Let's Go for a Drive: Exploring AI's Societal Impact in K-8 Education with an Interactive Self-Driving Car Tool

Pranathi Rayavaram
University of Massachusetts Lowell
Lowell, MA, USA
nagapranathi_rayavaram@student.uml.edu

Sashank Narain
University of Massachusetts Lowell
Lowell, MA, USA
sashank_narain@uml.edu

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

***Abstract*—This innovative practice full paper discusses the development of an interactive tool designed to educate middle school students on the ethical considerations and societal impacts of artificial intelligence (AI). As AI technologies become more embedded in our daily lives, the younger generation must grasp the implications of algorithmic bias and its societal effects. Our tool aims to deepen this understanding by focusing on self-driving cars—a relevant and significant example of AI technology. The interactive tool incorporates three advanced image recognition models trained on diverse datasets, including traffic cones, animals, and pedestrians. Through the tool's interface, students can choose one of these models to apply in a self-driving car simulation and select different obstacles for the car to encounter, such as traffic cones, animals, and pedestrians. This hands-on simulation highlights the importance of comprehensive AI model training, showcasing how well-trained models help avoid collisions and the risks associated with encountering untrained obstacles. It engages students by demonstrating how developers' AI training decisions can significantly influence end-user experiences. Moreover, the tool emphasizes the need for diverse and representative data in building fair and robust AI systems. To assess the effectiveness of this educational tool, we conducted a two-day AI exhibit attended by 26 middle-school students from grades six to eight. The effectiveness was evaluated through post-trial questionnaires to measure the students' understanding of several key concepts: the development of resilient AI models, the societal impacts of AI, and the ethical considerations of road safety in the context of AI. The results showed that 84.6% of the participants understood how poor training decisions could impact AI outcomes, and about 96% recognized the necessity for diverse data.**

***Index Terms*—Artificial Intelligence Education, Interactive AI Tools, AI Societal Impact, Self-Driving Cars Simulation**

I. INTRODUCTION

The advent of artificial intelligence (AI) offers significant potential to revolutionize industries and enhance human life. However, it also raises important concerns about fairness and safety. Currently, many AI applications exhibit biases and discriminatory behaviors, leading to notable adverse impacts on society [1]–[4]. This is particularly evident in the field of self-driving technology. A US National Highway Traffic Safety Administration report documented 392 accidents involving semi-autonomous driving systems and driver assistance technologies between 2021 and 2022 [5]. Determining liability for injuries caused by these evolving AI algorithms is a complex challenge that implicates both developers and end-users. Consequently, people of all ages must understand the wider societal implications of AI technologies.

In response, recent studies highlight the importance of teaching children in grades six to eight about the responsible use of AI. Educational initiatives such as AI4K12 [6] are designed to provide young people with a strong ethical foundation for responsible AI. Significant progress has also been made in educating children about AI ethics and biases [7]–[9]. However, exploring the broader societal implications of AI is still an emerging area. Notable work in this field includes Zhorai, which investigates the societal dimensions of AI-driven conversational agents [10]. Additionally, programs that involve students in activities such as stakeholder analysis and the redesign of platforms are being developed, such as reimagining YouTube with societal considerations in mind [11]. Despite these advancements, many initiatives tend to focus only on the development aspects or end-user impacts, which may not fully convey to students how decisions by developers directly affect user experiences.

Our objective with this effort is to bridge the knowledge gap for children regarding AI's societal impacts, examining both developer and end-user perspectives. We have open-sourced this educational tool to encourage integration into existing K-12 curricula and support researchers in developing further educational modules for comprehensive AI education [12]. Our tool uses self-driving cars as a case study, capitalizing on their relevance in the rapidly evolving AI landscape [13] and their ability to engage young learners' curiosity. In this tool, children take on the roles of both developers and end-users. They learn how the choices made in training machine learning models influence the experiences of those who use AI technologies. The tool's design encourages responsible AI usage and sheds light on developers' decision-making processes. It enhances interaction by allowing children to choose from three distinct AI models: the first model recognizes only traffic cones, the second both traffic cones and animals, and the third traffic cones, animals, and pedestrians. We intentionally designed these models to be simple to help students clearly

