The Design of a Critical Machine Learning Program for Young Learners

Tolulope Famaye, Cinamon Sunrise Bailey, Ibrahim Adisa, Golnaz Arastoopour Irgens
tfamaye@clemson.edu, cinamob@g.clemson.edu, iadisa@g.clemson.edu, garasto@clemson.edu
Clemson University

Abstract: Machine Learning (ML) is integrated into many of the technologies we use daily. However, biased training datasets have shown to be harmful for marginalized populations. As future consumers and producers of technologies, children should have the technical and social expertise to engage with such issues in ML. In this study, we describe a series of activities designed for elementary and middle-school aged children to learn concepts of machine learning (ML), bias, and the sociopolitical implications of ML. Grounded in critical constructionism principles, we describe how children engage in reflective discussions, tinker with existing ML tools, and build an ML-based robot for social good.

Introduction

Machine Learning (ML) is integrated into many of the technologies we use daily. However, the socio-ethical implications that arise due to biased training data sets have shown grave implications, especially for underrepresented minorities. For example, while working as an MIT computer scientist, Joy Buolamwini, uncovered racial and gender bias in artificial intelligence (AI) services from companies such as Microsoft, IBM, and Amazon (Kantayya, 2020). With other numerous documented instances of ML bias (Piano, 2020), it is essential to equip youth, as future consumers and designers of technology, with the technical expertise and experience necessary to engage with ML. To broaden youth engagement with ML, a number of studies have developed programs and curricula that teach young learners basic ML concepts and operations (Chan, 2019). However, few studies have integrated AI ethics (Payne, 2020) and sociopolitical issues in educational programs. To address this gap, we co-designed with learners between 3rd-5th grade and implemented a critical machine learning (CML) program that attempted to integrate teaching concepts of ML with sociopolitical issues in ML. In this paper, we describe critical constructivist activities that children engaged in and provide examples of how youth integrated ML concepts with sociopolitical issues. This paper is a part of a larger study on critical constructivist activities designed to enable children to think critically about the sociopolitical implications of ML and design ML applications for social good (Arastoopour Irgens et al., 2022). This paper presents a description and analysis of critical constructivism and an application of its active-learning principles to developing a CML program for elementary and middle school youth in afterschool programs.

Designing for CML education with a critical constructivist lens

Our design and pedagogical approaches for facilitating children's critical engagement with machine learning (ML) were based on critical constructionist perspectives (Holbert et al., 2020; Kafai et al., 2020). The critical constructionist design framework promotes a learning environment that builds upon learners' lived experiences and provides tools that mediate learners' development of creative ideas and artifacts that reflect and challenge their lived realities. In this framework, educators connect concepts to personal and communal structures of inequities that shape the meaning and application of such concepts. Learners reflect, tinker with tools, and design futures that challenge the identified structures. Holbert et al. (2020) argue that engaging students in a cycle of connection and critical reflection about knowledge gives learners the opportunity to infuse their own perspectives and values into their creation. Moreover, learners play an agentic role in reducing or creating awareness about systems of inequalities in society and the critical nature of knowledge-building leads to a better understanding of the content (Holbert et al., 2020).

Methodology and CML program design

The activities described in this paper were designed and implemented with youth from two after-school centers in the United States consisting of Black, Latinx, and White children with a mix of those who presented as girls and boys and ranged between 9 - 13 years of age. The goal of the program was to provide children with learning experiences that enable them to understand ML bias, critically reflect on the harmful effects of ML bias at the systemic level, and design ML systems while attempting to mitigate bias. The activities, inspired by MIT's AI + Ethics Curriculum for Middle School and AI Ethics Education Curriculum (Payne, 2020), have been designed and tested in three iterations with forty-four children and three staff counselors from two separate after-school youth programs. Each iteration lasted approximately seven weeks and the activities with the youth occurred 2-3 days per week and lasted 1-2 hours each. Through these bi-weekly sessions, the researchers evaluated and amended the activities in light of the children's behaviors, conceptual grasp, direct and indirect input, and observed level of involvement (Arastoopour Irgens et al., 2022). The children's input, perceptions, values, and needs were taken into account and contributed to re-designs during implementations and between iterations. In this way, the children engaged in the co-design process as testers, and informants (Druin, 2002). We describe the set of activities implemented in the third iteration of the project following the changes that were implemented in the first and second iterations (for examples of the iterative re-designs in this study, see Bailey et al., 2023). The section that follows describes the CML program activities in light of the program goals: 1) engage children in reflective discussions on bias in ML algorithms, 2) engage children in 'tinkering' with existing ML tools with adult guidance, and 3) promote opportunities for children to use their gained knowledge and experience to design and build an ML-based robot for social good. Alongside the activities, we present qualitative data from interviews with children and researcher observation notes to provide some context of learners' engagement in activities.

