Introducing Machine Learning with Scratch and Robots as a Pilot Program for K-12 Computer Science Education

Chan-Jin Chung
Lawrence Technological University/Department of Mathematics and Computer Science, Southfield, Michigan, USA
Email: cchung@ltu.edu

Lior Shamir
Kansas State University/Department of Computer Science, Manhattan, Kansas, USA
Email: lshamir@ksu.edu

***Abstract*—Machine learning (ML), a branch of artificial intelligence (AI), is a method that enables systems to learn from data for the purpose of recognizing patterns and making decisions without being explicitly programmed. In the past decade machine learning has been growing rapidly, and ML technologies such as speech recognition, spam filters, smart email reply, online recommendations, face recognition, fake news detection, and self-driving cars have become pivotal in modern daily life. However, computer science education has not yet fully adjusted to the tremendous growth in the sub-field of AI. This paper describes an approach of introducing K-12 students to ML through an on-line summer camp. The students are introduced to the concept of ML by hands-on activities of developing applications for recognizing text, numbers, sounds, images, and video data using a web-based cloud service tool "Machine Learning for Kids" and Scratch 3 programming language combined with Lego Mindstorms EV3 robots. The results show that the tools and technologies used in the camp are suitable for K-12 students, also when used in the form of online training. Pre and post surveys show that students express basic knowledge in ML and higher interest in coding and STEM after being exposed to the proposed training.**

***Index Terms*—artificial intelligence, computational thinking, computer science education, k-12 STEM education, online learning, machine learning, robots, scratch coding**

I. Introduction

Computer Science (CS), Computational Thinking (CT), and coding skills are gradually becoming part of the K-12 school curricula worldwide [1]. According to the Brookings Institute, out of 219 countries, 44 (around 20 percent) mandate that schools offer it as an elective or required course, 15 (around 7 percent) offer CS in select schools and some subnational jurisdictions (states, provinces, etc.) [2]. Multiple studies suggest that CS and CT education can help students beyond just computing, and CS education has been linked with improved problem-solving skills [3] and higher rates of college enrollments [4].

In the past decade, AI, and in particular machine learning, have been expanding very rapidly, and became one of the fastest growing sub-fields of computer science, with direct impact on billions of people [5]. Although AI and ML are mature fields of study, the relatively sudden and sharp increase in their impact on CS and society has been introducing new gaps in all levels of computer science education [6, 7].

While substantial attention has been given to undergraduate and graduate education in AI [7], the pivotal nature of AI in computer science and society introduces the need to also expose K-12 students to AI [8]. Exposing K-12 students to AI can provide them with AI literacy, and potentially attract them to academic education and consequently a career in AI. Therefore, it has been proposed that AI should be studied at an early age [9, 10].

An early example of efforts to develop educational experiences in AI was based on the development of expert systems [11]. More recent attempts were based on Google Teachable Machine [8], the Snap! block programming language [6], or simple programming environment such as HTML and JavaScript suitable for high school students [12].

Here we propose an online machine learning K-12 educational experience based on Machine Learning for Kids (ML4K) [13] to introduce K-12 students to AI/ML, as well as CS, CT, and coding skills with robots [14].

II. Educational Outcomes

The proposed educational experience makes use of multimedia data analysis to engage K-12 students in ML. The design is driven by the assumption that the use of multimedia data such as video can provide a broader variety of options for hands-on projects that can be of interest to K-12 students.

The primary expected outcome of the educational experience is providing K-12 students with ML and AI literacy. Additionally, it also provides the students with

Presented online in 3rd International Conference on Future Learning (ICFL 2020), December 17-19, 2020, London, United Kingdom

training in basic programming through the activities involved in implementing the ML tasks. The development activities involve basic practicing of logic and problem-solving skills. Other educational outcomes include collaboration among the students, implemented through data collection, verification, and testing.

A primary goal of the activities is to enhance the self-efficacy of the students, and increase their interest in CS, and particularly in ML/AI. Such experience can encourage more students to expand their CS education through academic studies, consequently leading to careers in CS, ML, AI, or other STEM fields.

