Preparing High School Teachers to Integrate AI Methods into STEM Classrooms

Irene Lee, Beatriz Perret
Education Development Center, Inc.
ilee@edc.org, bperret@edc.org

Abstract

In this experience report, we describe an Artificial Intelligence (AI) Methods in Data Science (DS) curriculum and professional development (PD) program designed to prepare high school teachers with AI content knowledge and an understanding of the ethical issues posed by bias in AI to support their integration of AI methods into existing STEM classrooms. The curriculum consists of 5-day units on Data Analytics, Decision trees, Machine Learning, Neural Networks, and Transfer learning that follow a scaffolded learning progression consisting of introductions to concepts grounded in everyday experiences, hands-on activities, interactive web-based tools, and inspecting and modifying the code used to build, train and test AI models within Google Colab notebooks. The participants in the PD program were secondary school teachers from the Southwest and Northeast regions of the United States who represented a variety of STEM disciplines: Biology, Chemistry, Physics, Engineering, and Mathematics. We share findings on teacher outcomes from the implementation of two one-week PD workshops during the summer of 2021 and share suggestions for improvements provided by teachers. We conclude with a discussion of affordances and challenges encountered in preparing teachers to integrate AI education into disciplinary classrooms.

Introduction

The rapid expansion of Artificial Intelligence is having unprecedented industrial and social impact and, as a result, is transforming our lives, and our futures. To productively participate in the age of AI, all youth must gain a fundamental understanding about how AI works and how it will impact their lives. Yet significant barriers to AI education exist. The current state of AI education lacks equity. Specialized programs offering fee-based AI courses for youth primarily serve students who are privileged with financial resources and parental support. US schools rarely offer AI education because AI curricula that are relevant and engaging to K-12 students are only now being developed, few teachers are prepared to offer AI education, and there are no National educational standards for AI education. Furthermore, AI has been traditionally taught in post-secondary education with a focus on its mathematical and computational underpinnings making it uninviting and intangible to students who lack early exposure and adequate preparation in these topics (Sulmont, Patitsas, & Cooperstock, 2019). Thus AI education as it currently exists is inaccessible to a majority of students—those who attend schools that do not have teachers prepared to offer AI education or cannot afford private lessons.

To advance the field of K-12 AI education and address the need to prepare diverse audiences for life and work in the AI-enhanced future, two innovations are needed. The first is an AI curriculum that is developmentally appropriate for high school students, can be integrated into existing courses, and is capable of engaging students from underrepresented groups in STEM and computing. The second is an approach to preparing teachers to offer the AI curriculum in a wide range of disciplinary settings.

To address these needs, EDC's Computational Sciences Pathway Option for Massachusetts High School Students project, or "Science+C" (NSF award #1934112) developed an AI curriculum called AI Methods in Data Science (AIMSinDS) that is designed to be integrated within existing STEM courses with the goal of elucidating how AI methods can be applied in STEM fields. The initial targets for integration were Science+C's three new courses (Biology+C, Chemistry+C and Physics+C) that integrate computational science methods into high school science classrooms. Four of the eleven units developed for each course featured Data Science (DS) and AI and were packaged separately as the "AI Methods in Data Science" (AIMSinDS) curriculum for piloting and field-testing purposes. Several partnerships were leveraged to support the project. Co-PI Lee, through her connections at MIT, recruited four

CS/AI undergraduate students as curriculum design collaborators. The EDC Science+C project also partnered with two CS-focused teacher PD providers (CS Alliance and CSforMA) to recruit teachers and offer our pilot PD workshops to teachers.

In this experience report we describe the AIMSinDS curriculum and associated teacher PD workshop then share results from an exploratory study of teacher learning and attitudes toward implementing AI lessons within their disciplinary courses.

Theoretical Foundations

Data science (DS) and AI are two important technologies that contribute to discovery and innovation in STEM fields. DS is an interdisciplinary data-driven field that combines statistics, analytics, and informatics to extract knowledge from data and apply that knowledge to understand phenomena in various fields of study. DS draws on various analytical and visualization techniques and theories from fields such as mathematics, computer science and information science and domain-specific knowledge and techniques germane to the field of application. AI has been framed as a method in DS - the ACM Computing Curricula 2020 (CC2020) report positions Machine Learning as a knowledge area within DS. AI methods are used to construct computer-based models and algorithms capable of making classifications and predictions of the future. In the AIMSinDS curriculum, we take an application-focused perspective (rather than a theoretical perspective) and frame AI methods as another set of tools and techniques for extracting insights and patterns from large datasets.

National computer science and science education organizations have promoted AI and DS education. The Computer Science Teachers Association revised 2017 Standards include two standards at the high school level that explicitly address AI (CSTA 2017). The National Research Council's Framework for K–12 Science Education and the Next Generation Science Standards stipulates that education in science should interweave content, modern scientific practice, and crosscutting concepts (NGSS Lead States 2013). A recently proposed California state standard in mathematics incorporates Data science as a subfield of mathematics (CADoE 2021). Additionally, the AI4K12 Initiative (ai4k12.org) has published guidelines for teaching AI in grades K-12 organized into 5 Big Ideas.

