ChemAIstry: A Novel Software Tool for Teaching Model Training in K-8 Education

Fred Martin
fred.martin@utsa.edu
University of Texas at San Antonio
San Antonio, Texas, USA

Vaishali Mahipal
vaishali_mahipal@uml.edu
University of Massachusetts Lowell
Lowell, Massachusetts, USA

Garima Jain
garima_jain@student.uml.edu
University of Massachusetts Lowell
Lowell, Massachusetts, USA

Srija Ghosh
srijaghosh2015@gmail.com
University of Massachusetts Lowell
Lowell, Massachusetts, USA

Ismaila Temitayo Sanusi
ismaila.sanusi@uef.fi
University of Eastern Finland
Joensuu, Finland

ABSTRACT

Machine learning (ML) systems are increasingly in use in society. For young learners to be informed citizens and have full career potential it is important for them to understand these concepts. To support this learning, we created "ChemAIstry," an interactive software tool for children which demonstrates training and classification in machine learning. Students select which everyday items are safe to bring into a chemistry lab (e.g., a lab coat is safe; pizza is not). These selections serve as training input for a decision tree classifier. After training, students see how the trained model performs in classifying new objects. ChemAIstry was tested with 40 students aged 7 to 14 years at a public K–8 school. The software captured student selections during training. We analyzed these interactions to yield a "Correspondence Score," a measure of student understanding of the classification task. We screen-recorded student use of the software and audio-recorded our conversations with them during this use. Our analysis of these data indicates that students were able to understand the concept of model training, including that items were subsequently classified based on their training input. More than half of the student trials indicated that students correctly understood the task. This suggests ChemAIstry was effective in introducing students to these ideas in machine learning. We recommend continued development of related tools for curriculum integration of AI in K–8 education.

CCS CONCEPTS

• Social and professional topics → K-12 education; **• Applied computing → Interactive learning environments**; **• Computing methodologies →** *Supervised learning by classification.*

KEYWORDS

Machine Learning, Artificial Intelligence, Decision Trees, Training Model, Models, K–8 Students, Software Tools

SIGCSE 2024, March 20–23, 2024, Portland, OR, USA
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0423-9/24/03.
https://doi.org/10.1145/3626252.3630804

ACM Reference Format:
Fred Martin, Vaishali Mahipal, Garima Jain, Srija Ghosh, and Ismaila Temitayo Sanusi. 2024. ChemAIstry: A Novel Software Tool for Teaching Model Training in K-8 Education. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1 (SIGCSE 2024), March 20–23, 2024, Portland, OR, USA*. ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3626252.3630804

1 INTRODUCTION AND MOTIVATION

Artificial Intelligence (AI) is having an unprecedented impact on society as the amount of data and processing capacity are expanding quickly. The widespread deployment of AI in many different disciplines and industries emphasizes the need to develop a workforce with strong computing abilities and the capacity to work with AI [5]. To give future generations proper AI training and related skills, AI should be introduced into school curricula [7, 19].

Developing useful mental models for exploring AI is crucial for children. A key component of this is understanding how machines learn [6]. Researchers are developing many tools to introduce different AI concepts to K-8 students, including machine learning (e.g., [1, 14]), but few specifically teach model training.

Here, we developed a novel software product which introduces students to model training. Students supply the training data in a scenario related to safety in a chemistry laboratory. Students see how their own training data causes an automated system to make classifications. Our work addresses Learning, the third of the five "Big Ideas in AI" identified by the AI4K12.org project [21].

2 RELATED WORK

The growing reliance on emerging technology such as AI has heightened the call for young learners' inclusion in understanding its basic functionalities. AI has recently been recommended and included as a subject in schools across different regions [18, 20]. While the emerging subject continues to find its way into classrooms based on researchers and practitioners' agreement of its relevance for K-12 students [7, 19], more work is required regarding resources to bring AI into schools. Meaningful effort has been made with respect to creation of resources such as curriculum, tools, and pedagogical designs, including assessment, to popularize AI learning within K-12 learning settings [16, 22]. Owing to the developing nature of the subject area, more resources are required to indeed broaden participation and democratize AI literacy.