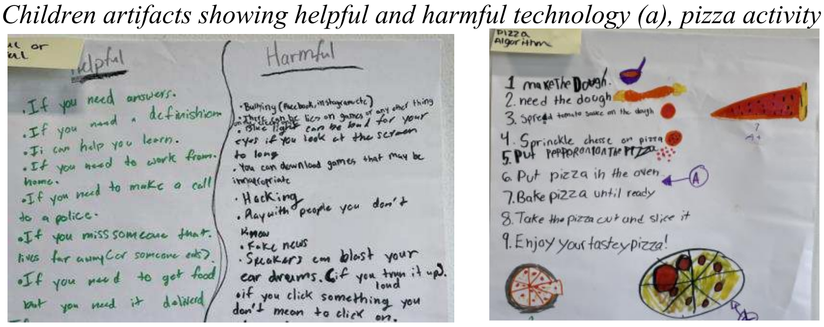
Engaging children in reflective discussions on bias in ML algorithms

The goals of the following discussion-based activities—*Harmful/Helpful Technologie*s, *Pizza Algorithm*s, and *Coded Bias*—were to guide children in critical reflections regarding the advantages and disadvantages of ML technologies, as well as understanding that algorithmic bias may exist within these technologies.

Harmful/Helpful Technologies: The children worked in groups of 3-5 participants to create a list of technologies they use. They were asked to document whether the identified technologies are helpful or harmful. Each group used large poster boards and markers to illustrate or list technologies within the categories of helpful or harmful (see Figure 1a). The youth worked in small groups and then in a whole group discussion with program facilitators.

Pizza Algorithm: We introduced the children to the concepts of algorithms and bias and the children worked in groups to write or draw a set of instructions on how to make the "best" pizza using markers and large posters hanging on the wall (see Figure 1b). They explored what represented the "best" pizza to each group and according to our observations, voiced how their preferences were reflected in their algorithms.

Figure 1
Children artifacts showing helpful and harmful technology (a), pizza activity (b)



(a) (b)

Coded Bias: Children watched the Coded Bias film trailer (Kantayya, 2020), which featured Joy Buolamini's realization of racist facial recognition technologies. Subsequently, children critically reflected on the concepts of bias, facial recognition technologies, and the socioethical implications of such bias with the researcher's guidance. In a group discussion, children responded to prompts such as: *What was interesting in this video? Who was being mistreated in this video and how? Who was creating the technology that was harmful?* During the discussion, one child commented: "*What I thought was interesting is a lot of people's life could be changed by, what the cameras [facial recognition software]..., on the streets pick up...*" Another child in the same group added, "*Just because of the software the person wasn't recognized to a certain, like house or where they live, they could get locked out of their house or they could be denied housing.* Statements made such as this one suggest engagement in the discussions around how technologies can be biased and the socio-ethical implications, such as AI making mistakes and people being denied housing.

Engaging children in tinkering with existing ML tools with adult guidance

The goals of the activities—*Google Search*, *Quick, Draw!*, and *Cat and Dog Teachable Machine*—were for children to tinker with tools and explore concepts of testing and training data, outputs, and algorithmic bias.