AI education is in its infancy and early findings on AI Literacy (for example, Druga et al. 2019; Cheng et al. 2020; DiPaola et al. 2020; Lee and Ali 2021; Long and Magerko 2020; Greenwald, Leitner & Wang 2021) are just beginning to shed light on how learners gain an understanding of AI concepts and processes and the ability to incorporate AI processes within their own applications. Various curricula have been introduced to teach fundamental AI concepts in K-12 (for example, Ali et al. 2019; Williams, Park and Breazeal 2019; Sabuncuoglu 2020; Lee et al. 2021). Yet, previous efforts in computer science (CS) education have shown that establishing new standalone CS courses is often difficult.

We embarked on constructing a curriculum to engage students in AI and DS through integration into core subjects as a means to prepare students with the knowledge, skills, and practices for future endeavors in STEM fields. This integration aims to bypass some of the difficulties associated with offering stand-alone courses and adding additional subjects to an already crowded school day. Taking an integration approach has many potential benefits. Since core subject classes are mandated for all students, integrating AI and DS into them serves as a way to introduce all students in an equitable fashion. Interjecting AI and DS into science classes in particular drives the modernization of science curricula to reflect modern scientific practices. The integration also aspires to raise teachers' and students' awareness of the relevance of AI and DS in STEM and of how STEM fields are rapidly changing through the integration of AI and DS methods.

While there are numerous potential benefits, there are also several substantial challenges to this integration approach. Teachers may have difficulty finding a fit and time in a crowded curriculum to teach AI and DS. Furthermore, since AI and DS content knowledge and skills are not yet tested in standardized assessments, their value is not measured or used to assess teaching, thus their inclusion into the school day is deprioritized.

Another significant hurdle to the implementation of AI education with K-12 is the dearth of teacher professional development in AI. A few recent papers report on attempts to characterize the types of expertise teachers need to teach AI (Kim 2021), and pedagogical approaches that have been effective in teaching AI (Sanusi and Oyelere 2020; Eguchi 2021). Yet the field is just beginning to investigate what preparation, in terms of knowledge, skills, and pedagogy, teachers will need to be able to teach AI.

Audience

AIMSinDS recruited high school STEM teachers to participate in its one-week pilot summer camps. A program partner serving the Southwest region of the US recruited 10 participants (Cohort 1) for the first summer workshop held online in June 2021. Another 9 participants (Cohort 2) were recruited through a program partner in the Northeast region of the US for the second summer workshop held online in August 2021. In total 19 in-service STEM teachers participated in our pilot professional development workshops. Participants' registration data show the two groups were very different. Cohort 1 teachers taught a wide range of grades: 2 taught exclusively at the middle school level, 4 taught exclusively at the high school level, and 4 taught at both the middle and high school levels. Cohort 1 teachers also taught a wide range of STEM subjects including general science, earth science, biology, chemistry, physics, environmental science, engineering, math, computer science, as well as business, health and social studies. Cohort 1 teachers had a range of prior experience with computer science having attended workshops on Project GUTS, Micro:bit, Scratch, NetLogo, Bootstrap Data Science, and Project Lead the Way CS offerings. One teacher had prior experience with Google Colab notebooks. In contrast, all but one of the Cohort 2 teachers taught exclusively at the high school level and six of the nine teachers were math teachers. Of the other three teachers one taught engineering, one taught CS, and one taught physics. Cohort 2's range of prior CS experiences included programming in R, NetLogo, Java, Python, Pascal, Basic, R and C/C++. One Cohort 2 teacher had no prior experience with programming.

Approach

The AIMSinDS curriculum was designed for high school students with no prior AI knowledge and skills. It was used as the basis for the teacher PD because we wanted teachers to experience the curriculum as learners prior to engaging in discussions of how to teach how to teach these lessons and ground them in concepts with which their students are familiar. We leveraged findings from MIT's "Developing AI Literacy" (DAILy) curriculum that found the interweaving AI concepts, Ethics and AI, and AI careers and futures was productive in engaging youth and supported their development as AI Literate citizens (Lee et al. 2021). Instead of the math- and theory-centric approach to teaching AI, we used a scaffolded approach wherein we ground AI concepts in everyday experiences, consider the ethical issues in AI, and engage in hands-on activities with AI tools and interactive games. After the introduction of an AI concept or method, interactive websites, participatory simulations, and games are used to further engage learners in the underlying constructs. Finally, we utilized Google Colab notebooks as interactive playgrounds in which learners can inspect and modify the code that was provided as an example, build AI models, train and test the model, and determine its predictive accuracy. This addition of inspecting code was made in response to requests from program partners and educators who asked to learn "how to code AI." We hoped that using Google Colabs to expose the scant code "under the hood" that generates a model and uses the model to make predictions would re-orient learners to think of AI as developing models rather than as coding. Google Colab was chosen because it is accessible to learners as part of the Google Suite, and provides them with the opportunity to modify the code that builds the model and see the impacts of those changes on the model's accuracy.