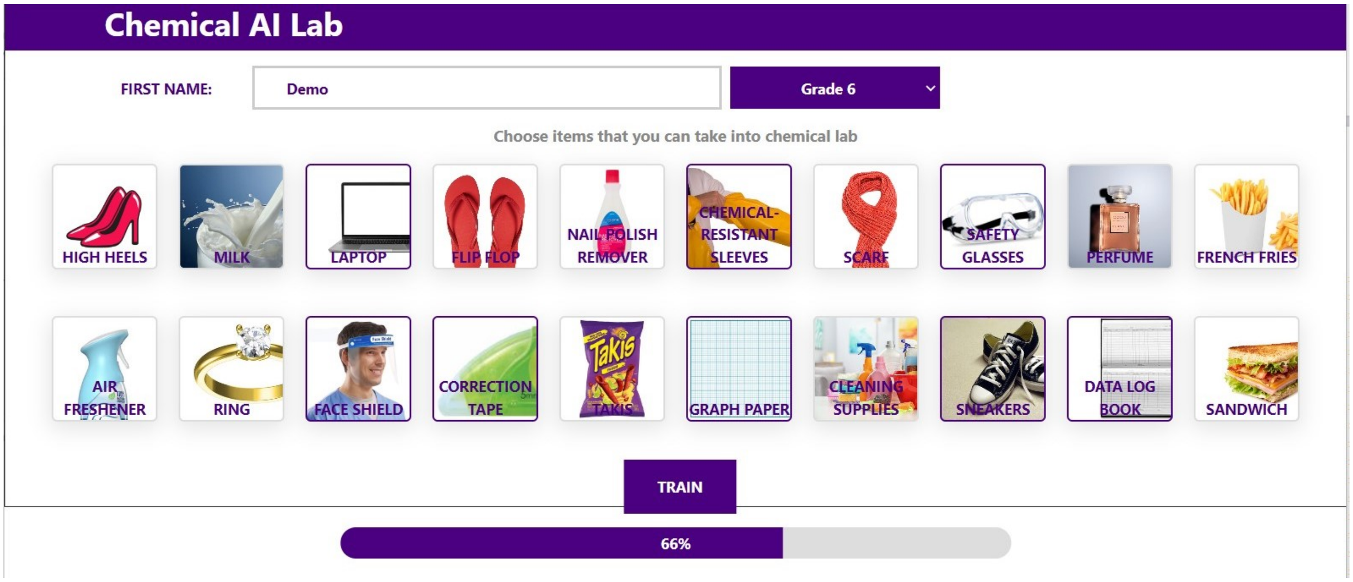

Figure 1: ChemAIstry Main Screen with Training Underway

Past studies have largely shown that software tools designed to learn AI and ML related concepts have been effective in promoting the knowledge of AI among K-12 students [4, 11]. To this end, software tools are being developed to teach different AI and ML related topics. Some systems that teach model training in ML include Google's Teachable Machine (GTM) [8], Machine Learning for Kids [2], LearningML [9], and Code.org's AI for Oceans [3].

These tools have been examined for their effectiveness in introducing the concepts for which they were designed. For instance, a study explored the use of GTM with young children in their homes and established children's ability to identify input-output relationships in AI technology [17]. An online program investigated how the creation of projects with LearningML impacts the knowledge of AI for students between 10 and 16 years; findings indicate that the initiative had a positive impact on participants' AI knowledge [15]. Mahipal *et al.* created Doodleit, a tool for teaching convolutional neural networks (CNNs), and demonstrated in a pilot study that middle school students were able to understand the functionality of kernels and feature maps involved in the CNN [11]. These findings across learning settings and modalities further reinforce the value of designing tools to demystify specific AI concepts to young students.

To contribute to the democratization of AI and inclusion of young learners in learning AI, we developed a tool, "ChemAIstry," that directly teaches about training in ML. The system is based on safety in a chemistry lab and was designed to introduce the concepts of model training in ML. While there are existing tools on ML as highlighted earlier, few tools specifically focus on model training and are situated within a STEM subject. This paper describes the tool's design and children's experiences learning with it.

