# The Seamless Integration of Machine Learning Education into High School Mathematics Classrooms

**Hyunju Oh, University of Florida**

Hyunju Oh is a Ph.D. student in School of Teaching & Learning, College of Education, University of Florida. Her research interests include Virtual Learning Environments, Learning Analytics, Artificial Intelligence in Education, and STEM education.

**Rui Guo, University of Florida**

Dr. Rui Guo is an instructional assistant professor of the Department of Engineering Education in the UF Herbert Wertheim College of Engineering. Her research interests include data science & CS education, Fair Artificial Intelligence and Experiential learning.

**Dr. Wanli Xing, University of Florida**

Wanli Xing is the Informatics for Education Associate Professor of Educational Technology at University of Florida. His research interests are artificial intelligence, learning analytics, STEM education and online learning.

**Zifeng Liu, University of Florida**

Zifeng Liu is a Ph.D. student and research assistant in School of Teaching & Learning, College of Education, University of Florida. Her research interests include educational data mining, artificial intelligence, and computer science education.

**Yukyeong Song, University of Florida**

Yukyeong Song is a PhD candidate and research assistant in School of Teaching & Learning, College of Education, University of Florida. Her research focuses on artificial intelligence in education, learning analytics, inclusive AI, computer science, and STEM education.

**Chenglu Li, The University of Utah**

Chenglu Li is an Assistant Professor in Educational Technology and Instructional Design at the University of Utah. He is interested in extending and developing algorithmic and design strategies to promote fair, accountable, and transparent (FAccT) AI in STEM education.

# Work-in-Progress: The Seamless Integration of Machine Learning into High School Mathematics Classrooms

## Introduction

Machine learning (ML) is the backbone of Artificial Intelligence (AI), infusing it with the remarkable ability to identify patterns autonomously and predict future outcomes. Nonetheless, as highlighted by the US Department of Defense's National Defense Strategy 2022 [1] and the National Security Commission on AI 2021 [2], a significant deficit in the AI and ML workforce has emerged as a pressing concern that threatens to hinder the realization of AI's potential. Given ML's growing influence across various sectors, it is increasingly vital for high school curricula to incorporate ML concepts, equipping students with the skills and knowledge essential for navigating and contributing to this rapidly evolving field [3].

However, integrating ML into K-12 education presents numerous challenges. One significant hurdle is the disparities in accessing current CS education, highlighting the urgent need for increased investment in formal education for underrepresented minorities [4]. Furthermore, the current ML education has heavily emphasized operational skills, such as coding in Python or other programming languages, which demands considerable time to grasp programming concepts and syntax [3, 5]. This focus often overshadows algorithmic thinking development, a critical aspect of ML education [6]. To widen student participation and maintain students' interest in ML, introducing early ML experiences for K-12 students by integrating ML into core subjects and providing no-code or low-code options is essential.

Incorporating ML into a high school core subject like mathematics can offer a seamless experience for all students. Understanding the mathematical theories underlying ML is vital for a more profound comprehension of its fundamental principles [7]. By leveraging the advancements in Generative AI (GAI), students can engage with ML by focusing on algorithmic thinking rather than on the intricacies of syntax [8-9].

Our team developed ML4Math, a curriculum that seamlessly integrates ML into high school mathematics education to test this approach. In this curriculum, students learned how machines "see" images by focusing on image classification, using real-life examples, and exploring the mathematical concepts underlying the process and the evaluation of the model. Furthermore, the students engaged in hands-on activity to build an image classifier. GAI was employed as a scaffold to facilitate students' construction of these classifiers.

