Hellenic Journal of STEM Education, 1(2), 59-66, 2020

Artificial Intelligence and Machine Learning in the STEAM classroom: Analysis of performance data and reflections of international high school students

Antonios Karampelas
American Community Schools (ACS) Athens, Greece
karampelasa@acs.gr

Abstract
This research aims in shedding light on international high school students' perceptions, awareness, and prior knowledge about Artificial Intelligence and Machine Learning, as well as to investigate the effect of taking a relevant learning unit within a STEAM course on students' dispositions toward and understanding about Artificial intelligence. The analysis of performance and reflection data from 62 individuals revealed low prior student engagement with Artificial Intelligence and Machine Learning content, a positive shift in the anticipated societal impact of Artificial Intelligence and an active engagement with online Artificial Intelligence applications during the unit, as well as no correlation between student performance and gender. It is suggested that the development and implementation of learning designs that focus on conceptual understanding of Artificial Intelligence and Machine Learning could benefit all students.

Keywords
Artificial intelligence education, machine learning education, high school, minimum spanning trees, artificial neural networks

◆Received 20 January 2021 ◆Accepted 30 June 2021

Introduction
The Fourth Industrial Revolution (Schwab, 2016) will transform societies and the global economy. Equipping students with knowledge and skills that are relevant in a cyber-physical professional and societal landscape is key for them to thrive in this uncharted environment, with a relevant initiative being the Education 4.0 Framework (World Economic Forum 2020). As Artificial Intelligence (AI) will probably be the most important technology to power this disruption, it comes naturally that AI should be considered for inclusion into K-12 curricula by educational heads and policy makers worldwide. Indicative initiatives include teaching AI in Chinese schools (Synced 2018; Jing 2018), and the development of AI curriculum guidelines by the AI4K12 collaboration (Touretzky et al. 2019) between the Association for the Advancement of Artificial Intelligence (AAAI) and the Computer Science Teachers Association (CSTA). In the latter, five big ideas are proposed: 1) "Computers perceive the world using sensors", 2) "Agents maintain models/representations of the world and use them for reasoning", 3) "Computers can learn from data", 4) "Making agents interact comfortably with humans is a substantial challenge for AI developers", and 5) "AI applications can impact society in both positive and negative ways". Karampelas (2020) suggests high school Artificial Intelligence blended and online learning experiences under three units, progressing from the broader to the narrower in scope: "Impact of Artificial Intelligence", "Machine Perception", and "Machine Learning". The integration of AI into K-12 curricula is still in its nascency, therefore relevant research and reflections are not readily available to educators and school administrators.

This article aims in contributing into a better understanding of Artificial Intelligence teaching and learning prospects for (but not limited to) the high school classroom through the analysis of performance data and reflections of international high school students of the American Community Schools (ACS) Athens. Data collection refers to an AI unit of a STEAM (Science, Technology, Engineering, Art, and Mathematics) course (Karampelas, 2021), offered for technology credits. Apart from AI, course topics and concepts include the Fourth Industrial Revolution, the Internet of Things, Electrical Engineering, Aerospace Engineering, Data-based investigations, Design Thinking, and Computer-Aided Design. The AI unit has been designed and delivered as a student-centered and project-based one. Following the three-unit approach outlined by Karampelas (2020), students were expected to demonstrate conceptual understanding of a) the impact of Artificial Intelligence on societies and economies, b) machine perception with a focus on computer vision and limitations of AI, and c) machine learning through the Artificial Neural Networks and Minimum Spanning Trees techniques. As the learning design intended to contribute into bridging the gap between AI supply and demand in K-12 for all students to address an AI-heavy society, the focus was on conceptual understanding rather than sophisticated programming or robotics (which should also be considered within a holistic curriculum).

The following sections describe the research methods used, the main characteristics of the student sample, the raw data and the respective analysis, as well as concluding remarks about the research findings.

Methods and student sample
In order to investigate a) students' perceptions, awareness, and prior knowledge about Artificial Intelligence and Machine Learning, b) the effect of AI learning experiences on students' dispositions toward AI, c) the effect of AI learning experiences on students' understanding about

 https://doi.org/10.51724/hjstemed.v1i2.13

the potential and limitations of AI, and d) the possible correlations between gender and grade level and performance scores, as well as to explore student satisfaction and areas of concern, a relevant questionnaire has been distributed to students upon the completion of the Artificial Intelligence unit of the STEAM course. Naturally, the learning design and delivery has been tailored to provide students with opportunities to learn, grow, and reflect in measurable ways, mostly through the alignment of teaching resources and assessments with the learning objectives. The language of instruction was English. Most of the topic was delivered virtually because of COVID-19 restrictions.