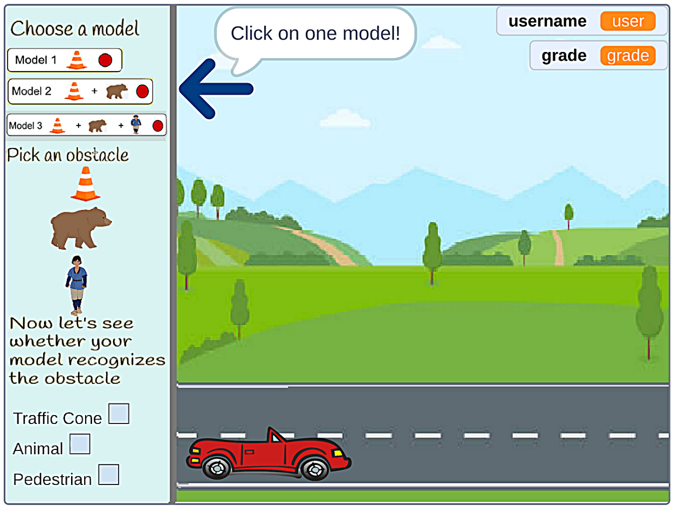

Fig. 1. Interface of the Self-driving AI Education tool.

grasp the societal impact of AI and encourage them to reflect on its real-world consequences. Students interact with the tool by selecting a model—playing the role of a developer—and then choosing an obstacle—acting as an end-user—such as a traffic cone, animal, or pedestrian. The tool demonstrates how self-driving cars with the selected AI model perceive and respond to the chosen obstacles through visual aids and a user-friendly interface (see Figure 1). By offering these choices, we aim to actively involve children in learning, helping them intuitively understand how each AI model operates and its broader societal implications.

We evaluated the effectiveness of our tool during an AI exhibit that involved 26 middle school students from grades 6 to 8. Each student engaged with the tool, experimenting with different models and obstacles to observe the simulated behavior of the self-driving car. To assess the educational impact of the tool, we conducted a survey that measured students' understanding of several key concepts: the broader societal consequences of AI, ethical considerations in AI development, and the importance of using diverse data to build robust machine learning models. The survey results were promising: 84.6% of the students demonstrated an understanding of how inadequate training data can negatively affect AI outcomes, and approximately 96% acknowledged the necessity of incorporating diverse data sets. These findings confirm that the tool effectively conveys critical perspectives from both developers and end-users. It highlights the importance of ethical and equitable model development, promotes responsible AI use, and warns against over-reliance on AI systems.

In summary, our contributions are as follows:

- We developed an interactive educational tool based on a self-driving car scenario, which helps children grasp the societal impacts of AI and the roles of both developers and end-users.
- We conducted a systematic evaluation with 26 middle school students from grades 6 to 8, demonstrating that our tool effectively communicates essential AI concepts to this age group.

The structure of this paper is as follows: Section II reviews related work; Section III details the design and implementation of our educational tool; Section IV outlines the user study design; Sections V to VII discuss the evaluation of the tool and the experiences of the students; and finally, Section VIII concludes our paper.

II. Related Work

A substantial body of research focuses on AI education, ethics, bias, and societal impacts within K-12 education [6], [14]–[17]. This section explores key studies and initiatives that are particularly relevant to our approach.

Significant advancements in AI curricula development for K-12 education have been made by various initiatives. The MIT Media Lab [18] has contributed notably by providing open-access curricula focused on AI, specifically designed for middle school students and teachers. Sabuncuoglu [19] has created a detailed one-year curriculum on AI and ethics tailored specifically for middle schoolers. Long and Magerko [8] have extracted design considerations from the literature, forming a conceptual framework that aids AI developers and educators in curriculum design. Additionally, Lee et al. [7] have implemented a 30-hour workshop introducing AI and machine learning concepts to children.

Nonprofit organizations like Code.org [20] have also dedicated significant efforts to increase student engagement in AI education. They provide activities such as "AI for Oceans," where students classify oceanic objects, train machine learning models, and connect these tasks to broader AI concepts. Additionally, platforms like Elements of AI [21] and OKAI [22] offer web-based interactive resources, including visual exercises, to aid in learning AI principles. These platforms use visualizations and animations in their activities to improve understanding and engagement with AI principles, making the learning process both educational and engaging.