A co-design process was used to create the curriculum. The four undergraduate selected as co-designers had recently studied courses on machine learning (ML) themselves and several of whom had been teaching ML in fee-for-service programs. As recent learners of ML, they had opinions on what precursor knowledge is needed before new ML concepts are introduced, and what information should be emphasized in teaching ML. For example, one co-designer emphasized the need to expose that the point of refining an ML model was to minimize the cost function, something she didn't learn until very late in the ML course she took. To bridge their experience to Lee's previous work, the co-designers were presented with how AI is taught at the middle school level using the DAILy curriculum as an example. Over a 6-week period, the project team and co-designers worked together to generate a skeleton progression that a) reflected commonalities in how the student co-designers wished they had been taught, b) included the activities they felt were most accessible and engaging for high school students, and c) included datasets aligned with STEM subject areas. Leveraging a promising practice from the DAILy middle school AI curriculum (Zhang et al 2021), a design criteria for the AIMSinDS curriculum was that it should interweave the issue of bias in AI and its ethical implications within each unit and also makes explicit links to real world uses of the various AI methods introduced.

A "stretch" goal for the curriculum was to enable learners to gain discernment on when to use different models and techniques. This goal responds to two emerging issues in AI education: science fairs and competitions are increasingly seeing student projects featuring AI methods but when asked students rarely know why the method they chose was better or more appropriate to the problem or dataset than others, and a finding from MIT's 6.S198 "Deep Learning Practicum" course that "choosing a model" is a necessary addition to the Use-Modify-Create progression (Lao, Lee & Abelson 2019; Martin et al. 2020) in the context of AI education. Simply put, it is not enough to use the tool provided or found in literature. Instead, students need to know when and under what conditions to choose one AI model over another.

Curriculum

The organizing framework for the curriculum was the categorization and nesting of different techniques under the umbrella of AI Methods in Data Science. (See Figure 1 below.) We divide DS into traditional data analytics and AI.

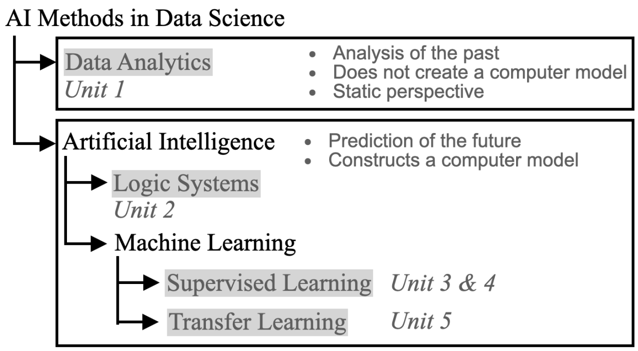

Figure 1. The Organizing Framework for the curriculum

Within the AI, we differentiate between logic systems that develop models using human-understandable rules to make predictions, and machine learning in which models are built and tuned or adapted from experience, not by writing better rules. Machine learning is further subdivided into Supervised learning that uses datasets with labels; and Transfer Learning that uses pre-built models as a base layer to build upon when classifying new data.

The AIMSinDS curriculum is composed of five units: Data Analytics, Logic Systems featuring decision trees, Machine Learning featuring Teachable Machine and perceptrons, Supervised Learning featuring neural networks (NNs), and Transfer Learning featuring the K Nearest Neighbor algorithm. The choice of the five units was based on wanting to expose learners to a variety of common methods, and to provide a trajectory of learning that went from human-inspired symbolic AI to sub-symbolic methods such as NNs. We sought to explain how transfer learning works to mitigate a common misconception that Teachable Machine was training a NN in real time.

Each unit is composed of five lessons that followed a trajectory from playful experiential learning and connecting to real world issues and careers, to the articulation of key concepts, to hands-on interactive activities, and finally to Google Colab activities in which students run summary statistics in R or Python, generate and use AI models to make predictions using Scikit-learn or PyTorch, and visualize the results. (See Table 1 below.)

An example: Supervised Learning / Neural Networks unit To begin the unit on NNs, we introduce an artificial neural network as a model that is based on the human neural network. We connect NNs to everyday life by pointing out that NN are key to common applications that students use. For example, a NN can take in different data sources (audio, image, video, text, etc.) and make recommendations such as making movie recommendations and classifications such as classifying images of cats and dogs. We also watch a video of a STEM professional of color introducing the structure and function of a NN.