III. Research Methodology

The research methodology aims at providing new knowledge regarding the following questions:

(1) How feasible is the "ML for Kids" environment as a tool for providing online-based training in data collection, training ML models, and using the models to create apps in Scratch.

(2) Whether the introduction of ML with robots leads students to have more positive view of CS and STEM subjects.

(3) Whether the experience of ML with robots increases students' interest in CS and STEM careers.

(4) Whether the hands-on experience of ML with robots improves students' learning in CS and STEM.

(5) What technical and pedagogical issues must be considered when designing CS curriculum with ML tools such as ML4K?

A. Tools

Tools considered for the camp were "Teachable Machine" [15] introduced by Google Creative Labs in 2018 and web-based Machine Learning for Kids (ML4K) cloud service [13] built using IBM Watson APIs in 2017. The latter was selected for this experiment due to its ability to provide an easy-to-use guided environment with Scratch 3 coding platform. Scratch 3 provides multiple extensions for robotics such as Lego Mindstorms EV3, Lego WeDo 2.0, and micro:bit.

Scratch [16, 17] is a visual programming environment that allows users to learn computer programming while working on media-rich projects such as interactive stories, games, and animations. Scratch is designed especially for ages eight to 16, but is used by people of all ages. A key design goal of Scratch is to support self-directed learning through tinkering and collaboration with peers. Images and sounds can be imported or created in Scratch using a built-in paint tool, webcam, and sound recorder. Programming is done by snapping together colorful command blocks to control 2D graphical objects called "sprites", moving on a background called the "stage". Any tablets or mobile devices can also be used to program in Scratch 3 through the device's web browser.

The camp was delivered using Zoom, the online conferencing software. Before the camp, instructions to prepare things such as robot sensors and motors before the meeting were sent to adults (teachers, robotics team coach, or parents) who registered K-12 students online. User accounts and passwords for the ML4K were also sent to the adults before the meeting day. In the beginning of the camp, presentation slides were provided to the participants. After the camp, recorded presentations were sent to the participants.

B. Research Participants

We offered two identical introductory level camps for three online contact hours on July 10 and July 24, 2020. The second workshop took place August 14th, lasted two hour and 15 minutes, and was focused on video image data applications.

We recruited students using the Robofest network. Robofest is an autonomous robotics competition primarily for 4th through 12 grade students focusing on learning computer science and programming. Any robotics kits were allowed in the construction of robots. Robots can be programmed with any programming language in Robofest [18, 19].

Twenty-four students were enrolled for the 1st camp on July 10th. Average grade was 7.7. 20 students were registered for the July 24 camp, where average grade was 7.6. 29 students were enrolled for Aug 10th camp, where average grade was 7.7. As a whole the total of 73 students were registered and the average school grade was7.7. However, 10 students did not show up to the Zoom meetings, and therefore the actual number of students who participated in the camps was 63.

C. Data Collection Methods

A pre-survey link was included in the preparation instruction file sent prior to the camp meeting day to coaches. The anonymous pre-survey had four questions as shown in Figure 1.

A post-survey link was sent directly to the students at the end of each camp day using the Zoom chat function. The anonymous post survey had exactly the same four questions as the pre-survey, as well as one additional new question as shown in Figure 2.