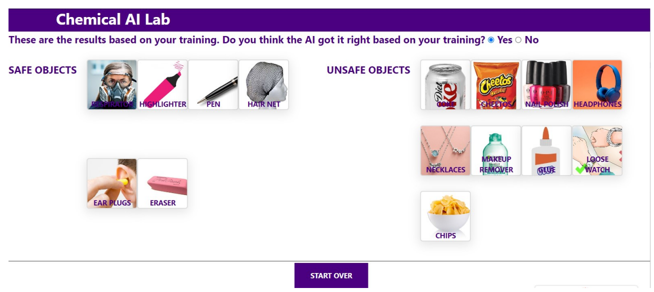

Figure 2: ChemAIstry Training Results Screen

3 TOOL DESIGN

ChemAIstry is an online tool designed to educate students on the principles of machine learning (ML) such as how computers are trained using data to make decisions. As as result of the domain, the tool also facilitates conversations with students about safety in chemistry laboratories.

3.1 User Experience

The user is presented with a screen with 20 items (Fig. 1). The user selects items they consider safe to take in a chemistry lab and then clicks the "TRAIN" button. After training, the user is shown the "Training Results" screen with a different set of 15 items which are automatically classified based on the training data (Fig. 2).

During the training stage:

- The app displays 20 objects with labels. The user clicks on the objects they think are safe to take into a lab. These 20 items are chosen at random from a dataset of 100 possible items. The app will present eight safe items and 12 unsafe items each time.

- Upon clicking on the "TRAIN" button, a progress bar proceeds for a few seconds to give the impression that the computer is doing computational work. In actuality, the training uses a simple decision tree algorithm and completes its process nearly instantaneously.

After training:

- The user is shown a results screen with 15 objects (Fig. 2). Based on the input from the previous screen, objects are classified by the trained ML model into two classes — safe and unsafe.
- The set of 15 objects is also randomly chosen from the dataset, excluding those that were previously presented for training.
- A question at the top of the results page asks the student to evaluate the performance of the trained AI model.

By randomly selecting the objects from the dataset, students encounter a different collection of objects each time they use the system. This approach helps them gain a better understanding of how the computer uses their information for training purposes.

For research purposes, the following student interaction data were automatically collected:

- When the "TRAIN" button was clicked, timestamped data were recorded: student first name; student grade level; and selection status (yes/no) for each of the 20 training items.
- On the Training Results screen, when the "START OVER" button was clicked, the student's answer to "Did the AI get it right?" question was recorded.

3.2 Implementation

To build ChemAIstry, a data set of 100 items to be classified was created. As the tool was implemented as a web app, the dataset was represented in JSON format. Each data point contains an item and five features: Flammable, Personal Protective Equipment (PPE), Food, Research Instruments, and Unsafe Wearable. Items belonging to the Research Instruments and PPE categories were considered as safe; items belonging to Flammable, Food, and Unsafe Wearable were considered as unsafe.

The items were divided so that 20 items belonged uniquely to each one of the five feature-categories. In our training data, that feature is coded as one and others as zero. An excerpt of the dataset is presented in Table 1 on the next page. For instance, Perfume is flammable, so this feature is coded as one and the others are coded as zero. The dataset contains 40 safe items and 60 unsafe items.

The web app was developed using HTML, JavaScript, CSS, and PHP. To train the model, we used the `ml`-cart package which provides a machine learning algorithm based on Decision Trees and using the CART implementation [13].

We used the decision tree machine learning algorithm owing to its simplicity and effectiveness of making decisions. The objective was to demonstrate to students how the computer assigns objects to specific categories and then further classify these categories as safe and unsafe objects. Fig. 3 shows a sample decision tree created by the trained machine learning model. This table was not displayed to the students.

ChemAIstry's source code is available at a GitHub repository [10]. The repository includes a link to a live browser-based demo.