Next, we play the "Neural Network game" (Lee and Ali, 2021) as an online interactive game within a Google Drawing. The goal of the game is for the NN to predict the caption of an image. Students acting as nodes in the network select and pass words through the network based on whether they think the words will be part of the caption for the image. The processes of feeding forward, evaluation, and back propagation are mimicked in each round of the game. Connections between nodes that helped to predict the caption are strengthened, whereas those that do not are weakened. Students graphically carry out the weighting of links between nodes by thickening or thinning the lines drawn between nodes. In subsequent rounds of play, words provided by the nodes connected through thicker lines are to be preferred to words coming from nodes that are only

Lessons in Unit 1: Data Analytics & Visualization
1: (P) "Be the Dataset" activity, Careers: Data scientist.
2: (K) Sources of data, data exhaust,privacy; Summary statistics
3: (C) Intro to summary statistics in R using Pokemon dataset.
4: (K/I) Data visualizations and Visualization Matching game.
5: (C) Loading, analyzing & visualizing the Mushroom dataset. Colab using R. Ethical considerations in Data Science.
Lessons in Unit 2: Logic Systems / Decision Trees (DT)
1: (P) Introduction to AI/ AI or not; "Is it for Winter?" game
2: (K) Human vs Computer Decision Trees; Selecting features; Predictive accuracy; Bias in DTs. Use of DTs in science.
3: (I/K) Slice of ML; features, labels, training & testing a DT.
4: (I/K) Split a dataset, set accuracy goal, train & test DT.
5: (K/C) CART, Gini index; train & test w Scikit-learn in Colab.
Lessons in Unit 3: Machine Learning (ML) / Perceptrons
1: (P) Intro to ML; Tech that use ML; Teachable Machine (TM); train & test a model in TM, determine accuracy; Image recog; Ethics of facial recognition.
2: (K) Labels; Introduction to Supervised Learning; Functions.
3: (K/I) Investigating Bias in ML (case studies/investigations).
4: (K/I/C) Cost function; Gradient descent exploration;
5: (K/I) Perceptrons; Summing weighted inputs; Activation func. Will the dog bite me? Game / participatory simulation.
Lessons in Unit 4: Supervised Learning / Neural Networks (NN)
1: (P) Intro to NNs; NN Game; NNs in STEM fields / careers.
2: (K) Feed forward; evaluation; back propagation; learning;
3: (C) Split datasets, over/underfitting; build a NN; Loss func.; Minimize loss using gradient descent; bias in NNs;
4: (C) Train & test a NN. Predicting an unknown value.
5: (K/I) Problems NN can solve; NNs in science and medicine. Ethics and AI. Where in AI dev. pipeline can bias creep in.
Lessons in Unit 5: Transfer Learning / K Nearest Neighbors (KNN)
1: (P) Training large models: data req., power req., env. impact
2: (K) How does Teachable Machine work? Transfer Learning;
3: (K/I) Vectors; High-Dim. spaces; Neural numbers activity and compact vector representations; Embedding Projector (EP). Ethical issues w.r.t mappings, analogies in embedding space.
4: (K/I) Reducing models; KNN and ex, of Clustering in EP.
5: (C) KNN for Classification; Visualization of KNN Colab.

Table 1. Units & Lessons in the AIMSinDS curriculum.

weakly connected. After playing the game, we introduced the participants to how skewed datasets can introduce bias in a neural network model. We viewed and discussed Buolamwini's work on bias and facial recognition systems.

Finally, after having been exposed to the structure and function of a neural network, how learning occurs, and where bias might creep in, students work through a Google Colab containing code and instructions to build, train and test a NN on a numerical dataset. The Iris dataset was used because it had been inspected in a prior unit and we prioritized learning a new AI method and comparing methods over introducing and gaining familiarity with a new dataset. In the Google Colab they learn the steps in the process of building, training and testing a NN to predict the length of an iris flower's petal based on three other measurements (sepal width, sepal length, and petal width). In this final step, the instructor can refer back to the activities in the first two lessons in this unit to discuss how each feature is an input of the NN and speculate how the weights in the network may have changed after the training stage, based on each feature's importance. The familiarity of the dataset facilitates a discussion about the AI method in question but creates a barrier to demonstrating how to discern between methods. Connections are also made to gradient descent (an algorithm introduced in the previous unit) as the method NNs use to adjust the weights to increase the algorithm's accuracy. Within the five-day progression we also include an activity wherein learners read about different applications of neural networks in STEM fields and to share their findings with their classmates.

Professional Development

The AIMSinDS PD workshop design was informed by research on effective PD in general (Gaible and Burns, 2005; Hassel, 1999) and effective PD when integrating Computational Thinking across the curriculum (Yadav et al. 2016). Our PD supports core areas of teaching such as content, curriculum, instructional practices and assessment. The PD also takes teachers' implementation contexts and needs into account by providing modular units that can be inserted singly within existing curriculum or as a full set and enabling teachers to use Google Colabs designed around different subject area datasets. While the mushroom dataset was used generically to introduce data analytics, the iris dataset was used to make connections to Biology and the elements dataset was used to make predictions related to Chemistry and Physics.