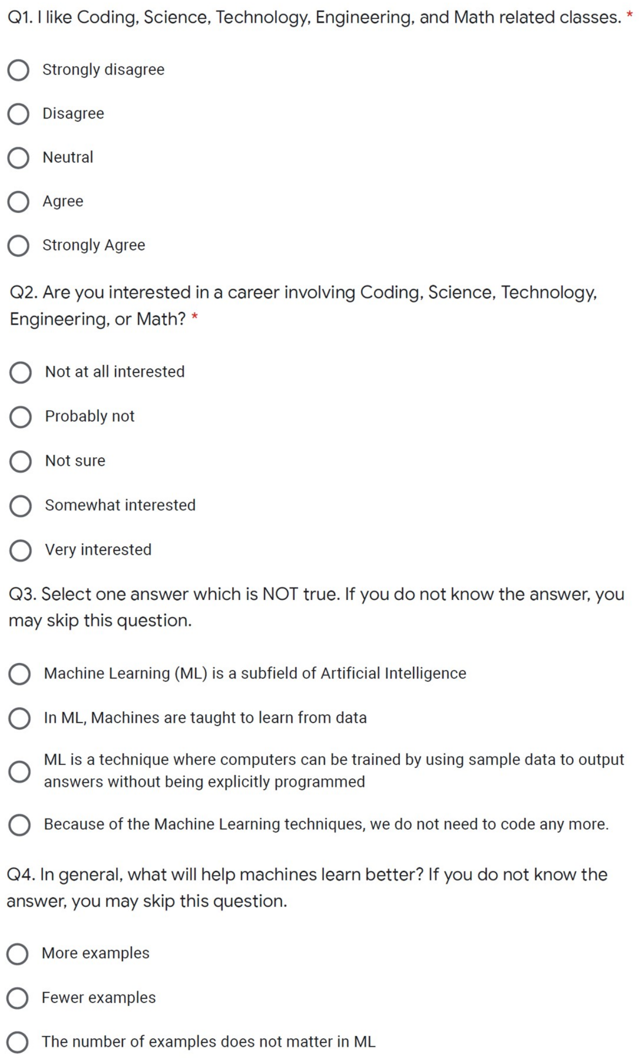
IV. In Class Hands-on and Minds-on Projects

Figure 3 shows the Lego robot setup requested to prepare the camp. One color sensor, one distance sensor, and a Lego EV3 motor were used. EV3 intelligent brick is connected to the computer via Bluetooth connection.

During the camp, after explaining concepts and showing sample solutions for the following projects, the instructor allowed some time to complete the project during the session. Students reported completion through the Zoom Chat.

A. Types of Projects

The projects included a variety of tasks that involve different types of data, demonstrating a relatively broad range of options that can engage a variety of students.


Q1. I like Coding, Science, Technology, Engineering, and Math related classes. *

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

Q2. Are you interested in a career involving Coding, Science, Technology, Engineering, or Math? *

- Not at all interested
- Probably not
- Not sure
- Somewhat interested
- Very interested

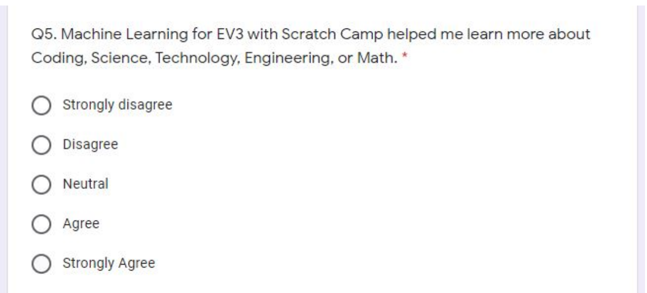
Q3. Select one answer which is NOT true. If you do not know the answer, you may skip this question.

- Machine Learning (ML) is a subfield of Artificial Intelligence
- In ML, Machines are taught to learn from data
- ML is a technique where computers can be trained by using sample data to output answers without being explicitly programmed
- Because of the Machine Learning techniques, we do not need to code any more.

Q4. In general, what will help machines learn better? If you do not know the answer, you may skip this question.