The student sample included 62 international high school students, the majority being U.S., Greek, and Chinese. Most of the students are girls, while about three quarters of them are 9th graders. **Table 1** summarizes the main characteristics of the student sample.

Table 1. Overview of the student sample in terms of sample size, gender, and grade level.

Category	Classification	Number of students	Fraction of students (%)
Gender	Female	36	58.1
	Male	26	41.9
Grade Level	9th	48	77.4
	10th	8	12.9
	11th	6	9.7
	Total	62	100.0

Research data

The questionnaire raw data and data analysis are presented below, occasionally including methodologies and data evaluation. They are structured around the students' views on and prior knowledge about AI, student performance in relation to gender and grade level, student understanding per delivered AI topic (societal impact, machine perception, and machine learning), and student satisfaction and challenges.

Students' general views on Artificial Intelligence learning

About two thirds of the students (66.1%) reported that they have not engaged with Artificial Intelligence content and activities in the school before. The rest were asked to elaborate further, with many of those students demonstrating false or doubtful understanding about the question itself or the claimed AI-related experiences (e.g. mentioned the assembly of robots, taking the STEAM course, or using science simulations), therefore the percentage above most probably underestimates the sparsity of prior AI-related learning experiences. Eventually, all 62 students recognized they have become more knowledgeable about Artificial Intelligence after completing the STEAM course's Artificial Intelligence Unit, while almost all of them (96.8%) agreed that students should engage with Artificial Intelligence content in high school.

Assessment scores

Figure 1 illustrates the distribution of the average student scores in the assessments of the Artificial Intelligence unit. Overall, there were four assessments, each one of them been graded from 5 (50% performance level representing insufficient deliveries or lack of) to 10 (100% performance level representing deliveries at or beyond expectations) at a precision of 0.5, hence the histogram's 0.5 bin size. The distribution is skewed toward the higher end of the grade spectrum.

Figure 2 presents a breakdown of the student scores per assessment through the respective percentiles. Typical for box-and-whisker plots and from the bottom to the top, the horizontal box lines represent the first (Q1), second (Q2: median) and third (Q3) score quartiles, boxes include the middle half of the data, while the whiskers visualize the range of scores. The median of the student assessment scores in the Cyber-Physical Interactions assessment coincides with the third quartile. Even though the scores refer to assessments different in scope and means of delivery therefore they cannot be directly compared, the students performed better in the Cyber-Physical Interactions assessment. Assessment descriptions are included in the "Societal impact of Artificial Intelligence", "Machine perception", and "Machine learning" sections.

Figure 3 implies better performance for the 10th graders followed by 9th graders, but the small sample sizes of 10th and 11th grade students that took the AI unit (**Table 1**) do not allow for further exploration.

Figure 4 summarizes students' average assessment scores per gender. A two-sample Kolmogorov-Smirnov statistical test was conducted to investigate a possible correlation between a student's gender and their average assessment score. The Null Hypothesis H_0 states that the two samples are drawn from the same distribution (i.e. scores are independent of gender), while the alternative hypothesis H_A states the two samples are not drawn from the same distribution. H_0 was accepted at the 5% significance level (α), because the calculated p-value (0.433) was larger than α (0.05). The Test is summarized in **Table 2.**

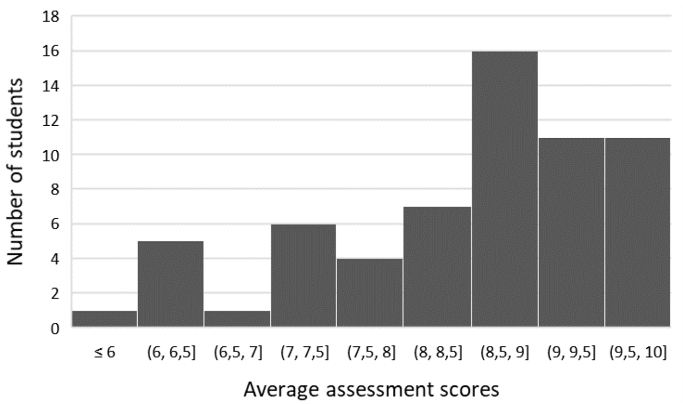

Figure 1. Distribution of average student scores in the assessments of the Artificial Intelligence unit.