The majority of the PD activities featured a student-centered instructional approach in which teachers experience and reflect upon learning activities that they will ultimately lead for their students, a best practice in PD. After experiencing the activities, teachers reflected on the pedagogy we demonstrated as well as the places where the units might fit into their curricula, and critiqued formative assessments (daily exit tickets) that were designed to capture student learning outcomes.

Our primary strategy to support a variety of subject area teachers during the PD was to use an identical progression of activities and units across subjects but provide specialized questions and datasets aligned to different subjects. This strategy enabled us to run a single workshop covering all five units and provide breakout groups during some activities where subject-specific teachers could work with datasets aligned with their subject. In the two latter pilots we report upon in this paper, we used generic questions and datasets that were not specific to a subject area because we had two instructors per workshop and thus could not run separate breakout groups per subject due to the large number of subject areas represented, and insufficient numbers of teachers in each subset to form meaningful clusters. Another strategy to promote relevance was to make links to STEM fields and professionals whenever possible during the PD.

On each day of the five-day PD workshop we taught a full unit over the course of 5 hours. Synchronous collaborative work time and solo work time were offered such that participants could choose to work alone or in mixed teams within breakout rooms. Practice with making modifications to code was prompted through directed activities in the Colabs and supported by facilitators who moved between the breakout rooms and the main room.

The AIMSinDS curriculum and PD reflect that data scientists need to become familiar with numerous analytical and predictive methods and they must be able to choose the right method to use to solve the problem at hand with the available data. Thus, at the beginning and end of each unit, we discuss the uses and types of data required that differentiate the new method from the previously introduced methods. At times, we experimented with using the same dataset across multiple methods to explain that two or more methods could be used to answer the same question. Even then, we sought to discuss trade offs between the methods including processing time, accuracy, and adaptability.

Teacher Outcomes

Our findings on teacher PD outcomes were drawn from data collected from Cohort 1 and 2 teachers, including (1) exit tickets administered at the end of each day during the PD that asked questions about teachers' impression of the workshop activities, (2) an assessment administered before and after the workshop that assessed teachers' AI content knowledge, and (3) a pre/post-survey that examined teachers' attitudes toward AI and interest and excitement.

Teachers found the professional development highly engaging and participated fully. Cohort 1 and 2 teachers attended all 5 days of their respective workshops. In the daily exit tickets, more than 90% strongly agreed that the workshop's goals were clear, pacing was appropriate and the workshop day was worthwhile. When asked how much they felt their voices were heard and questions were answered, 100% strongly agreed. The teachers were less satisfied with a sense of community with roughly 50% strongly agreeing with the statement "I felt a sense of community among the participants."

Participants gained knowledge and skills through their participation. In an assessment of AI content knowledge administered before and after the workshop, teachers showed improvement from pre- to post scores. Across both cohorts the number of correct answers improved significantly from baseline to exit (from 50% to 64.4%, matched pair t-test, n=19, p-value=0.023). Interestingly, Cohort 2 teachers had higher mean scores than Cohort 1at baseline (12 vs. 10.1 points out of 22 respectively) and larger gains in overall scores. Cohort 2 teachers gained an average of 5.3 points whereas Cohort 1 teachers gained an average of 1.2 points. Participants showed marked improvement (5+ points) on items related to Decision Trees and Neural Networks, and lesser percentile gains in the area of Data Analytics. For the items related to Machine Learning, only Cohort 2 showed improvement. It is notable that Cohort 2 showed significant gains in the area of Neural Networks with a growth of 3.5 points. (See Figure 2 below.)

Another component of the pre/post-assessment is a set of six scenario-based questions designed to examine whether learners develop the ability to discern which AI method or technique to use given the question at hand and the data available. We note that this is a higher-order cognitive skill than learning about individual methods, but one we feel is important to address due to its relevant to the professional

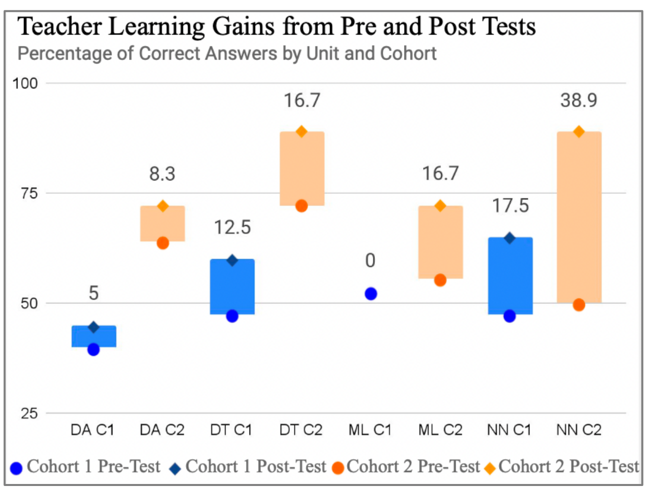

Figure 2. Teacher avg. percentile gains per unit by cohort.