- More examples
- Fewer examples
- The number of examples does not matter in ML

Figure 1. Pre-survey questions


Q5. Machine Learning for EV3 with Scratch Camp helped me learn more about Coding, Science, Technology, Engineering, or Math. *

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

Figure 2. Post-survey questions

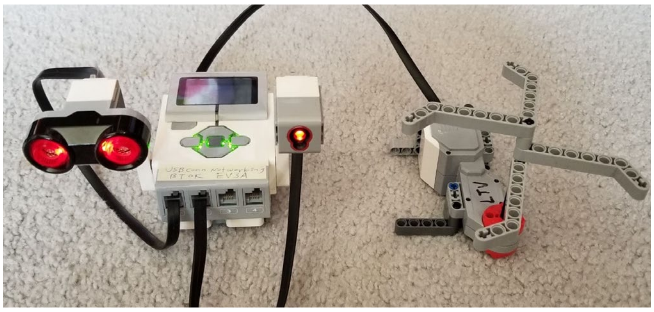


Figure 3. EV3 robot setup for the camp

Recognizing numbers. The purpose of the project is to create a simple model to recognize whether a number representing the value from an ambient light sensor is dark or bright. If dark, the color of the Scratch stage backdrop should be changed to yellow.

Text recognition. The students need to train a model with possible texts that can be used to turn on/off fans such as “Turn on the fan”, “It is hot here”, “Can we turn off the fan?”, “Please turn the fan off”, and “It’s cold”. The students are asked to write a Scratch code to turn on the Lego motor based on keyboard text entered using the trained model.

Recognizing sounds. In this project the students are asked to collect recordings of your own voice when you say “on” and “off” and train a model to recognize them. For sound projects, students are also required to collect examples of background noise to help the computer to learn to recognize your sounds. Then, the students write a code to turn the Lego motor fan on or off according to voice command using the trained model.

Recognizing Images. The students collect images of “Rock” and “Paper” by showing their hand to the webcam. Then, the students train them to recognize when their hand makes either “Rock” or “Paper”. The students then create costumes of the sprite to use as the image inputs of the model by showing their hands to the webcam.

Recognizing two Lego sensor data. This project involves the recognition of two number values problem to turn on the Lego motor fan only when the light sensor detects dark and the distance sensor detects a near object. A Scratch code must collect data through a light sensor and a distance sensor to train a model.

Color your nose red. The students use a pre-trained machine learning model to detect a face on a live webcam video, and code animated effects on their nose.

Alarm if no mask detected. The students first collect images of people wearing face masks, as well as images of people not wearing masks. The students then use the images to train a model to recognize whether a person on an image wears a mask or not. No code was required.

Alarm if the user is not wearing a mask. The students collect images of themselves wearing face masks, as well as images of themselves when not wearing a mask using the webcam function in ML4K. After training the model, the code should recognize whether the student is wearing a mask or not in real-time. Figure 4 describes blocks and process of this project.

Pose detection – Wrist Up or Down. The students use a pre-trained ML model provided by ML4K to detect the location of the right-hand wrist. Students then create and train their own ML model to recognize whether their wrist position is up or down. The app should make a special sound only when the right hand is up. Figure 5 shows the high-level blocks of the project for training and inference processes. Figure 6 shows two scripts that collects Wrist-up images and wrist-down images.