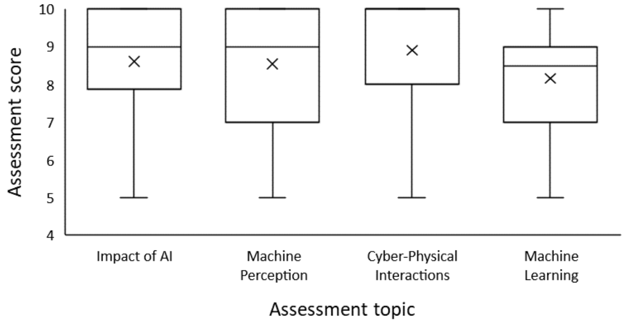

Figure 2. Box-and-whisker plots of the students' scores per assessment. The X symbol denotes the average score for the respective assessment.

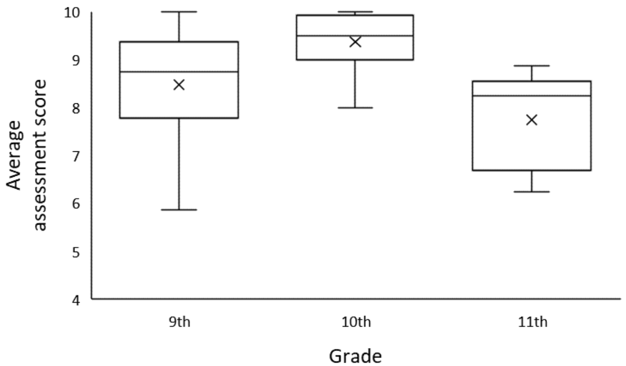

Figure 3. The Box-and-whisker plots of the students' average assessment scores per grade level.

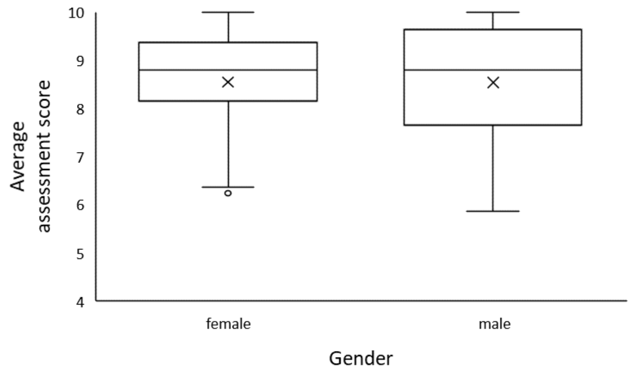

Figure 4. Box-and-whisker plots of the students' average assessment scores per gender. The circle denotes a possible outlier value.

Table 2. A summary of the two-sample Kolmogorov-Smirnov test that was conducted to investigate a possible correlation between the students' average assessment score and their gender.

H_0 (Null Hypothesis)	Samples are drawn from the same distribution
H_A (Alternative Hypothesis)	Samples are not drawn from the same distribution
α (significance level)	0.05
p-value	0.433
Result	Accept H_0

Societal impact of Artificial Intelligence

In this first topic of the Artificial Intelligence unit, students were introduced to basic concepts of AI and indicative cases of AI impact. Following, they were asked to work in groups to research and present the impact of AI on an area of focus, e.g. employment, ethics, healthcare, education, natural disasters, global economy, and transportation. Students were expected to understand the societal impact of Artificial Intelligence via three different modes of learning: a) personal mode of independent research about the given focus area, b) collaborative mode in blending and summarizing the group findings, and c) observer mode by watching the presentations about the rest of the areas of focus.

Figure 5 showcases the anticipated impact of Artificial Intelligence on societies as selected by students with respect to experiencing the AI unit, including positive, mostly positive, negative, mostly negative, and no anticipated societal impact, as well as cases of students who have not thought about the AI impact before or knew nothing about AI. The latter two options were excluded from the evaluation after the AI unit as being meaningless. An inspection of the graph reveals an existing optimism about Artificial Intelligence that grew stronger during the AI unit. Figures 6 and 7 present the findings through the different genders with notable deviations not being observed. Both genders exhibit a change toward the positive end of the defined AI societal impact spectrum, although male high school students appear to be more positive prior to taking the AI unit. Figures 8 and 9 do not imply a correlation between students' AI impact anticipation (either before or after taking the AI unit) and their average assessment score, with the possible exception of top performing students expecting a "mostly negative" AI societal impact before taking the AI unit. Low number statistics do not allow for further validation. Small sample sizes do not allow for all Figure 5-7 anticipation categories to be statistically summarized in Figures 8-9 either.