Google Search: To explore how algorithms have the creator's biases and opinions embedded within them, a researcher used a laptop and projector to guide the children through a group activity that required running a Google image search for words such as “basketball player.” The search returned images of men and the researcher asked: *Who might not be represented in this group? Why do you think we are seeing the images we are seeing? Who or what decides what we see?* During the discussion that followed this search, one girl stated “*If you look up basketball… you should be able to see both of them [men and women]. If you look up anything,* ***you should see women and basketball.****”* After completion of the program, another girl revisited the activity in her post-interview: *“Like, when we did the Google algorithm… When we searched up the basketball player, I liked how we could like to express our opinions on if it should be male or female.”* Here, the girl liked expressing her opinions about the output of the search algorithms. These girls felt comfortable in the space to provide their opinions and display a critical lens towards existing search algorithms.

Quick, Draw!: The children explored Google’s "Quick, Draw!" to see an example of how a neural network predicts in real time. The application randomly assigned six drawing prompts. The user draws the object and after 20 seconds the AI guesses the drawing by extracting from a database of drawings that other players had contributed to the application in the past. In their post-interview, when the researcher asked what children liked best about the entire program, one child stated “*Oh, the Quick Draw! I thought that was interesting.”* Another child answered, “*I really like the drawing thing like where you gotta draw something and then you had to guess it.”* These responses suggest they enjoyed the drawing and that this type of hands-on activity could initially engage children to participate in learning activities around ML.

Cat and Dog: Children built, trained, and tested ML models as they tinkered with Google’s Teachable Machine (GTM). This web-based tool allows users to develop simple classification models using images, audio, or video. Children were given photographs of both cats and dogs to use as a training and testing set for their models, but they were unaware that the researchers had given them a biased dataset that contained images of more cats than dogs. The goal was for children to discover the bias in the training dataset and then retrain their machines with new and less biased datasets. One child directly connected concepts of biased data sets from this particular activity with her final project. In her post-interview, she remarked, “*I don’t think my [final project] machine is biased because I tried to get the same amount of pictures so it could, like, make it even. So, it wasn't biased, like the cat and the dog one.”* Here she was comparing her final project with the cat and dog activity.

Using gained knowledge and experience to design an ML-robot for social good

The goals of the three final activities—*Build your own GTM, Superhero Robot Story, and Build Superhero Robot*—were to design an ML-based robot machine through storytelling and block-based programming. The children were given the challenge to identify issues in their communities and apply concepts of what they had learned in previous activities to create an ML robot for social good.

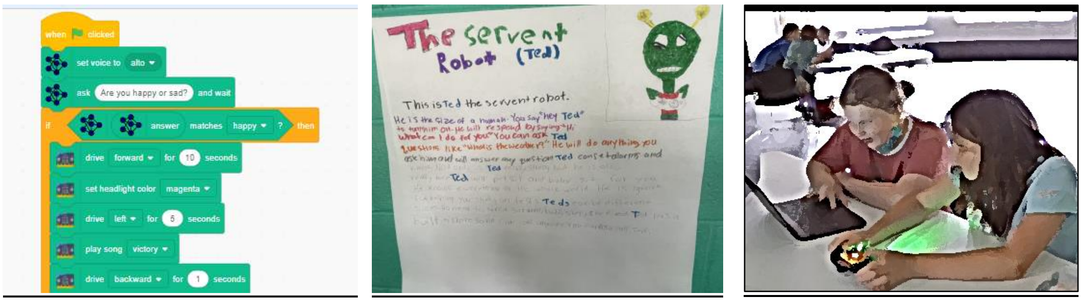
Build a GTM: After introducing the children to supervised ML in the cat and dog activity, children were challenged to create their own machine that recognized images, poses, or sounds using GTM. One child described her machine saying, *"I did a lot of pictures of guinea pigs and hamsters because they look alike. And I tested it… And it didn't work as well. So, I just added more pictures of different hamsters and different guinea pigs… I think 16 each… from the web. I put a lot of different pictures of different hamster and different guinea pigs and not just like one type of hamster and one type of guinea pig.”* Here, the child explained how she applied her knowledge of how to mitigate bias in ML algorithms by using a balanced training set of a variety of both hamsters and guinea pigs.