practice of Data Scientists. Each scenario-based item provided a data set and a question to be answered then asked the respondent to pick the AI method they would be most suitable for answer the question. For example, one scenario was determining whether or not there was a stop sign in an image, and the dataset available was hundreds of images labeled as either having a stop sign or not. The choice of methods included Descriptive analytics, Decision tree, Clustering/KNN, and Neural Networks. Teachers' ability to discern which method to use varied greatly between the cohorts. In Cohort 1 we saw a decrease of 3.3% (from 43.3% to 40%) between baseline and exit while in Cohort 2 we saw an increase of 35.2% (from 38.9% to 74.1%) between baseline and exit. Further inspection by topic shows overall gains in the recognition of decision trees and neural networks in both cohorts, while Cohort 1 decreased its discernment on descriptive statistics and clustering/KNN by 15% and 10% respectively.

Teachers' attitudes toward AI and interest and excitement about learning and teaching AI were also captured in the pre- and post- survey. Overall, interest was high at baseline (average of 4.7 out of 5) and rose moderately (3%) with the greatest gain seen in the item: "I will talk to other teachers about what I learn about Data Science and Machine Learning." Excitement was similarly high at baseline and increased 5% in the item " I am excited to teach kids about Data Science and Machine Learning." Two items in the relevance of AI scale changed by -8% on the negatively framed items "I am uncertain why students need to learn about Data Science and Machine Learning" and " I don't interact with Data Science and Machine Learning on a regular basis" indicating an increase in teachers' perception of the relevance of AI in their own and their students' lives. The Anxiety toward AI items showed teachers had mixed emotions about AI. Teachers were less nervous and less worried about the impact of AI at exit than at baseline but they were also less certain that AI would make the world a better place.

In general, teachers' comfort with leading the activities decreased slightly as the units progressed. This was expected as the units progressed from topics most likely familiar to teachers (descriptive statistics) to ones most likely novel or obscure to them (Neural Networks). Across both cohorts, teachers' comfort with leading the lessons was high with averages ranging from 3.76 to 4.75 on a 5-point scale (from very uncomfortable to very comfortable). Teachers were least comfortable with the Google Colabs and most comfortable leading activities through hands-on activities such as "Be the Dataset" and online tools such as Google's Teachable Machine. Cohort 2 respondents had a higher comfort range (4.0 to 5.0) than Cohort 1 (3.64 to 4.71).

Suggested Improvements

While overall teachers felt engaged with the content, they also shared suggestions to further improve the PD workshop and curriculum. First, some teachers felt dumbfounded with some of the questions that were posed by the instructors. One teacher stated *"I think one thing that may help with engagement would be refining the questions you ask. There were many times throughout the week where I was engaged but didn't answer a question because I wasn't sure what it was really asking."* This points to a need to clarify the language and to be more specific on the intent of the questions we ask.

Second, several teachers expressed the need for more examples of integrating the units within a science class, and requested to use the subject specific materials; in particular they wanted to adapt the biology, chemistry and physics-specific versions of the Google Colabs, answer keys, and presentation slide decks for their classrooms.

Third, learning with Google Colabs was exciting but posed many challenges to teachers. One teacher felt more pre-learning about Colabs was needed prior to using them. Another wanted more instruction on Python codes and libraries as part of the lessons. Several teachers commented that the Colabs "could have been more user friendly" and that it would have been helpful if they had been more thoroughly debugged beforehand. The PD instructors also found Colabs difficult to maintain due to a lack of file protection; all viewers have write access and thus on occasion the content was accidentally corrupted when teachers forgot to make copies before editing. Overall, teachers expressed an interest in including more Colab activities within the curriculum.

Fourth, teachers expressed a need for further situating each new concept within real world contexts and applications. One teacher commented that "*The only thing I would change would be framing some of the concepts in a bigger picture or real world before we jumped into what something did. It usually became clearer as we progressed through a new concept but I'm a big picture learner who tends to be lost until I can see where we are going."*

We were encouraged by teachers' remarks that the pacing of the PD workshop was reasonable and "helpful to those with no or limited prior knowledge." One teacher remarked on the social learning aspect of the PD workshop stating: *"My favorite parts of the week were when we were able to explore and discuss activities in pairs or in small groups."* This statement reinforces the need to attend to providing opportunities for teachers to learn from each other and develop a supportive community among teachers embarking on this new endeavor.

Discussion

In this paper, we discussed an innovative curriculum and accompanying PD workshop for the integration of AI methods into STEM classrooms. Then we reported on initial findings on secondary school teachers' overall impressions of the PD, what they learned about AI, their shifts in attitudes toward AI, and their level of comfort with leading the lessons within their disciplinary classrooms.