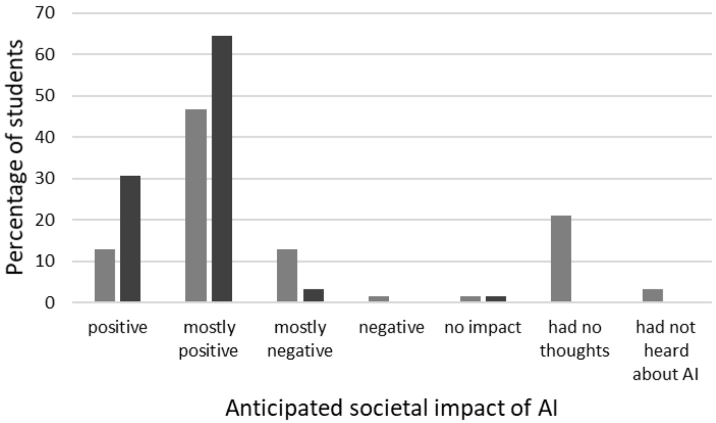

Figure 5. The anticipated societal impact of Artificial Intelligence as reported by students before (light gray) and after taking the AI unit (dark gray).

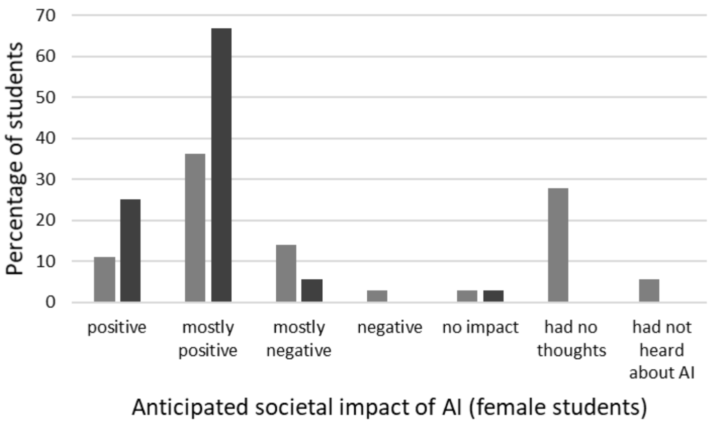

Figure 6. The anticipated societal impact of Artificial Intelligence as reported by female students before (light gray) and after taking the AI unit (dark gray).

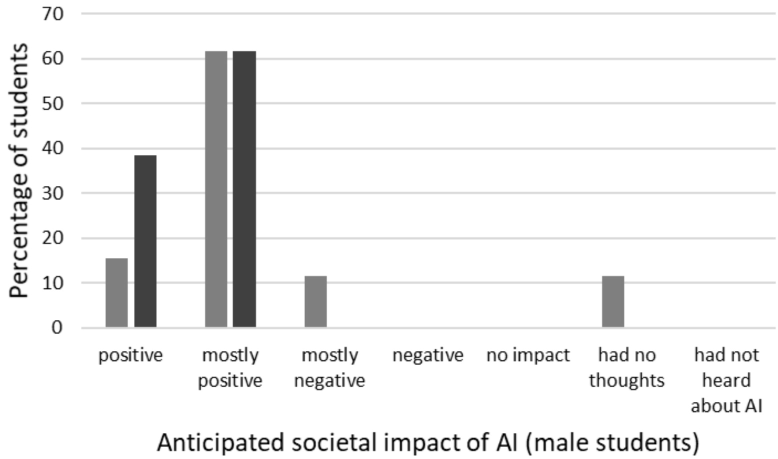

Figure 7. The anticipated societal impact of Artificial Intelligence as reported by male students before (light gray) and after taking the AI unit (dark gray).

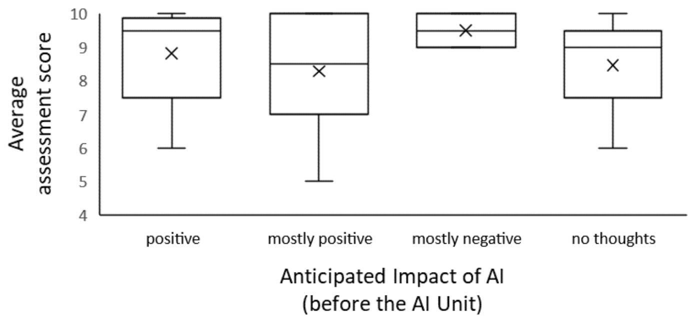

Figure 8. The box-and-whisker plots of the students' average assessment scores per reported anticipated societal impact of Artificial Intelligence before taking the AI unit. Low frequency anticipations have not been included because they cannot be visualized through their quartiles.

