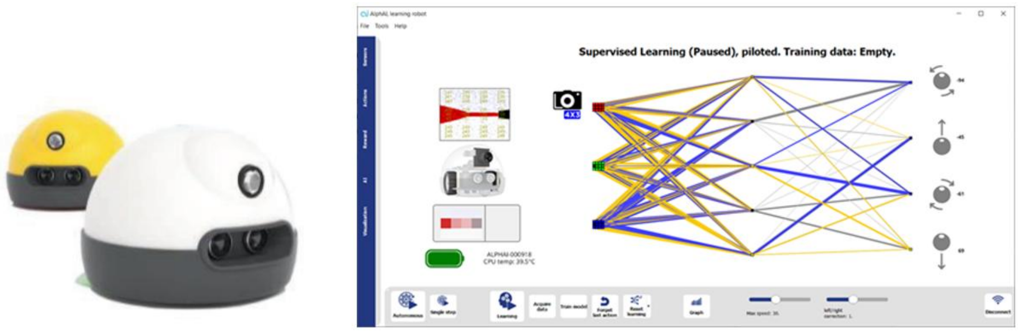


Figure 1. The relationship among commonly used terms of artificial intelligence, machine learning, and deep learning.

In particular, the course introduced foundational concepts beginning with ANNs—the basis of deep learning architectures—and supervised learning, a core machine learning method utilising labelled datasets to train predictive models. Subsequent units gradually progressed to advanced topics, including the role of weights and backpropagation in ANNs, reinforcement learning, and KNN, a machine learning algorithm distinct from ANN-based approaches. The course culminated in exploring the connections between core machine learning concepts—including the methods covered—and real-world applications, such as Google's Teachable Machine, which leverages GPT-like models.

3.1.2 Hands-On Learning Experience with AlphAI Robots

Hands-on activities were integrated into each session using AlphAI learning robots (Martin et al., 2023), providing students with experiential learning opportunities. These robots (Fig. 2(a)), connected to AlphAI software (Fig. 2(b)), offered a graphical interface that visualised AI algorithms like ANNs, Reinforcement Learning, and KNN. Students engaged in experiments such as training robots to race, modifying pixel inputs and movement outputs, and comparing learning algorithms, which helped reinforce abstract concepts.



(a). The AlphAI learning robots

(b). Visualisations of algorithm in the form of an artificial neural network on screen (i.e. supervised learning)

Figure 2. The AlphAI Learning Robots and Software Interface.

3.2 Participants

The study utilised a convenience sample of 145 sixth-grade students (94 boys, 51 girls; ages 10-13, *M*=10.94, *SD*=0.530) from five classes at a primary school. Class sizes ranged from 23 to 37 students. All participants had approximately two years of prior programming experience. Following ethical guidelines, we obtained written informed consent from both the school administration and parents.

3.3 Data Collection and Analysis

The study used a mixed-methods approach, collecting both quantitative and qualitative data: (1) pre- and post-tests on machine learning and deep learning concepts and (2) focus group interviews with 15 students. The 13-item concept test assessed understanding of machine learning and deep learning, showing moderate reliability (Cronbach's alpha = 0.656). Focus group interviews captured detailed student experiences with the machine learning course. Five focus groups of three students each, totalling 15 participants, were facilitated by a trained researcher for approximately 30 minutes. To address the first research question, a Wilcoxon signed-rank test was conducted to analyse pre- and post-test score differences in conceptual understanding. For the second question, thematic analysis explored qualitative feedback from focus group interviews.

4. Results

4.1 Analysis of Students' Conceptual Understanding of AI Literacy

The Wilcoxon signed-rank test revealed a significant improvement in students' AI literacy, with an increase from pre-test total scores (M = 5.03, SD = 2.072) to post-test total scores (M = 7.07, SD = 2.790), Z = 5.860, $p < 0.001$. Table 1 shows the descriptive data of each concept item. These results underscore a substantial enhancement in students' AI conceptual understanding following course completion. Gender-based analysis using the Mann-Whitney U test found no significant differences in score improvements between boys and girls (U = 1769.50, p = .365).