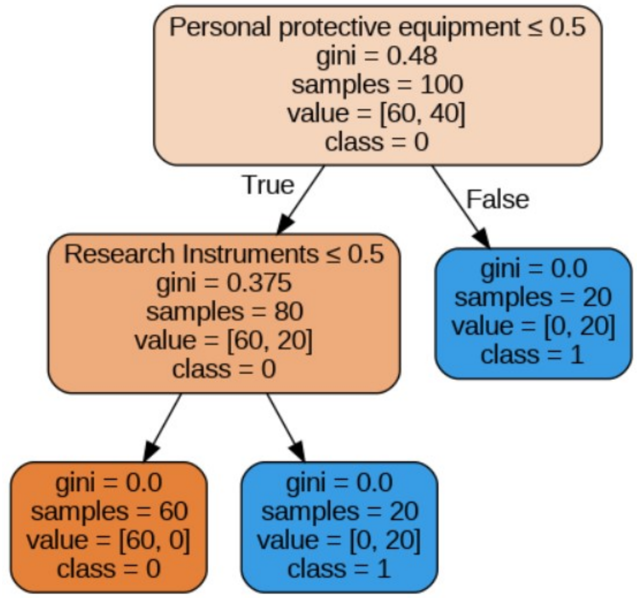

Figure 3: Decision Tree

The machine learning algorithm decided to first split the data set based on whether an item is Personal Protective Equipment. If it's not PPE (value less than 0.5 is true), it tests whether it is a Research Instrument (and so on). Otherwise, it is determined to be PPE and classified as safe, following the right-hand "False" arrow (class=1). The gini values represent the "impurity" of the items remaining to be classified; once the subset is fully classified, this value goes to zero.

4 PILOT STUDY

ChemAIstry, along with four other AI exhibits, was developed in Spring 2023 as part of a special-topics course offered by the first author at his university [12]. To test the work, we collaborated with a partner public K–8 school. The school and university are located in a gateway city in the Northeast region of the United States.

In five class groups over two days, a total of 125 children with ages ranging from 7 to 14 years old interacted with the tools. Each class group had a class period (about 45 minutes) to interact with the five AI tools. Each tool had two university students or faculty members dedicated to helping the children. Facilitation and assistance was also provided by school teachers and staff. IRB permission to conduct the study and parental consent and student assent was obtained. ChemAIstry was used by 40 students during these sessions.

4.1 Student Use

The tool was introduced to children with the explanation of them acting as teachers while the tool played the role of the student. Therefore, whatever the student (as a teacher) taught the machine, the machine would learn accordingly. It would be up to us how we taught the machine, similar to the important role a teacher plays in a student's life, teaching them what is right and wrong. Similarly, the students would teach the machine which items were acceptable to carry and which were not acceptable to carry in a science or

Table 1: Excerpt of ChemAIstry Dataset

Object	Flammable	Personal protective equipment	Food	Research Instruments	Unsafe Wearable Things
Perfume	1	0	0	0	0
Cookies	0	0	1	0	0
Lab Coat	0	1	0	0	0
Lab Manual	0	0	0	1	0
Open-toed shoes	0	0	0	0	1

chemistry laboratory. The term "science" was used for 1st graders as they were not aware of chemistry, and "chemistry" was used with the students in grades six, seven, and eight.

After being introduced to the ChemAIstry, the students were given the opportunity to further test it for themselves and assess whether or not computer had learned effectively. After training, students were prompted to indicate whether they thought the "AI got it right based on your training?"

4.2 Data Collection

As the students interacted with the tool, we recorded their screen activity and the audio of our interactions with them as well as the conversations among the groups of students. We also collected data from the tool itself: first name and grade of the student; items selected as safe objects; items that were not selected as unsafe items; and the response to the question on the results page about performance of trained model. After tool use, we conducted a brief paper-based survey and structured interview with the students.

5 DATA ANALYSIS

We used quantitative and qualitative methods to analyze student learning. Our quantitative approach is based on the definition of a "Correspondence Score," a metric created to indicate students' correctness in selecting safe and unsafe items during training. Our qualitative approaches are based on interpreting students' statements during their engagement with our tool.

5.1 Correspondence Score

A Correspondence Score was defined to indicate student understanding of the classification task. During training, 20 items were displayed on the screen: eight safe items and twelve unsafe items. A maximum score of 20 is possible, representing that the student correctly identified all items, by selecting the eight safe ones and not-selecting the twelve unsafe ones.