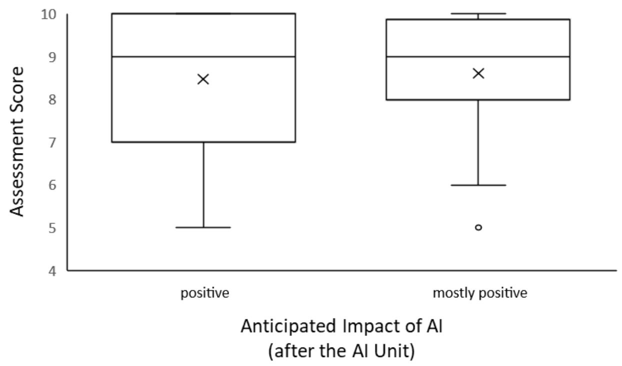

Figure 9. The box-and-whisker plots of the students' average assessment scores per reported anticipated societal impact of Artificial Intelligence after taking the AI unit. Low frequency anticipations have not been included because they cannot be visualized through their quartiles.

Machine perception

In this second topic of the AI unit, students were initially introduced and discussed about Machine perception and the correspondence between sensors and human senses, followed by a discussion around Computer Vision and autonomous vehicles. Most of the students (93.5%) reported prior familiarity with the concept of autonomous vehicles. The relevant assignment required students to brainstorm about and design a self-driving cart to be used in the school premises, and describe its potential functionalities, limitations, and safety concerns. The majority of them recognized an enhancement of their understanding about the potential of computer vision (**Figure 10**).

Then, students interacted with Artificial Intelligence online applications (chatbot, voice recognition apps, computer vision apps, indicatively: Experiments with Google), explored their functionalities and limitations, and were assessed in demonstrating the above. Indicatively, they explored the level at which a) *Mitsuku* the chatbot processes natural language and carries out chat conversations, b) *Poem Portraits* uses input words to generate poems, c) *Akinator* guesses real or fictional characters by asking effective questions, d) *MixLab* creates music based on voice commands, e) *Art Palette* matches pictures of similar color patterns, f) *Quick, Draw!* recognizes user drawings, g) *Algorithmia* automatically colorizes grayscale photographs, h) *Petalica Paint* automatically colorizes user sketches, i) *Giorgio Cam* recognizes objects captured by a camera and incorporates their identified name into an ongoing song's lyrics, j) *Semi-Conductor* creates music based on the detection of body motion through a camera, and k) *Teachable Machine* recognizes images, sounds, and poses upon training provided by the user. The students were encouraged to not only familiarize themselves with the variety of the AI approaches represented above, but also to explore their performance. Most of the students recognized that their interaction with AI online applications helped them realize the limitations of Artificial Intelligence (**Figure 11**). The "other" category regarding the limitations of AI represents two student answers beyond simply confirming or rejecting the improvement (one stated that AI has no limitations, while the other recognized improvement despite a reported existence of prior knowledge). Overall, students enjoyed interacting with AI apps the most (34%) out of the entirety of the AI unit (**Figure 14**).

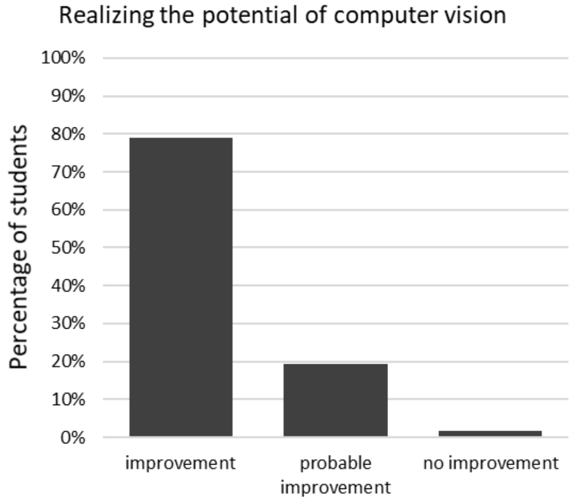

Figure 10. The improvement in realizing the potential of computer vision as reported by students.