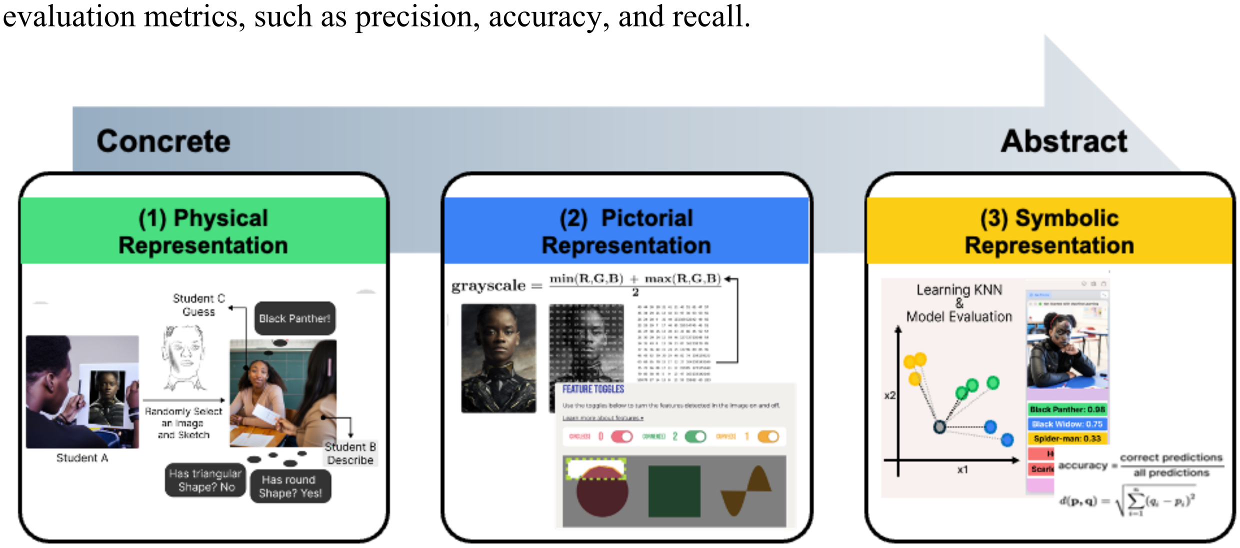
To assess the impact of ML4Math on students, we conducted a two-day pilot study with high school students, evaluating their knowledge through tests and gathering their reactions to the curriculum. In this study, we aim to address three research questions: 1) To what extent does integrating ML into high school mathematics curricula enhance students' understanding of both concepts? 2) What are the experiences of high school students with the ML4Math program? 3) How do students utilize GAI to build their image classifiers in the ML4Math program?

## Methods

### 1) ML4Math program

We developed the ML4Math program using a concreteness-fading approach, which includes three stages: *Physical Representation*, *Pictorial Representation*, and *Symbolic Representation*. The *Physical Representation* phase introduces physical representations linked to abstract symbols, thereby fostering an effective understanding by connecting abstract concepts with imaginary examples. The *Pictorial Representation* phase transitions from physical examples to graphic or iconic diagrams and models. This intermediate phase facilitates a smoother connection between concrete and abstract materials. The final phase, *Symbolic Representation*, introduces abstract concepts, such as numerical calculations and symbolic equations. The concreteness-fading approach facilitates students' understanding of ambiguous mathematical concepts by offering practical contexts [10].

Figure 1 shows the concreteness-fading approach in the ML4Math program. During the *Physical Representation* phase in our program, students explore image classification applications in real-world situations, such as object recognition in self-driving cars and automated photo categorization. Another activity, 'sketch-guessing', involves groups of three students; one student sketches a chosen image, another describes it using geometric concepts, and then the third guesses the original image based on these clues. In the *Pictorial Representation* phase, students explore the conversion process of colors in images to numerical values, exemplified by grayscale transformation. This process exemplifies machines' capability to analyze pixel data. Students also explore how machines identify objects, ranging from low-level features (e.g., edges, dark spots) to high-level features (e.g., facial structure). In the *Symbolic Representation* stage, students learn the K-nearest neighbors algorithm, an ML method for image classification, with a calculation of the Euclidean distance to find the nearest label. Furthermore, students explore the significance of comparing predicted and actual outcomes and calculate model evaluation metrics, such as precision, accuracy, and recall.