Superhero Robot Story: Children created narratives about robots that could help their community. They used large poster boards to visualize their imagined superhero robot. Afterward, they shared their Superhero Robots story with the group by responding to prompt questions such as: *What does your robot do? Who does your robot help? Why did you decide to design this robot?* Children told stories about robots that could contribute to social good in a way that they could relate to in their own lives. For example, when describing her robot design, one girl stated, “*My cousin, when she was growing up, she didn't have the opportunity to watch TV shows that teach her colors and stuff… So I thought to myself that could be happening to multiple other kids all over America. So, I thought, I could make a machine that can help kids with that."*

Build a Superhero Robot: Children worked in groups to build a minimal viable prototype of their superhero robot. By integrating the principles of training and testing data, the children built their ML models using GTM and linked their model with block-based programming (see Figure 2a) with a micro:bit robot. For example, one group described “Ted” as a helpful robot that answers questions and helps with household chores (see Figure 2b). Using this baseline story, the girls created a minimum viable version of Ted (see Figure 2c) by training a ML algorithm that combines with a micro:bit robot to ask a question: “Are you happy or sad?” The robot then responded to the answer by making movements, displaying lights, and playing a sound depending on the feedback.

In this activity children designed an artifact that incorporated their own perspectives and values while also practicing ML concepts and, in some cases, mitigating bias.

Figure 2
Children artifacts, a code block (a) superhero robot story (b), training and programming a robot (c)



(a) (b) (C)

Future study and redesign of the CML program

We observed children as they engaged in what Holbert and colleagues (2020) described as "cycle of connection and critical reflection about knowledge." Children critically contributed to discussions in activities like Coded Bias and Google Search and integrated their personal experiences while creating their superhero projects and building ML algorithms. They used knowledge gained from prior activities for building and evaluating their designs for bias, with some youth making reference to the earlier activities. The final activity; *Superhero Robot Story* helped us realize storytelling is a powerful way to help young learners engage actively with CML concepts and integrate a critical lens with their design projects (Chan, 2019). Thus, future work includes designing a narrative-driven role-play experience that incorporates ML bias activities but is a more cohesive set of activities and requires a critical lens for *every* activity. We believe these efforts are vital for children to develop a critical consciousness as the future generation of consumers and producers of technologies.

References

Arastoopour Irgens, G., Vega, H., Adisa, I., & Bailey, C. (2022). Characterizing children's conceptual knowledge and computational practices in a critical machine learning educational program. International Journal of Child-Computer Interaction, 34, 100541. https://doi.org/10.1016/j.ijcci.2022.10054

Bailey, C. S., Arastoopour Irgens, G., Famaye, T., & Mulholland, K. (2023). Using iterative participatory design to develop a machine learning bias program *with* youth *for* youth. Proceedings at the 2023 Annual AERA conference. Chicago, IL.

Chan, C. (2019). Using digital storytelling to facilitate critical thinking disposition in youth civic engagement: A randomized control trial. *Children and Youth Services Review 107*, (December 2019), 104522. DOI:https://doi.org/10.1016/j.childyouth.2019.104522

Druin, A. (2002). The role of children in the design of new technology. *Behaviour and information technology*, *21*(1), 1-25.

Holbert, N., Dando, M., & Correa, I. (2020). Afrofuturism as critical constructionist design: Building futures from the past and present. Learning, *Media and Technology*, 45(4), 328–344. https://doi.org/10.1080/17439884.2020.1754237

Kafai, Y., Proctor, C., & Lui, D. (2019). From Theory Bias to Theory Dialogue: Embracing Cognitive, Situated, and Critical Framings of Computational Thinking in K-12 CS Education. Proceedings of the 2019 ACM Conference on International Computing Education Research, 101–109. https://doi.org/10.1145/3291279.3339400.

Kantayya, S. (2020). *Coded bias* [Documentary]. New York: Netflix Studios.

Piano, S. L. (2020). Ethical principles in machine learning and artificial intelligence: Cases from the field and possible ways forward. *Humanit Soc Sci Commun,* 7(1), 1–7.

Payne, B. H. (2020). Can my algorithm be my opinion? An AI + ethics curriculum for middle school students. Thesis. *Massachusetts Institute of Technology.* Retrieved February 28, 2022 from https://dspace.mit.edu/handle/1721.1/127488