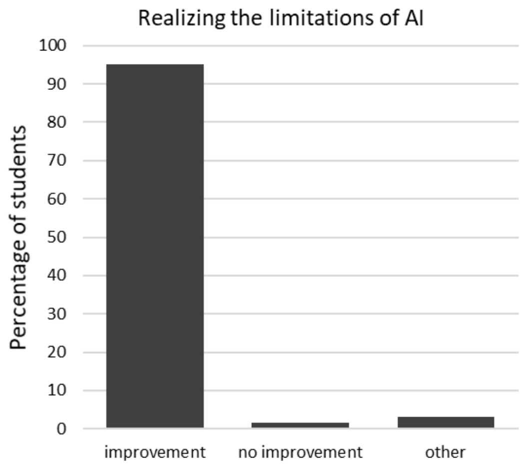

Figure 11. The improvement in realizing the limitations of Artificial Intelligence as reported by students.

Machine learning

In the third topic of the AI unit students were introduced to the concept of machines learning from data as opposed to machines performing tasks they are explicitly programmed to, with 59.7% of the students reporting that there were not aware of this fact before taking the AI unit. Then, students were engaged with activities on two machine learning algorithms selected to represent different computational approaches: Minimum Spanning Trees (MST) and Artificial Neural Networks (ANN). More specifically, students were asked to find MSTs that would connect nodes without loops by using either the Kruskal's or the Prim's algorithms (**Figure 12**), and consider the functionality of simple 2-layer neural networks focusing on the activation of a neuron in terms of its minimum signal strength and the weights of the neurons of the previous layer (**Figure 13**), including engaging with a ANN simulation game (Karampelas, 2018). The two assessments of the aforementioned machine learning algorithms were graded as one (average of MST and ANN score).

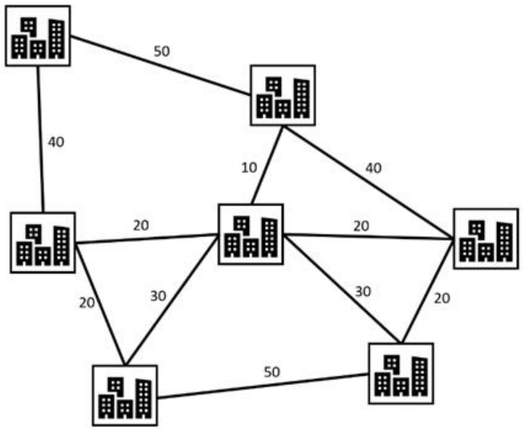

Figure 12. Assessment visual: students were asked to find the Minimum Spanning Tree that includes all cities. The numbers denote the edge weights of the graph and represent distance in kilometers.

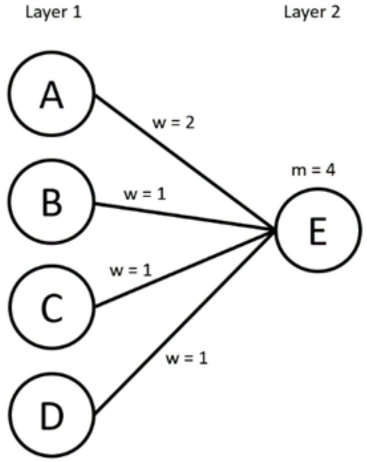

Figure 13. Assessment visual: students were asked to identify whether neuron E activates. The w values denote the weights of the Layer 1 neurons (A through D) while the m value denotes the minimum signal strength of activation for the Layer 2 neuron (E).

Only two students (3.2%) reported that they have heard about the Minimum Spanning Trees algorithm before taking the AI unit, while three students (4.8%) said they had engaged with relevant learning activities in the past. This logical inconsistency might have been caused by one student's misunderstanding of the relevant questions.

Students reported a higher still relatively low percentage of prior knowledge of the term and engagement with Artificial Neural Networks (17.7% and 6.5%, respectively) than Minimum Spanning Trees, while 95.2% of them recognized that the aforementioned Neural Networks game simulation helped them further understand how Artificial Neural Networks work.

Most of the students also reported their belief that high school students are capable of conceptually understanding Machine Learning algorithms like Neural Networks and Minimum Spanning Trees (69.4%), with the rest answering "maybe". A correlation between the students' assessment scores in the Machine Learning assignment and their answers regarding the capacity of high school students to conceptually learn Machine Learning is highly unlikely, given the similarity in the grades' profiles: Affirmative answers correspond to grades with an average of 8.2 and a standard deviation of 1.6, while the hesitant ones represent a score average of 8.1 and a standard deviation of 1.4.