**FL Math Standards**
Geometry: Expressing Geometric Properties with Equations – HSG.GPE.B; MA.8.GR.1; MA.912.GR.3
Data Analysis and Probability: Solve problems Involving Categorical Data – HSS.ID.B; MA.912.DP.3

Figure 1. The ML4Math Program with a Concreteness-fading Approach

After the conceptual learning, students form small groups to create a ‘hero image classifier’ in collaboration with the GAI tool Ghostwriter on Replit. To successfully build the image classifier, students must solve five tasks related to writing the necessary code. The GAI tool assists in generating and refining code, as well as detecting and fixing errors. This makes it easier for beginners to start programming and allows students to explore ML concepts without needing extensive programming knowledge. After constructing their classifier, students evaluate the classifier’s performance, discuss potential biases, and explore retraining strategies to enhance accuracy and fairness.

### 2) Pilot Study

We conducted a pilot study with twenty-five high school students in Florida from May 30 to May 31, 2023, to explore their experiences with ML4Math. During this period, two researchers taught all the courses to students. Out of the 25 students, a total of 15 completed the program. Most students (60%) reported that they had 'sometimes' programmed a computer in the past; 20% of students indicated 'most of the time'; 13.3% of students indicated 'about half the time'; and 6.7% indicated 'never.’

The pilot study was set up in four phases: (1) A pre-knowledge test and a pre-survey were conducted to determine students’ baseline knowledge and conceptions; (2) Students learned the image classification using a concreteness-fading approach, integrating high school math concepts into their learning (see Figure 1); (3) Students built their image classifiers using the GAI tool; and (4) at the end of the study, a post-knowledge test and a post-survey were conducted.

Knowledge test questions were designed to assess students’ basic knowledge of ML and their understanding of the mathematical concepts related to ML. This test comprised seven questions, divided into three multiple-choice, two short answers, and two open-ended questions. Multiple-choice questions asked students to demonstrate their understanding of computer vision, ML model validation methodologies, and the importance of image classification. Two of the short-answer questions asked students to calculate the Euclidean distance between two points and determine a model’s accuracy. The two other questions required explanations of the causes of bias in ML models and the fairness of model evaluations.

In addition to the knowledge test, students completed a 5-point Likert scale survey before and after the program. This survey was designed to compare their self-efficacy in learning math by learning AI and using GAI technologies. The post-survey included additional questions focusing on engagement and self-efficacy of the program to elicit feedback from students. We adapted survey items from previous articles [11-12] after we modified them to fit our context.

## Results

### 1) Knowledge test results

We analyzed the data using SPSS 29.0, employing a paired t-test to examine the difference between the pre-knowledge and post-knowledge test scores. The t-test was chosen as the data followed a normal distribution. The significance values of the Kolmogorov-Smirnova and the Shapiro-Wilk tests were above 0.05, indicating a normal distribution. Table 1 reveals that the mean score for the post-knowledge test (Mean = 3.77, SD = 1.739) was significantly higher than the pre-knowledge test (Mean = 3.00, SD = 1.080).

The study revealed that students have developed an understanding of bias in ML models. In the pre-test, more than a third of the students answered "*I don't know*" in response to the open-ended question 'What is the main reason for bias in machine learning models?' However, in the post-survey, some students provided specific reasons, such as "*Incorrect evaluations based on improper data, which can be resolved by closing the errors and making it more accurate, training it on the model*." While others believed "*The programmers who code it, the data sets it is using, etc.*" Moreover, students demonstrated the ability to assess whether a model is biased. When asked to 'Examine the results below, and answer whether the model is fair or not, and why?' students who initially answered "*I don't know*" in the pre-test, after completing the ML4Math program, identified bias, saying, "*No, as the model predicts one data set correctly, while consistently predicting another data set incorrectly.*"