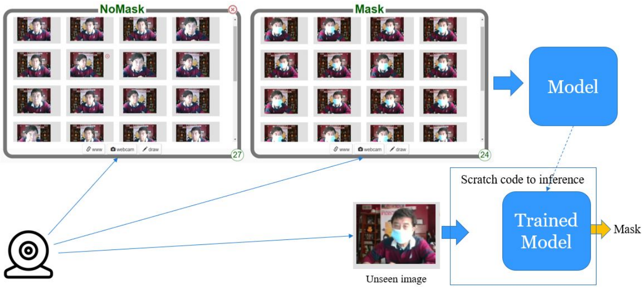


Figure 4. Project to alarm if no mask is detected

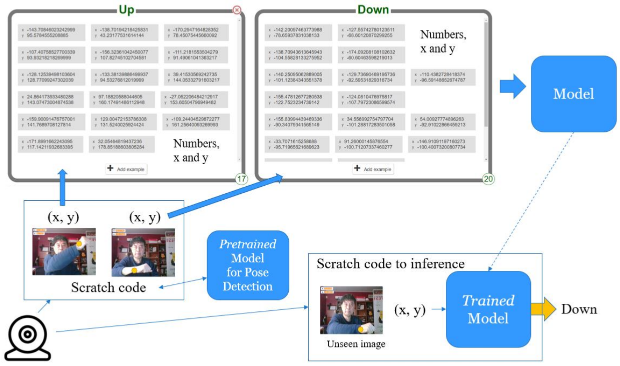


Figure 5. Project to detect wright wrist is up or down

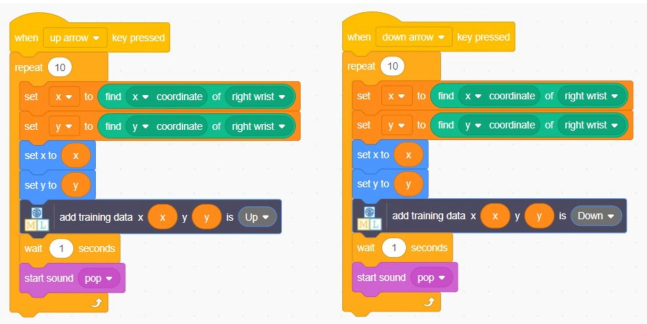


Figure 6. Scratch scripts to collect wrist up and down images

Figure 7 shows the code that trains the model, get (x, y) positions of the wrist on the current image, move the Sprite to that location, and then ask the local trained model if the wrist is up by sending the current (x, y) position. If that is true, you will hear a pop sound.

B. *Mini Project Competition*

Robofest defines a “camp” as “hands-on workshops + Mini project competitions” [18]. It is a teaching method for STEM+CS education in the context of constructivist learning theory [20]. The camp was delivered in an online format, and the students were asked to submit source code files as well as video links of their own mini projects for prizes within 60 hours, after the meeting.

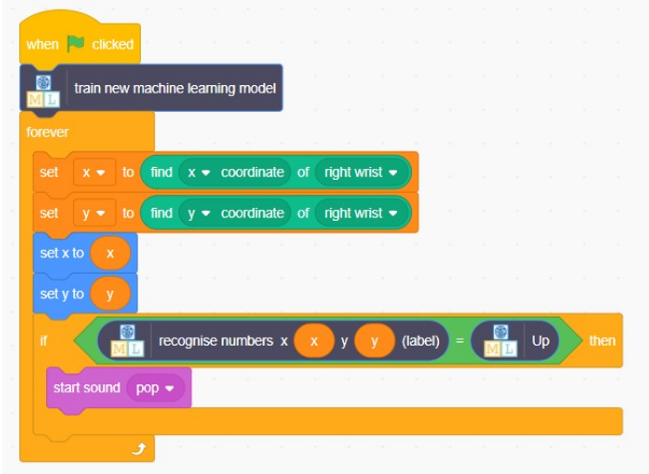


Figure 7. Scratch code to inference if the wrist is up

V. Results

According to the Zoom chat log, approximately 40% of the students completed the in-class projects. Students who did not complete their projects were asked to complete their projects after the camp. However, because the ML4K account only saves training data and trained models, code for projects completed after the sessions could not be collected.

A total of 41 students participated in the pre-survey and 45 students participated in the post-survey. The survey had two questions directly related to basic knowledge in ML concepts. For Question 3 in Figure 1, 91% of students answered immediately after taking the camp. Fifty eight percent of the students answered correctly the same question in the pre-survey. For Question 4 in Figure 1, 78% of the students answered correctly in the pre-survey, while 99% answered correctly in the post survey.

The percentage of students indicating that they like STEM/Coding increased from 80.5% in the pre-survey to 91.1% in the post-survey. Figure 8 shows the result of pre and post surveys for Question 1. Using t-test statistics, the probability to have such difference by mere chance is (P≃0.016).