An interesting finding was that the Machine Learning topic was accompanied both by students' satisfaction and struggle ("Student reflections" section and **Figures 14** and **15**): 23% of them named Machine Learning, ANN, or MST (or combinations of) as having enjoyed the most while experiencing the Artificial Intelligence topic, while a comparable fraction (27%) named ANN and/or MST as their biggest challenge, with most of those students reporting they overcome the challenge through hard work.

Student reflections

Figures 14 and **15** illustrate the emerging themes from the students' free responses on what they enjoyed the most and what they found most challenging (including actions taken to overcome the challenges) regarding the Artificial Intelligence unit. Interacting with online AI apps was the main source of students' satisfaction (34%), followed by engaging with fun and/or creative projects in general (27%, including designing a self-driving cart), and Machine Learning (23%, including MST and ANN). The latter was also the most reported challenge (27%), even though most of them also reported they had eventually overcome it. Furthermore, and as most of the Artificial Intelligence unit of the STEAM course was delivered virtually because of COVID-19 restrictions, 13% of the students named online learning and time management as challenges.

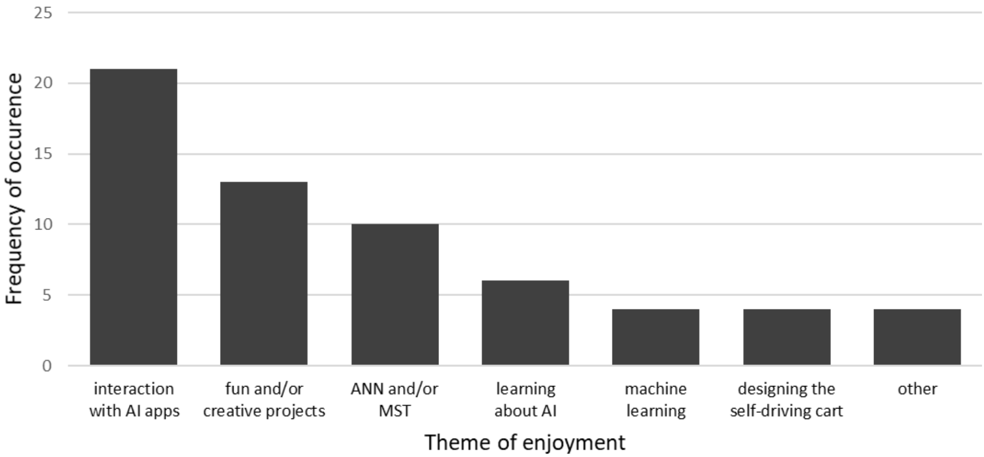

Figure 14. Themes of student-reported enjoyment regarding the Artificial Intelligence unit.

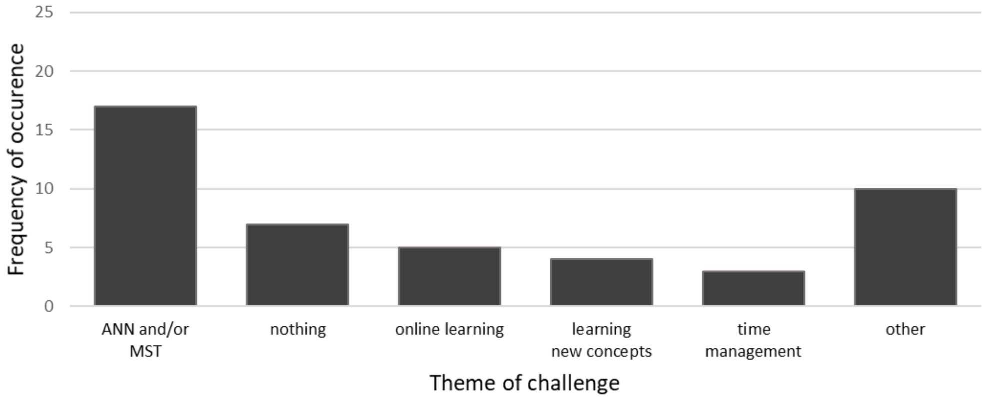

Figure 15. Themes of student-reported challenges regarding the Artificial Intelligence unit.

Concluding Remarks

Artificial Intelligence, in all its disruption capacity and interdisciplinarity, could be introduced into K-12 education in several ways. This study presents performance and reflection data of international high school students that took an Artificial Intelligence unit within a STEAM course. The learning design allowed for a) content and skills feasible to be obtained by students of no relevant background, b) a student-centered, teacher-facilitated approach of independent and collaborative inquiries, c) opportunities for students to be creative through brainstorming solutions and interacting with online AI apps, and d) opportunities for students to conceptually understand selected Machine Learning algorithms. Coding and robotics were not considered.