Table 1. The t-test on the Pre-Post Knowledge Test (n = 13)

| | Mean | SD | t | df | Cohen's d | P-value |
|---|---|---|---|---|---|---|
| (pre) Knowledge test | 3.00 | 1.080 | -2.739 | 12 | 1.013 | .018* |
| (post) Knowledge test | 3.77 | 1.739 | | | | |

*p<.05

### 2) Survey results

As shown in Table 2, the Wilcoxon signed rank test was utilized to analyze the survey results concerning students' self-efficacy. The analysis revealed a significant improvement in students' self-efficacy in using GAI, as compared to the pre-survey results (D; $p < .05$). Although not statistically significant, the ML4Math program appeared to enhance students' confidence in learning ML with mathematics (B) and using GAI for learning purposes (C). The post-only survey results in Table 3 also indicated that the ML4Math program effectively incorporates ML into math education. 60% of the participants agreed that they enjoyed learning math and AI together, with 6.7% expressing disagreement (E). Furthermore, 60% of the students were motivated to learn more about AI, whereas 6.7% disagreed with this sentiment (I).

### 3) Building image classifier using GAI

In this research, students utilized the GAI functions for generating, improving, and debugging code while developing an image classifier. We observed that students completed the unfinished segments of the image classifier code by requesting specific code snippets following the guidelines in the to-do list. An example of such a request is, "In the retrain function, call the fit

Table 2. The Wilcoxon Signed Rank Test on the Self-efficacy Pre-post Surveys (n = 13)

| Question | Mean (SD) | | Statistic Value | |
|---|---|---|---|---|
| | Pre | Post | Z | Sig |
| A. I can learn math by learning AI. | 3.77 (0.60) | 3.77 (0.83) | .000 | 1.000 |
| B. I would like to learn math in school through learning AI. | 3.31 (0.95) | 3.38 (0.87) | -.333 | .739 |
| C. I feel confident using the code generation AI (e.g., Replit). | 3.23 (0.93) | 3.38 (1.19) | -.702 | .483 |
| D. I know what information is needed to use code generation AIs (e.g., Replit) effectively. | 2.69 (1.03) | 3.54 (0.97) | -2.299 | .022* |

* p <.05

Table 3. Post-only survey on Engagement and Self-efficacy (n = 15)

| Question | Mean (SD) | Disagreement | Agreement |
|---|---|---|---|
| E. I enjoyed learning math and AI together. | 3.73 (0.88) | 1 (6.7%) | 9 (60%) |
| F. Using the code generation AI (Replit) in the course enhances the effectiveness of programming. | 3.40 (1.18) | 2 (13.3%) | 8 (53.3%) |
| G. I would recommend this course to people I know as a great resource to learn math and AI. | 3.60 (0.74) | 0 (0%) | 7 (46.7%) |
| H. Learning in this course gives me a feeling of personal accomplishment. | 3.27 (0.96) | 2 (13.3%) | 6 (40%) |
| I. This course motivates me to learn more about AI. | 3.67 (0.82) | 1 (6.7%) | 9 (60%) |
| J. This course motivates me to learn more about Math. | 3.00 (0.76) | 2 (13.3%) | 3 (20%) |

E - G: Engagement; H- J: Self-efficacy

function of this class to retrain the model." Despite having these commands at their disposal, students still encountered challenges, particularly with errors often stemming from indentation issues and inconsistent responses from the GAI. When such errors occurred, students utilized the GAI function for troubleshooting. They asked it to review the script, copy and paste either the entire script or the specific parts where the error manifested and the error messages themselves to debug the problem effectively.