One point is given for a correctly selected item—an item which is coded as safe and was also selected by the student while training/teaching the machine. A point is also given for an unsafe item that the student did *not* select. A point of zero was given if a student selected an item coded as unsafe. Similarly, a point of zero was given when a student did not select an item coded as safe. Thus:

$$CorrespondenceScore = \sum_{i=1}^{20} (\text{selected} \wedge \text{safe}) \vee (\neg\text{selected} \wedge \neg\text{safe})$$

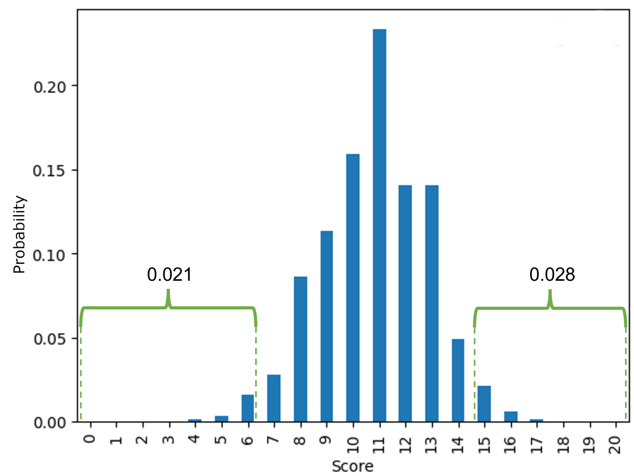

Figure 4: Monte Carlo Simulation of Random Training
The graph shows the distribution of expected Correspondence Scores if items were selected randomly during training. The left-bracketed tail, with 0.021 probability due to chance, indicates deliberate "reverse training" (selecting unsafe items on purpose). The right-bracketed tail, with 0.028 probability due to chance, indicates intentional correct training.

5.2 Monte Carlo Simulation

We used the Monte Carlo method to evaluate the likelihood of students' training performance being a result of intentional understanding of the task or as due to chance. We simulated the outcomes of classifying items as safe or unsafe if selected randomly.

The unique counts of items picked by students were tabulated. This frequency distribution played an important role in determining how many times each scenario was simulated. Based on the frequency distribution, the Monte Carlo simulation was executed. For instance, if in four of the 95 trials a participant chose exactly eight items, then in the simulation, eight random selections were made in four of every 95 simulated runs. This mirrored the actual distribution of student choices, ensuring that the simulation was representative of real-world behaviors. We implemented the simulation in Python and ran approximately one million trials.

The result of the simulation is shown in Fig. 4. The analysis revealed that 2.09% of the simulated trials produced Correspondence Scores ranging from 0–6, while 2.84% resulted with Scores between 15–20. The cumulative percentage of these tails is 4.93%. From a statistical perspective, these scores, being on the extreme ends of the distribution, are less likely to have occurred by chance. Thus, with

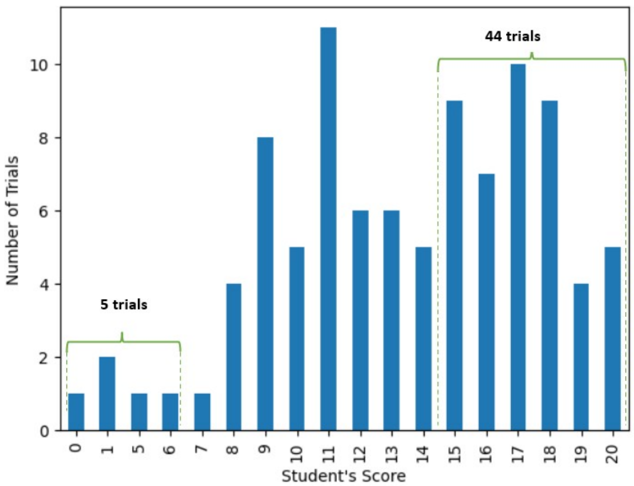

Figure 5: Student Correspondence Scores
Students conducted a total of 95 trials in our study. Of these, five trials (5.3%) had highly unlikely low scores, indicating students were intentionally "reverse training"—choosing known unsafe items as safe. 44 trials (46%) represented statistically unlikely high scores, indicating understanding of the task and correct training. The remaining 46 trials (48%) fell within the likely-due-to-chance scoring range, indicating that the students conducting them did not yet understand the intention of the tool.

95% confidence, students with such scores exhibited a intentional understanding of their activity.