Overall, the integration approach was well received by STEM teachers and framing of AI as a method in DS made it seem less esoteric and more applicable to a range of subject areas. Providing modular units that could be incorporated piecemeal into existing curricula and providing customized units per subject area provided teachers with choice and the discretion to select lessons based on their disciplinary context and their goals for student learning.

The design of the curriculum, informed by undergraduate students and the development team's prior work on DAILy as well as the author's previous work on Science+C, NM-CSforALL, and Project GUTS integration curricula, supported teachers who were novices in AI to learn fundamental concepts and gain skills in applying AI methods. The progression of units ensured that learners inspected and analyzed datasets in the first unit prior to using the dataset in subsequent units for the training and testing of AI models. Importantly, the inclusion of Google Colabs as interactive playgrounds enabled us to move beyond AI Literacy and offer a view into how DS and ML libraries are used to perform data analysis, visualization and AI model building in code.

While the PDs were well regarded by participants, there is an ongoing tension between providing subject specific units and offering differentiation by subject area within the PD workshops. While providing the subject specific units and associated Colabs and datasets is necessary to build strong disciplinary connections and increase teachers' buy in, providing PD for the various disciplines often requires more facilitators than are available. In the future we might consider offering a choice of either generic STEM or subject specific workshops.

Participants' learning outcomes and reactions to the curriculum and PD workshops signify the promise of an integration approach for offering AI learning experience within high schools. The notable differences between the two cohorts of teachers and their respective outcomes lead us to believe that a strong foundation in mathematics and prior exposure to CS concepts and professional programming languages may advantage some teachers over others. Furthermore, we found that the discernment between AI methods as a cognitive task proved to be well within reach of high school teachers. Irrespective of teachers' learning outcomes and core subjects they teach, all teachers expressed a great deal of comfort in leading the lessons.

Thus, we are encouraged that the progression of activities in the units and the overall breadth and depth of the curriculum are suitable for implementation in high school classrooms. Yet, one cannot assume that this curriculum and PD will support all teachers equally and further, the learning outcome findings suggest that explicit review of mathematical concepts and providing additional support to interpret the code within Google Colabs will be needed to help those without such background succeed.

Future Work

Our next steps are to respond to teachers' suggestions for improvement of the curriculum and PD workshops, such as clarifying the intent of the questions we ask, sharing detailed examples of integrating the units within a science class, situating each new concept in real world applications, increasing opportunities for collaborative learning, and generating and support a community of practice. Additionally, we hope to build new instructional materials in response to teachers' needs. Several teachers asked us to provide more Colab activities within the curriculum in particular, ones that teach the basics of model building from a constructionist perspective.

In the longer term, we plan to follow up with the participating teachers at the end of the semester and school year to learn which lessons they were able to implement: when and how, and to what effect. We also hope to determine which mathematical concepts are most salient in teachers' and students' learning of AI concepts at the high school level and how we can support learning of those concepts using an hands-on and playful approach.

Acknowledgments

We gratefully acknowledge the National Science Foundation STEM+C program support of grant #1934112, Computational Sciences Pathway Option for Massachusetts High School Students.The authors would like to thank Nisha Devasia, Claire Bao, Erica Zhou, and Grace Kim for co-designing the initial version of the curricular units and for their teaching assistance during early pilots. We are grateful for the feedback on the AIMSinDS curriculum provided by the EDC Science+C facilitator corps and pilot teachers. We also like to thank Susan Gibbs and Gabrielle Rabinowitz for facilitating the teacher PD workshops, and CS Alliance and CSforMA for sponsoring the two PD workshops described in this paper.

References

Ali, S; Payne, B. H.; Williams, R.; Park, H. W.; and Breazeal, C. 2019. Constructionism, Ethics, and Creativity: Developing Primary and Middle School Artificial Intelligence Education. In International workshop on education in artificial intelligence k-12 (eduai'19).

Ali, S.; DiPaola, D.; Lee, I.; Jackson, D. W.; Kiel, J.; Beal, K.; Zhang, H.; Cheng, Y.; and Breazeal, C. 2021. Adapting K-12 AI Learning for Online Instruction. 2nd International Workshop on Education in Artificial Intelligence K-12. *Proceedings of German Journal of Artificial Intelligence 2/2021*.

California Department of Education. 2021. Mathematics Framework. First Field Review Draft (Chapter 5, Data Science). Downloaded at https://www.cde.ca.gov/ci/ma/cf/

CC2020 Task Force. 2020. *Computing Curricula 2020: Paradigms for Global Computing Education*. Association for Computing Machinery, New York, NY, USA.