**Discussion and Conclusion**

This study highlights the potential benefits of integrating ML with math in high school education, particularly in improving learning performance and self-efficacy and making the learning process engaging. The ML4Math program based on the concreteness-fading approach allowed students to engage with ML from fundamental concepts to practical applications [13]. This approach not only enriches their learning experience but also has the potential to spark a deeper interest in learning about AI in the future. GAI has played a pivotal role in making programming more accessible to beginners by lowering technical barriers [14]. Despite most students reporting minimal prior experience with programming, they were able to contribute to developing the image classifier by utilizing the GAI function. Yet, it's important to acknowledge that GAI can exhibit inconsistencies, like providing varied responses to the same prompts. Also, students had to understand basic programming principles, including proper indentation techniques.

The study's limitations, primarily due to the small number of participants, suggest that further study with a larger sample size is needed to confirm and generalize these findings. This additional research would help assess the impact and scalability of integrating ML into high school mathematics education.

## References

[1] U.S. Department of Defense. (2022). *2022 National Defense Strategy*. https://media.defense.gov/2022/Oct/27/2003103845/-1/-1/1/2022-NATIONAL-DEFENSE-STRATEGY-NPR-MDR.PDF

[2] National Security Commission on Artificial Intelligence. (n.d.). *National security commission on artificial intelligence final report*. https://reports.nscai.gov/final-report/introduction

[3] Martins, R. M., & Gresse Von Wangenheim, C. (2023). Findings on teaching machine learning in high school: A ten-year systematic literature review. *Informatics in Education*, *22*(3), 421-440.

[4] Barretto, D., LaChance, J., Burton, E., & Liao, S. N. (2021, June). Exploring why underrepresented students are less likely to study machine learning and artificial intelligence. In *Proceedings of the 26th ACM Conference on Innovation and Technology in Computer Science Education V. 1* (pp. 457-463).

[5] Chua, H. X., Chua, E. L., & Soo, K. (2019). Budding data scientists hackathon. *ACM SIGKDD Explorations Newsletter*, *21*(1), 38-40.

[6] Kirçali, A. Ç., & Özdener, N. (2023). A comparison of plugged and unplugged tools in teaching algorithms at the K-12 level for computational thinking skills. *Technology, Knowledge and Learning*, *28*(4), 1485-1513.

[7] Dara, D. S., Surmara, S., Reddy, S. K., Vani, K., & Varma, A. S. (2022). Role of Mathematics in Machine Learning. *IRJMETS*, *4*(04), 2543-2548.

[8] Denny, P., Prather, J., Becker, B. A., Finnie-Ansley, J., Hellas, A., Leinonen, J., Luxton-Reilly, A., Reeves, B. N., Santos, E. A., & Sarsa, S. (2024). Computing education in the era of generative AI. *Communications of the ACM, 6*7(2), 56-67.

[9] Porter, L. (2024). *Learn AI-assisted Python programming: With GitHub Copilot and ChatGPT*. Simon & Schuster.

[10] Fyfe, E. R., McNeil, N. M., Son, J. Y., & Goldstone, R. L. (2014). Concreteness fading in mathematics and science instruction: A systematic review. *Educational Psychology Review, 26*(1), 9-25. 10.1007/s10648-014-9249-3

[11] May, D. K. (2009). *Mathematics self-efficacy and anxiety questionnaire* (Doctoral dissertation). University of Georgia.

[12] Huang, C. Y., & Wang, J. C. (2022). Effectiveness of a three-dimensional-printing curriculum: Developing and evaluating an elementary school design-oriented model course. *Computers & Education*, *187*, 104553.

[13] Xing, W., Huang, X., Li, C., & Xie, C. (2023). Teaching thermodynamics with augmented interaction and learning analytics. *Computers & Education, 196*, 104726.

[14] Sun, J., Liao, Q. V., Müller, M., Agarwal, M., Houde, S., Talamadupula, K., & Weisz, J. D. (2022, March). Investigating explainability of generative AI for code through scenario-based design. In *Proceedings of the 27th International Conference on Intelligent User Interfaces* (pp. 212-228).