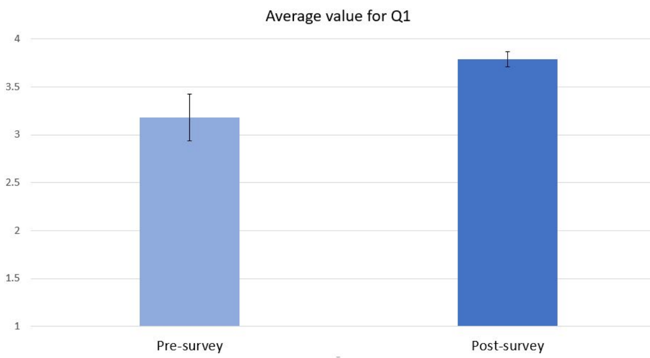


Figure 8. The result of pre and post surveys for Question 1. The error bars are the standard error of the means.

In the post survey, 100% of the students also expressed that they would consider a career involving Coding, Science, Technology, Engineering, or Math. The change was from 95.1%. Figure 9 shows the result of pre and post survey for Question 2 about STEM career plan. The probability of such difference to occur by chance is (P 0.043).

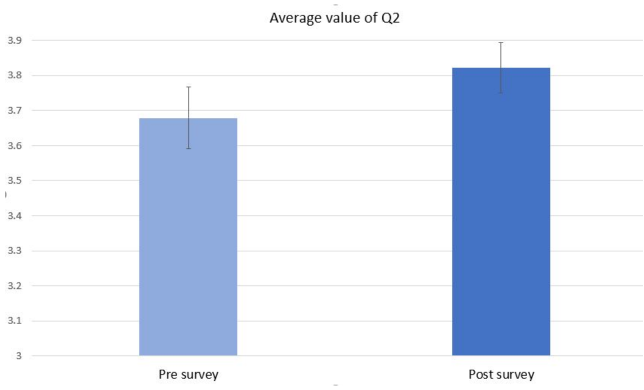


Figure 9. The result of pre and post surveys for Question 2.

A majority (44.4+37.8=82.2%) of students indicated that the ML camp experience helped them learn more about Coding, Science, Technology, Engineering, or Math as in Figure 10.

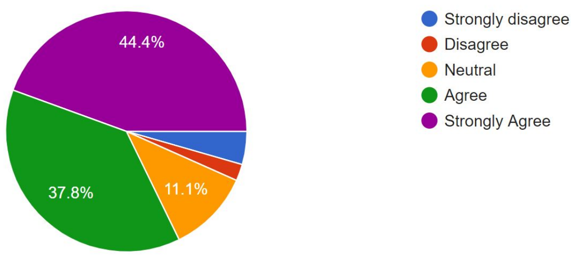


Figure 10. The result of post survey for Q5: This camp helped me learn more about Coding/STEM.

Regarding the home work mini projects, a total of 9 projects were submitted. Selected projects for prizes are as follows:

- Smart Mask Detector: https://youtu.be/kkpFaDUhXj8
- Smart Shopper: https://youtu.be/CmnwfPzYXIE
- Recognizing handwriting to turn on/off a Lego robot fan: https://youtu.be/UEgeDiKJ6rc
- Mood Recognition: https://youtu.be/LH5pInLapMo
- Counting the number of sit-ups: https://youtu.be/RVTMIQs_MWI
- Flappy bird game: https://youtu.be/kvJSJuzONsk
- Wear a mask: https://youtu.be/-bedMApwKA8

VI. Conclusion

AI and ML technologies have been growing significantly over the past decade, impacting billions of people in virtually all aspects of modern lives, and the trend of increasing impact of AI is bound to continue. The growing importance of AI and ML in computer science and society reinforces the introduction of AI and ML concepts to students at an early stage.

Here we described an online educational experience designed for K-12 students, based on Machine Learning for Kids (ML4K)" system with Scratch and Lego EV3 robots. The preliminary results show that using ML4K in combination with Lego EV3 is feasible, engages K-12 students, and leads to positive attitude of the participants towards ML, AI, CS, and STEM in general.

Based on the surveys, projects completed, and account data, it is clear that the use of ML4K environment is suitable for teaching the concept of ML so that students can create apps using the models they created and trained. Also, it is shown that the introduction of ML with robots improves students' likeness in CS and STEM subjects. The experience of ML with robots also increases students' interest in CS and STEM careers, and the hands-on experience of ML with robots improves students' learning.