5.3 Student Results

There were a total of 95 trials by 40 students as depicted in Fig. 5. Of these, 49 trials fall within the 95% likelihood of not being due to chance:

- 5 trials (5.3%) scored in range of $[0, 5]$—Statistically unlikely low scores which indicate intentionally doing it wrong to see the results of a "reverse training"
- 44 trials (46%) scored in range of $[15, 20]$—Statistically unlikely high scores which indicate that students understood the scenario and intentionally selected safe items

The remaining 46 trials (48%) scored in range of $[6, 14]$, indicating that the students conducting these trials did not (yet) understand the task. Thus, we can report that in more than half of our trials, we are confident that students did understand the purpose of ChemAIstry, and were able to properly train the model with safe items.

5.4 Vignettes of Student Use

We analyzed the students' behavior from the screen and audio recordings that we collected. We present representative conversations between students and between students and researchers, showing how students engaged with the task.

Conversation 1. A group of two students were having a conversation about the things that are safe to bring in a chemistry lab. They began by discussing sunscreen. Student 1 remarked, "I mean sometimes." Student 2 asked, "What if it catches on fire though?" They decided that sunscreen was not safe.

Then they discussed a highlighter pen. Student 1 asked, "What if you need a highlighter to highlight some points?" They agreed that would be a good idea and then they selected following as safe items: calculator, sticky notes, lab coat, ruler, safety glasses, and highlighters. They got a score of 18 of 20.

After the training, they analyzed the results properly and made remarks like "yeah you need sneakers"; "[The pencil] sharpener could be useful"; and "That looks great."

While answering the question, "These are the results based on your training. Do you think AI got it right based on your training?", Student 1 answered "I say yes" and the other student selected the Yes option.

A research member asked the group to deceive the machine by training it incorrectly. The same pair of students selected the following items as safe: sunscreen, aerosol spray, sunglasses, lighter, necklaces ("Definitely, not supposed to have necklace"), apple, sandals, peanut butter, nail polish, pizza. They got a score of 2.

The students analyzed the results properly and made remarks for items that were predicted as safe items like gasoline, gum—"nope"; "yeah the bad ones." When we asked, "Does this makes sense, what it is showing and tell me what it is doing?" Their response was "Yeah, it analyzes the object that you pick to see which is helpful in a chemistry lab or something, yeah that's about it."

The same pair of students did one more iteration and this time selected all the safe items correctly and got a score of 20 and looked at the results closely and when asked what do you think about it, one student answered "Well, in my opinion this is pretty cool."

Conversation 2. One of the students got a score of 16. His choices demonstrated a clear comprehension of the activity's objectives. He was selective in his decisions, opting for items he was confident were appropriate for lab use and avoiding those he was unsure about. When presented with a challenge by our research member to train with a flawed dataset, the student's enthusiasm was palpable. He exclaimed, "I got you," and proceeded to consistently select items that were clearly unsuitable for a lab environment, such as makeup remover, nail polish remover, and various food items. With these incorrect choices, his consistency in selection demonstrated an understanding of the task's altered goal. He performed this three times, showing persistence in his learning process.

Some other snippets of conversations between students while selecting items included: for hair cap—"No, they are safe"; for surgical hoods—"It is like something, you know when doctors going to surgery, they have to put that over their hair so that it doesn't get in their body/hair"; for open toed shoes—"Naah"; for aerosol spray—"I don't know what that is." Screen activities like clicking/unclicking the items, hovering the mouse over each item while selecting, going back and forth and taking time while selecting the items suggests that they were choosing the items on purpose.

Remarks like "That makes sense" and "That's cool" were made sometimes while analyzing the results page that suggested they recognized the computer was making decisions based on their selection. These examples illustrate the range of understanding and decision-making skills among the children, as well as the potential for learning from peers.

5.5 Post-Experience Interviews

After students used ChemAIstry, we administered a written post-survey where we asked them about the performance of the tool and how they see AI as being present in their lives. While students were filling out the survey, we encouraged students to reflect on their role as AI trainers and consider the ethical implications of AI.

Particularly for the older students, this interaction provided a deeper understanding of their perspectives on AI. Many students recognized that "Google," "Siri," and their phones were examples of AI in their life.

When considering ethical implications of AI, several students made insightful remarks. One student highlighted the importance of correct and ethical training in AI, stating that ChemAIstry "should learn properly otherwise it may get hacked." Another noted that their phone, laptop, and electronics all consisted of AI, and that "more people will use AI once they understand [its] abilities."