Cheng, Y.; Zhang, H.; Jackson, D. W.; Lee, I. A.; Brown, N. J. S.; Szendey, O.; Ali, S.; and DiPaola, S. 2020. Raising minoritized middle schoolers' AI career awareness and adaptability: Findings from two online summer camps. American Educational Research Association 2021

DiPaola, D.; Ali, S.; Hong, J.; Zhang, H.; Breazeal, C.; and Lee, I. 2020. Youth as Investigators of Bias in Artificial Intelligence. American Educational Research Association 2021

Druga, S.; Vu, S. T.; Likhith, E.; and Qiu, T. 2019. Inclusive AI literacy for kids around the world. Fablearn'19, March 2019, New York City, NY USA. https://doi.org/10.1145/3311890.3311904

Eguchi, A.; Okada, H.; and Muto, Y. 2021. Contextualizing AI Education for K-12 Students to Enhance Their Learning of AI Literacy Through Culturally Responsive Approaches. *Künstl Intell 35*, 153–161. https://doi.org/10.1007/s13218-021-00737-3

Gaible, E. and Burns, M. 2005. Using Technology to Train Teachers [Online]. Available from infoDEV: http://www.infodev.org/en/Publication.13.html

Greenwald, E.; Leitner, M.; and Wang, N. 2021. Learning Artificial Intelligence: Insights into How Youth Encounter and Build Understanding of AI Concepts. *Proceedings of the AAAI Conference on Artificial Intelligence*, *35*(17), 15526-15533. Retrieved from https://ojs.aaai.org/index.php/AAAI/article/view/17828

Hassel, E. 1999. Professional development: Learning from the best. Oak Brook, Illinois: North Central Regional Educational Laboratory.

Kim, S.; Jang, Y.; Choi, S.; et al. 2021. Analyzing Teacher Competency with TPACK for K-12 AI Education. *Künstl Intell 35*, 139–151. https://doi.org/10.1007/s13218-021-00731-9

Lao, N.; Lee, I.; and Abelson, H. 2019. A Deep Learning Practicum: Concepts and Practices for Teaching Actionable Machine Learning at the Tertiary Education Level. ICERI 2019 Conference. Retrieved from: http://appinventor.mit.edu/explore/research

Lee, I.; Ali, S.; Zhang, H.; DiPaola, D.; and Breazeal, C. 2021. Developing Middle School Students' AI Literacy. *SIGCSE '21:*

Proceedings of the 52nd ACM Technical Symposium on Computer Science Education. https://doi.org/10.1145/3408877.3432513

Lee, I. and Ali, S. 2021. The Contour to Classification Game. *Proceedings of the AAAI Conference on Artificial Intelligence*. 35 (17) 15583-15590. Retrieved from https://ojs.aaai.org/index.php/AAAI/article/view/17835

Long, D. and Magerko, B. 2020. What is AI Literacy? Competencies and Design Considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–16. https://doi.org/10.1145/3313831.3376727

Martin, F.; Lee, I.; Lytle, N.; Sentance, S.; and Lao, N. 2020. Extending and Evaluating the Use-Modify-Create Progression for Engaging Youth in Computational Thinking. In The 51st ACM Technical Symposium on Computer Science Education (SIGCSE '20), March 11–14, 2020, Portland, OR, USA. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3328778.3366971

National Research Council. 2012. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, DC: The National Academies Press. https://doi.org/10.17226/13165

National Research Council. 2013. *Next Generation Science Standards: For States, By States*. Washington, DC: The National Academies Press. https://doi.org/10.17226/18290

Sabuncuoglu, A. 2020. Designing One Year Curriculum to Teach Artificial Intelligence for Middle School. In *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education* (*ITiCSE '20*). Association for Computing Machinery, New York, NY, USA, 96–102. https://doi.org/10.1145/3341525.3387364

Sanusi, I. T. and Oyelere, S. S. 2020. Pedagogies of Machine Learning in K-12 Context. In 2020 IEEE Frontiers in Education Conference (FIE) (pp. 1-8). IEEE. https://doi.org/10.1109/FIE44824.2020.9274129

Sulmont, E.; Patitsas, E.; and Cooperstock, J. R. 2019. Can You Teach Me To Machine Learn? In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education (SIGCSE '19)*. Association for Computing Machinery, New York, NY, USA, 948–954. https://doi.org/10.1145/3287324.3287392

Williams, R.; Park, H. W.; and Breazeal, C. 2019. A is for Artificial Intelligence: The Impact of Artificial Intelligence Activities on Young Children's Perceptions of Robots. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Paper 447, 1–11. https://doi.org/10.1145/3290605.3300677

Yadav, A.; Hong, H.; and Stephenson, C. 2016. Computational thinking for all: pedagogical approaches to embedding 21st Century problem solving in K-12 classrooms. *TechTrends 60*, 565–568. https://doi.org/10.1007/s11528-016-0087-7

Zhang, H.; Lee, I.; Ali, S.; DiPaola, D.; Cheng, Y.; and Breazeal, C. 2021. Integrating Ethics and Career Futures with Technical Learning to Promote AI Literacy for Middle School Students: An Exploratory Study. IJAIED, in publication.