Observations of communication with the students suggested that students working on projects related to multimedia content such as video, images, sound, and natural language data were more motivated compared to the other students.

The ML4K environment allows students to collect and verify training data examples through collaboration with each other. Future work will include design that better enables collaboration between students through an online platform.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Chan-Jin Chung organized the summer camps, developed curriculum, taught the online camps, and collected the data; Lior Shamir assisted with the camps, analyzed the data, and wrote the paper; Both authors conducted the research and approved the final version.

Acknowledgment

This work was supported in part by a sponsorship from DENSO, Southfield, Michigan, USA. The camps were hosted by Lawrence Technological University's Robofest® (https://www.robofest.net) program.

References

[1] Amanda Sullivan and Marina Bers. 2019. Computer science education in early childhood: the case of Scratch Jr. Journal of Information Technology Education: Innovations in Practice 18, 1 (2019), 113–138.

[2] Emiliana Vegas and Brian Fowler. 2020. What do we know about the expansion of K-12 computer science education? A review of

the evidence. Available: https://www.brookings.edu/research/what-do-we-knowabout-the-expansion-of-k-12-computer-science-education/
[3] Ruqayya Toorawa Shima Salehi, Karen D. Wang and Carl Wieman. 2020. Can Majoring in Computer Science Improve General Problem-solving Skills?. In SIGCSE '20: Proceedings of the 51st ACM Technical Symposium on Computer Science Education. ACM, 156–161.
[4] Emily Anne Brown and Richard S. Brown. 2020. The Effect of Advanced Placement Computer Science Course Taking on College Enrollment. Available: http://www.westcoastanalytics.com/uploads/6/9/6/7/69675515/ longitudinal_study_-_combined_report_final_3_10_20__jgq_.pdf
[5] Giang Nguyen, Stefan Dlugolinsky, Martin Bobák, Viet Tran, Álvaro López García, Ignacio Heredia, Peter Malík, and Ladislav Hluch`y. 2019. Machine Learning and Deep Learning frameworks and libraries for large-scale data mining: a survey. Artificial Intelligence Review 52, 1 (2019), 77–124.
[6] K Megasari Kahn, Rani Megasari, Erna Piantari, and Enjun Junaeti. 2018. AI programming by children using snap! block programming in a developing country. 13th European Conference on Technology Enhanced Learning (2018).
[7] Weidong Liao and Osman Guzide. 2019. Teaching artificial intelligence for undergraduate students: a project based approach. Journal of Computing Sciences in Colleges 34, 3 (2019), 132–132.
[8] Tapani Toivonen, Ilkka Jormanainen, Juho Kahila, Matti Tedre, Teemu Valtonen, and Henriikka Vartiainen. 2020. Co-Designing Machine Learning Apps in K–12 With Primary School Children. In 2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT). IEEE, 308–310.
[9] David Touretzky, Christina Gardner-McCune, Fred Martin, and Deborah Seehorn. 2019. Envisioning AI for K-12: What should every child know about AI?. In Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 33. 9795–9799.
[10] Randi Williams, Hae Won Park, Lauren Oh, and Cynthia Breazeal. 2019. Popbots: Designing an artificial intelligence curriculum for early childhood education. In Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 33. 9729–9736.
[11] CF Reynolds. 1988. Introducing expert systems to pupils. Journal of Computer Assisted Learning 4, 2 (1988), 79–92.
[12] Radu Mariescu-Istodor and Ilkka Jormanainen. 2019. Machine Learning for High School Students. In Proceedings of the 19th Koli Calling International Conference on Computing Education Research. 1–9.
[13] D Lane. 2018. Machine learning for kids. Available: https://machinelearningforkids.co.uk/
[14] Kaiqin Yang, Xin Liu, and Guang Chen, "Global Research Trends in Robot Education in 2009-2019: A Bibliometric Analysis," International Journal of Information and Education Technology vol. 10, no. 6, pp. 476-481, 2020.
[15] 2018. Teachable Machine. Available: https://teachablemachine.withgoogle.com/
[16] John Maloney, Mitchel Resnick, Natalie Rusk, Brian Silverman, and Evelyn Eastmond. 2010. The scratch programming language and environment. ACM Transactions on Computing Education (TOCE) 10, 4 (2010), 1–15.
[17] Mitchel Resnick, John Maloney, Andrés Monroy-Hernández, Natalie Rusk, Evelyn Eastmond, Karen Brennan, Amon Millner, Eric Rosenbaum, Jay Silver, Brian Silverman, et al. 2009. Scratch: programming for all. Commun. ACM 52, 11 (2009), 60–67.
[18] ChanJin Chung, Christopher Cartwright, and Joe DeRose. 2017. Robotics festival and competitions designed for STEM+ C education. In Robotics in STEM education. Springer, 131–170.
[19] CJ ChanJin Chung, Christopher Cartwright, and Matthew Cole. 2014. Assessing the impact of an autonomous robotics competition for STEM education. Journal of STEM Education: Innovations and Research 15, 2 (2014).
[20] Andreas Zendler, "Teaching Methods for Computer Science Education in the Context of Significant Learning Theories," International Journal of Information and Education Technology vol. 9, no. 7, pp. 470-476, 2019.