When students were asked if machines are excellent students, one student responded, "It depends. They are good learners but it depends on the person who is training and what they are teaching if it is good/ bad," underscoring the role of the trainer in shaping AI outputs. Another student noted, "Kind of. They give us wrong information too. So more training data would have made it better with giving the information," demonstrating an understanding of the relationship between training data and the accuracy of the AI model.

6 DISCUSSION

Through ChemAIstry, students learned to train an ML model, using it to distinguish between safe and unsafe items to bring into a laboratory. Students across grade and age levels exhibited strong engagement.

Students experimented with the tool, training the AI with both correct and incorrect information and observing the varied outputs. They recognized that the AI model's output was largely consistent with their training, thereby reinforcing the principle that the quality of an AI model's output is directly influenced by the quality of its training.

One of our youngest participants (aged 7 years old) initially interacted with the tool in a less structured way, clicking randomly and observing the changes in results. However, the student made connections between their choices and the AI model's output. When asked if machines are good learners, the student agreed. Another student pointed out that the machine had correctly classified food items as safe, which they had specifically taught the machine.

Conversations revealed that for some students, ChemAIstry sparked thoughtful reflections on the ethics of AI. Students understood their role in shaping the AI model's behavior, highlighting the importance of ethical considerations in AI training.

The tool presented a promising approach to introducing complex AI concepts to students in an engaging and accessible way.

7 LIMITATIONS

This is an exploratory study. Results need to be confirmed by additional studies, including working with more students to increase the sample size.

8 CONCLUSION AND FUTURE DIRECTIONS

ChemAIstry was developed to introduce students to the concept of model training in AI. A "Correspondence Score" was defined to measure student understanding of the task. A Monte Carlo simulation informed by the student data shows that selections made in 49 out of 95 trials were intentional and showed student understanding of the task. Our qualitative analysis from students' screen activity and their conversations with us also suggests that students were able to understand how computer made decisions based on the items selected by them during training. The conversations also displayed their enthusiasm and excitement while using the tool.

For future work, ChemAIstry can be extended to domains beyond safety in chemistry lab. ChemAIstry could become a toolkit where students import their own material and create training tools for each other. The decision tree that is created after training could be presented to students so that they could see how the items they selected were assigned to different categories to perform classification.

More generally, ChemAIstry points towards a genre of software "toys" that allow students to playfully engage with the core ideas of machine learning: training and automated classification. These tools can be integrated into existing school curricula, expanding access and equity. As more systems like this are created and become available, we will be able to introduce more learners to machine learning and related ideas, improving education broadly.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant IIS-2112633. We thank the children who participated in this study and the teachers and administrators who facilitated our work at our partner school.

REFERENCES

[1] Michelle Carney, Barron Webster, Irene Alvarado, Kyle Phillips, Noura Howell, Jordan Griffith, Jonas Jongejan, Amit Pitaru, and Alexander Chen. 2020. Teachable machine: Approachable Web-based tool for exploring machine learning classification. In *Extended abstracts of the 2020 CHI conference on human factors in computing systems*. 1–8.
[2] Machine Learning for Kids. 2023. https://machinelearningforkids.co.uk/
[3] AI for Oceans. 2023. https://code.org/oceans
[4] Christiane Gresse von Wangenheim, Jean CR Hauck, Fernando S Pacheco, and Matheus F Bertonceli Bueno. 2021. Visual tools for teaching machine learning in K-12: A ten-year systematic mapping. *Education and Information Technologies* 26, 5 (2021), 5733–5778.
[5] Irene Lee, Safinah Ali, Helen Zhang, Daniella DiPaola, and Cynthia Breazeal. 2021. Developing Middle School Students' AI Literacy. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education* (Virtual Event, USA) *(SIGCSE '21)*. Association for Computing Machinery, New York, NY, USA, 191–197. https://doi.org/10.1145/3408877.3432513
[6] Phoebe Lin, Jessica Van Brummelen, Galit Lukin, Randi Williams, and Cynthia Breazeal. 2020. Zhorai: Designing a Conversational Agent for Children to Explore Machine Learning Concepts. *Proceedings of the AAAI Conference on Artificial Intelligence* 34, 09 (Apr. 2020), 13381–13388. https://doi.org/10.1609/aaai.v34i09.7061
[7] Duri Long and Brian Magerko. 2020. What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–16.
[8] Google Teachable Machine. 2023. https://teachablemachine.withgoogle.com/
[9] LML Artificial Intelligence made easy. 2023. https://web.learningml.org/en/home-spanish-en-translation/
[10] Vaishali Mahipal, Srija Ghosh, Garima Jain, Ismaila Temitayo Sanusi, and Fred Martin. 2024. *ChemAIstry GitHub repository*. https://github.com/engaging-computing/ChemAIstry
[11] Vaishali Mahipal, Srija Ghosh, Ismaila Temitayo Sanusi, Ruizhe Ma, Joseph E Gonzales, and Fred G Martin. 2023. DoodleIt: A Novel Tool and Approach for