Table 1. *The Descriptive Data of Machine Learning and Deep Learning Concepts*

	Pre		Post		
	M	*SD*	*M*	*SD*	*Z*
Item 1	.31	.46	.56	.50	-4.185***
Item 2	.25	.44	.59	.49	-5.742***
Item 3	.69	.46	.76	.43	-1.050
Item 4	.48	.50	.63	.49	-2.121*
Item 5	.22	.41	.53	.50	-4.672***
Item 6	.43	.50	.54	.50	-2.018*
Item 7	.40	.49	.49	.50	-1.638
Item 8	.40	.49	.71	.46	-4.965***
Item 9	.29	.46	.35	.48	-.802
Item 10	.45	.50	.60	.49	-2.357*
Item 11	.48	.50	.52	.50	-.762
Item 12	.17	.38	.28	.45	-1.938
Item 13	.36	.48	.51	.50	-2.393*

*$p < .05$, **$p < .01$, ***$p < .001$

Significant improvements were observed in eight assessed items, particularly in understanding machine learning fundamentals. The results revealed enhanced comprehension of general machine learning processes, supervised learning, and specific algorithms including KNN and ANNs (including layer structures and neurons). Students also developed better awareness of backpropagation as a weight-adjustment mechanism and recognized overfitting as a common training challenge. The most notable improvement appeared in understanding backpropagation, with significantly more students correctly identifying it as a process where output errors are fed back to adjust neuron weights in hidden and input layers to optimize performance. These findings confirm the course's effectiveness in teaching core machine learning concepts.

4.2 Student Perceptions of the Course's Impact on Their Conceptual Understanding

Thematic analysis of five focus group interviews revealed three key themes about students' learning experiences: (1) enhanced conceptual understanding of machine learning, (2) valuable hands-on experience with AlphAI learning robots, and (3) recognition of real-world machine learning implementation challenges.

All 15 students reported benefiting from the course's structured approach. Many used analogies to explain concepts, like one student (ID9) comparing methods to classroom scenarios: "*Supervised learning is like a teacher guiding a student with labels, reinforcement learning uses rewards like in school, and unsupervised learning is when the robot explores alone—no labels, no teacher.*" Others demonstrated sophisticated insights, with a student (ID8) noting machine learning's data dependence: "*Machine operate by directly processing correct and incorrect knowledge (well- and poor-quality data) to achieve their purpose, but it needs to be given data to analyse (by human).*"

Students particularly valued AlphAI robots' physical interactivity. A student (ID15) contrasted passive lectures with active experimentation: "*Listening to the teacher explaining the concepts would be very abstract but actually doing it and testing it out would give a deeper impression.*" The iterative trial-and-error learning process was also a recurring theme in the interviews, with four students emphasising its educational value.

Practical challenges emerged during activities, particularly regarding data quality. A student (ID10) emphasised the critical role of high-quality training data: "*When we are making a model of ourselves with Teachable Machine, we need to take pictures of rock, paper, and scissors (gestures) from different angles (as inputs) to train the model to recognise them. However, the problem is that we may not be shooting from the right angle, and the model will not recognise my gestures accurately when it is trained and tested.*"

Most students expressed strong satisfaction with the course. A student (ID13) summarised: "*Training models and competing was fun and memorable—I even thought about how to optimise them at home.*" These findings demonstrate the course's success in making machine learning accessible while fostering critical understanding of its real-world constraints.

5. Conclusion and Implication

This study demonstrated the effectiveness of a machine learning course incorporating AlphAI robots in enhancing AI literacy among sixth-grade students. The quantitative analysis, notably the Wilcoxon signed-rank test, revealed significant improvements in students' understanding of machine learning and deep learning concepts. Additionally, the qualitative insights from focus group discussions underscored the importance of hands-on learning experience in demystifying complex AI concepts for young learners.

These findings have important implications for AI literacy education in the K-12 setting, particularly for young students. The results indicate that a well-designed curriculum that balances theoretical knowledge with practical application can enable students aged 10 to 13 to grasp fundamental machine learning concepts and critically evaluate the training process. To uncover the "black box" nature of machine learning, educational frameworks should emphasise scaffolding complexity and fostering active reflection among students.

However, the generalisability of this study is limited by its small sample size. Future research should consider expanding the sample size and including more diverse educational settings to validate and extend these findings. Additionally, extending instructional time and refining educational materials could further improve comprehension and engagement. By overcoming these limitations, future studies can further improve AI literacy courses for young learners, equipping them with the skills and critical insight necessary to thrive in an increasingly AI-driven world.