Chan-Jin Chung attended HongIk University in Seoul, Korea, where he earned a B.S. Computer Science degree. While he was working for Electronics and Telecommunication Research Institute (ETRI) as a research scientist from 1982 to 1992, he was involved in developing TDX switching systems that became later the base system for the first commercialized CDMA system in the world. Chung also worked as a visiting researcher to develop telecommunication software modules for L.M. Ericsson in Sweden in 1983-1984. He received his Ph.D. in Computer Science from Wayne State University in 1997. His doctoral research was the development of a self-adaptive AI system motivated by cultural evolution process, which was then applied to solve various optimization problems. He was an advisor of LTU team who won 1st place award in 3D design optimization competition sponsored by HONDA R&D Europe GmbH as a part of World Congress in Computational Intelligence Conference in 2002. His research interests include evolutionary computation, cultural algorithms, evolutionary-neuro-fuzzy algorithms, deep neural network learning, evolutionary robotics, and robotics in education. Prof. Chung founded a world-wide autonomous robot competition called Robofest (www.robofest.net). Over 28,000 students from 16 US States and 28 countries have participated in the competition since 1999. He launched numerous STEM+CS education programs such as RoboParade, a parade of autonomous vehicles in 2006, Vision Centric robot Challenge in 2007, and CS+PA2: Learning Computer Science with Physical Activities and Animation. He is a faculty advisor of LTU's IGVC (Intelligent Ground Vehicle Competition) teams since 2003. His H2Bot team won 1st place design award in 2007. His team was also selected to represent USA to compete at RoboCup Four-legged robot soccer division in 2007. BigBoot II team won Grand Award LESCOE Cup in 2016. He served as the USA National Organizer for World Robot Olympiad in 2014 and 2015. Currently, he leads ACTor (Autonomous Campus TranspORt) project using a drive-by-wire electric vehicle. The ACTor vehicle team won the Self-Drive Challenge competition at the IGVC in 2017 and 2018. In 2011, IEEE USA honored Dr. Chung with its citation of honor award for his leadership in STEM education.

Lior Shamir Graduated with a Bachelor's in computer science from Tel Aviv University in 1998, M.S. in computer science from Open University in 2003, and Ph.D in computational science and engineering from Michigan Technological University in 2006.

He was research fellow at the National Institute on Aging, Assistant and Associate Professor of computer science at Lawrence Technological University, and Associate Professor and Nick Chong Research Faculty Scholar at Kansas State University, department of computer science. Dr. Shamir authored over 100 peer-reviewed publications, and a member of numerous scientific collaborations such as the Midwest Big Data Hub (MBDH), the Vera Rubin Observatory (formerly LSST), the Astrophysics Source Code Library (ASCL), and more.