Teaching How CNNs Perform Image Recognition. In *Proceedings of the 25th Australasian Computing Education Conference*. 31–38.

[12] Fred Martin, Saniya Vahedian Movahed, James Dimino, Andrew Farrell, Elyas Irankhah, Srija Ghosh, Garima Jain, Vaishali Mahipal, Pranathi Rayavaram, Ismaila Temitayo Sanusi, Erika Salas, Kelilah Wolkowicz, and Sashank Narain. 2024. Perception, Trust, Attitudes, and Models: Introducing Children to AI and Machine Learning with Five Software Exhibits. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education* (Portland, OR, USA), Vol. 2. Association for Computing Machinery, New York, NY, USA, 2 pages. https://doi.org/10.1145/3626253.3635512

[13] npm. 2022. *ml-cart-npm*. https://www.npmjs.com/package/ml-cart?activeTab=readme

[14] Juan David Rodríguez-García, Jesús Moreno-León, Marcos Román-González, and Gregorio Robles. 2020. Introducing artificial intelligence fundamentals with LearningML: Artificial intelligence made easy. In *Eighth international conference on technological ecosystems for enhancing multiculturality*. 18–20.

[15] Juan David Rodríguez-García, Jesús Moreno-León, Marcos Román-González, and Gregorio Robles. 2021. Evaluation of an online intervention to teach artificial intelligence with learningml to 10-16-year-old students. In *Proceedings of the 52nd ACM technical symposium on computer science education*. 177–183.

[16] Ismaila Temitayo Sanusi, Solomon Sunday Oyelere, Friday Joseph Agbo, and Jarkko Suhonen. 2021. Survey of resources for introducing machine learning in K-12 context. In *2021 IEEE Frontiers in Education Conference (FIE)*. IEEE, 1–9.

[17] Ismaila Temitayo Sanusi, Kissinger Sunday, Solomon Sunday Oyelere, Jarkko Suhonen, Henriikka Vartiainen, and Markku Tukiainen. 2023. Learning machine learning with young children: exploring informal settings in an African context. *Computer Science Education* (2023), 1–32.

[18] Jiachen Song, Linan Zhang, Jinglei Yu, Yan Peng, Anyao Ma, and Yu Lu. 2022. Paving the Way for Novices: How to Teach AI for K-12 Education in China. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 36. 12852–12857.

[19] David Touretzky, Christina Gardner-McCune, Cynthia Breazeal, Fred Martin, and Deborah Seehorn. 2019. A year in K-12 AI education. *AI Magazine* 40, 4 (2019), 88–90.

[20] David Touretzky, Christina Gardner-McCune, Bryan Cox, Judith Uchidiuno, Janet Kolodner, and Patriel Stapleton. 2022. Lessons Learned From Teaching Artificial Intelligence to Middle School Students. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 2*. 1371–1371.

[21] David Touretzky, Christina Gardner-McCune, Fred Martin, and Deborah Seehorn. 2019. Envisioning AI for K-12: What Should Every Child Know about AI? *Proceedings of the AAAI Conference on Artificial Intelligence* 33, 01 (Jul. 2019), 9795–9799. https://doi.org/10.1609/aaai.v33i01.33019795

[22] Qi Xia, Thomas KF Chiu, Min Lee, Ismaila Temitayo Sanusi, Yun Dai, and Ching Sing Chai. 2022. A self-determination theory (SDT) design approach for inclusive and diverse artificial intelligence (AI) education. *Computers & Education* 189 (2022), 104582.