Students reported low to no prior engagement with Artificial Intelligence and (especially) Machine Learning content, reflecting the gap between supply of and demand for AI education. Taking the AI unit had students reporting a shift in their anticipated impact of AI on societies toward a more positive future outcome, as well as an overall personal growth and improved understanding regarding the AI content. No correlation has been found between student performance and their gender.

Machine Learning activities and assessments were categorized both as challenges and sources of enjoyment by the students of the sample under investigation. Given the diversity of Machine Learning algorithms (beyond ANN and MST to, indicatively, K-Means Clustering, Regression, and Decision Trees) and their capacity to be introduced to students at different levels of abstraction (conceptual functionality that promotes computational thinking, statistics and probability, programming and data analysis), the topic appears to be a promising one for the AI-themed courses of the future.

Finally, students performed very well and reported enjoyment about exploring the functionality and limitations of AI through interacting with online AI apps. Such readily available, diverse, game-like, often art-focused applications could enhance student understanding and become an important aspect of the application of AI in education. For example, the students demonstrated curiosity and creativity interacting with the *Teachable Machine*, a user-friendly educational AI tool purposefully designed to allow for an intuitive understanding and customization of machine learning classification models (Carney et al., 2020). Additionally, students were observed to considerably engage and enjoy interacting with *Mitsuku*, the chatbot. The duration of the interaction with AI could affect student engagement though. Croes & Antheunis (2021) found that the initially high expectations of people interacting with *Mitsuku* decreased after several interactions with the chatbot over a three-week period (in contrast to the twenty-minute interaction outlined in this paper), with participants feeling less socially attracted toward the chatbot over time.

This study shed light on the high school AI classroom at the crossroad between applying AI in education and educating students about AI, with insights from the students' experiences and perspectives. As national K-12 curricula will be enriching their offerings with Artificial Intelligence and Machine Learning, and as relevant educational and technological initiatives from the public and the private sector will be coming to fruition, more classroom data will be available for the AI scholars' community to assess and evolve the corresponding learning designs.

References

Carney, M., Webster, B., Alvarado, I., Phillips, K., Howell, N., Griffith, J., Jongejan, J., Pitaru, A., & Chen, A. (2020). Teachable Machine: Approachable Web-Based Tool for Exploring Machine Learning Classification. CHI '20: CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery

Croes, E. A. J., & Antheunis, M. L. (2021). Can we be friends with Mitsuku? A longitudinal study on the process of relationship formation between humans and a social chatbot. *Journal of Social and Personal Relationships, 38* (1) 279-300, SAGE

Experiments with Google. (2020). Retrieved from https://experiments.withgoogle.com/

Jing, M. (2018). China looks to school kids to win the global AI race. Published May 3, 2018. Retrieved from https://www.scmp.com/tech/china-tech/article/2144396/china-looks-school-kids-win-global-ai-race

Karampelas, A. (2021). *Building a design-centered STEAM course.* In Avgerinou, M.D., & Pelonis, P. (Eds.), Handbook of research on K-12 blended and virtual learning through the i²Flex classroom model. IGI Global

Karampelas, A. (2020). *Developing and delivering a high school Artificial Intelligence course in blended and online learning environments*. European Distance and E-Learning Network (EDEN) Proceedings, European Distance and E-Learning Network, pp. 255-261.

Karampelas, A. (2018). Introducing Artificial Intelligence and Machine Learning in Secondary Education. Astrolavos. *Journal of New Technologies, Hellenic Mathematical Society*, Issue 29-30

Schwab, S. (2016). The Fourth Industrial Revolution: what it means and how to respond. Published Jan. 14, 2016. Retrieved from https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/

Synced (2018). *Chinese publisher introduces AI textbooks for preschoolers*. Published Dec. 5, 2018. Retrieved from https://medium.com/syncedreview/chinese-publisher-introduces-ai-textbooks-for-preschoolers-b95e1a89cfa0

Touretzky, D., Gardner-McCune, C., Martin, F., & Seehorn, D. (2019). Envisioning AI for K-12: What Should Every Child Know about AI? *Proceedings of the AAAI Conference on Artificial Intelligence, 33*, 9795–9799. https://doi.org/10.1609/aaai.v33i01.33019795

World Economic Forum. (2020). *Schools of the Future. Defining New Models of Education for the Fourth Industrial Revolution* (REF 09012020)