Acknowledgements

The authors would like to acknowledge the funding from the Central Reserve Allocation Committee, The Education University of Hong Kong of the Hong Kong Special Administrative Region, China (Uncovering the "Black Box" of Machine Learning: Promoting Artificial Intelligence Literacy with AlphaAI robots in Senior Primary/Junior Secondary Schools across Hong Kong and France). (Project No. EdUHK 04A55).

References

Absalon, M., & Deneux, T. (2025, April). "AlphAI": Teaching AI algorithms to K12 by training learning robots and visualizing how it works. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 39, No. 28, pp. 29096-29103). https://doi.org/10.1609/aaai.v39i28.35181

Gibellini, G., Fabretti, V., & Schiavo, G. (2023, April). AI education from the educator's perspective: Best practices for an inclusive AI curriculum for middle school. In *Extended abstracts of the 2023 CHI conference on human factors in computing systems* (pp. 1-6). https://doi.org/10.1145/3544549.358574

Hong, J., & Kim, K. (2025). Impact of AIoT education program on digital and AI literacy of elementary school students. *Education and Information Technologies, 30*(1), 107-130. https://doi.org/10.1007/s10639-024-12758-0

Kim, S., Jang, Y., Kim, W., Choi, S., Jung, H., Kim, S., & Kim, H. (2021, May). Why and what to teach: AI curriculum for elementary school. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 35, No. 17, pp. 15569-15576). https://doi.org/10.1609/aaai.v35i17.17833

Kong, S. C, & Yang, Y. (2023, December). Designing and evaluating an Attention-Engagement-Error-Reflection (AEER) approach to enhance primary school students' artificial intelligence literacy and learning-to-learn skills: A pilot study. In W.Q. Chen, et al. (Eds.), *Workshop Proceedings of the 31st International Conference on Computers in Education ICCE 2023* (pp. 394-399). Matsue, Shimane, Japan: Asia-Pacific Society for Computers in Education. https://doi.org/10.58459/icce.2023.1417

Kong, S. C., & Yang, Y. (2024, May). Using the robot-assisted attention-engagement-error-feedback-reflection (AEER) pedagogical design to develop machine learning concepts and facilitate reflection on learning-to-learn skills: Evaluation of an empirical study in Hong Kong primary schools. In *Proceedings of the 16th International Conference on Computer Supported Education - Volume 2* (pp. 155-162). Angers, France. https://doi.org/10.5220/0012505700003693

Martin, M., Chevalier, M., Burton, S., Bonvin, G., Besançon, M., & Deneux, T. (2023, April). Effects of introducing a learning robot on the metacognitive knowledge of students aged 8–11. In *International Conference on Robotics in Education (RiE)* (pp. 169-183). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-38454-7_15

Martin, M., Deneux, T., & Chevalier, M. (2024, April). From automaton to AI robot: The added value for learning. In *International Conference on Robotics in Education (RiE)* (pp. 403-410). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-67059-6_35

Maspul, K. A. (2024). Using interactive language development tools to enhance cognitive and literacy skills in K-12 education. *J-SHMIC: Journal of English for Academic, 11*(1), 20-35. https://doi.org/10.25299/jshmic.2024.vol11(1).15835

Royal Society. (2017). *Machine Learning: The Power and Promise of Computers That Learn by Example: An Introduction.* Royal Society.

Tedre, M., Toivonen, T., Kahila, J., Vartiainen, H., Valtonen, T., Jormanainen, I., & Pears, A. (2021). Teaching machine learning in K–12 classroom: Pedagogical and technological trajectories for artificial intelligence education. *IEEE access, 9*, 110558-110572. https://doi.org/10.1109/ACCESS.2021.3097962

Yang, Y., & Kong, S. C. (2025, April). Teaching machine learning for senior primary and junior secondary students: A case study of teacher professional development in Hong Kong. In *Proceedings of the 17th International Conference on Computer Supported Education* - Volume 2: CSEDU, (pp. 35-42). Porto, Portugal. https://doi.org/10.5220/0013200300003932

Yue, M., Jong, M. S. Y., & Dai, Y. (2022). Pedagogical design of K-12 artificial intelligence education: A systematic review. *Sustainability, 14*(23), 15620. https://doi.org/10.3390/su142315620