Several initiatives have concentrated on enhancing AI education through visual and interactive tool development. Pang et al. [23] have developed a curriculum and toolkit specifically designed to explore the societal impacts of AI. Zumi AI [24] contributes to this field by offering a self-driving car robot toolkit, which allows students to program the robot using either Scratch or Python, providing a hands-on experience with concepts from the autonomous vehicle industry. Furthermore, Zhorai [10] focuses on deepening understanding of the ethical issues related to natural-language-based conversational agents by facilitating structured discussions.

Web-based drag-and-drop platforms such as Cognimates [25], Machine Learning for Kids [26], and Google Teachable Machine [27] offer visual environments that enable children to build AI models without needing coding or machine learning experience. Additionally, CreateML [28], specifically designed for Macintosh users, allows for creating custom image and text recognition models. These platforms are instrumental in supporting the development of tools like ours, providing interactive and user-friendly methods for training models and integrating them with programming languages.

Our educational tool leverages the widely adopted Machine Learning for Kids platform [26] and the block-based programming environment Scratch [29] to develop a simulated self-driving car tool that is both accessible and cost-effective. Unlike physical AI educational tools such as Zumi [24], which can be expensive, our web-based solution integrates seamlessly into existing K-12 curricula, making it an ideal resource for educational institutions. This approach allows students to directly experience the consequences of AI training decisions through an interactive, user-friendly interface, enhancing their understanding of AI's impact. The design of our tool is intentionally simple, visual, and intuitive, drawing on successful strategies from previous AI curricula developments [7], [8], [18], [19]. While tools like "AI for Oceans" and "Zumi" primarily focus on object detection, our tool introduces a novel element by emphasizing the significance of diverse training data, developer decisions, and their profound impact on end-users. This not only educates students on how AI systems operate but also instills a critical awareness of these technologies' ethical dimensions and societal implications, preparing them to be responsible creators and users of future AI systems.

III. Tool Design

In this section, we describe our tool's conceptual framework and practical application. Section III-A delves into the design philosophy behind the user interface and student engagement. Following that, Section III-B provides an overview of the tool's implementation. We have released the tool as an open-source repository, serving K-12 teachers and facilitating integration into various research projects [12]. This approach fosters collaboration, ensuring adaptability and continual enhancement for diverse educational contexts.

A. Overview

Design Choices & Adoption: We opted for the Machine Learning for Kids (ML4Kids) platform from various choices like Cognimates and Google Teachable Machine. Our decision was influenced by ML4Kids' simple model training process and seamless integration of the trained models into Python, Scratch, and App Inventor. Scratch's widespread use in K-12 computer science education aligns exceptionally well with our educational objectives by offering visual interfaces through block-based coding. Within Scratch, ML4Kids allows for integrating trained models as visual, block-based elements. This enables several important functionalities for students: they can upload and utilize their models, identify image labels, assess confidence levels of predictions, and add new training images directly from Scratch. This robust integration solidified our preference for ML4Kids.

A key criterion in designing our tool was ensuring ease of classroom adoption. We aimed to create an extensible tool, enabling K-12 educators to integrate it effortlessly into their curricula. The user-friendly interface of ML4Kids empowers

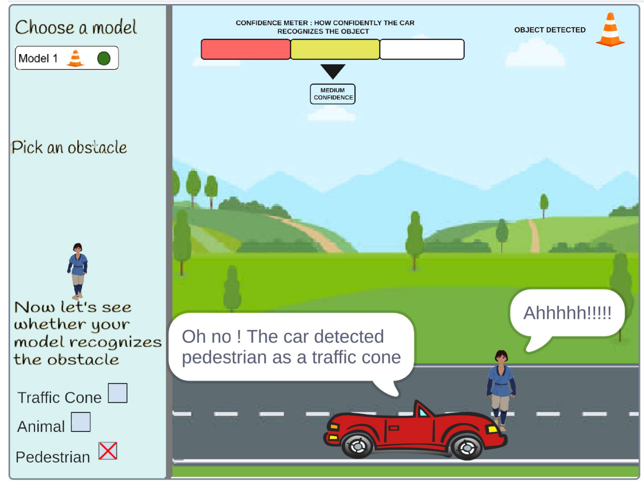

Fig. 2. Example illustrating our self-driving car tool using Model 1 trained on traffic cones and a pedestrian as the chosen obstacle. Here, misclassification leads to an accident, emphasizing the need for a well-trained model.

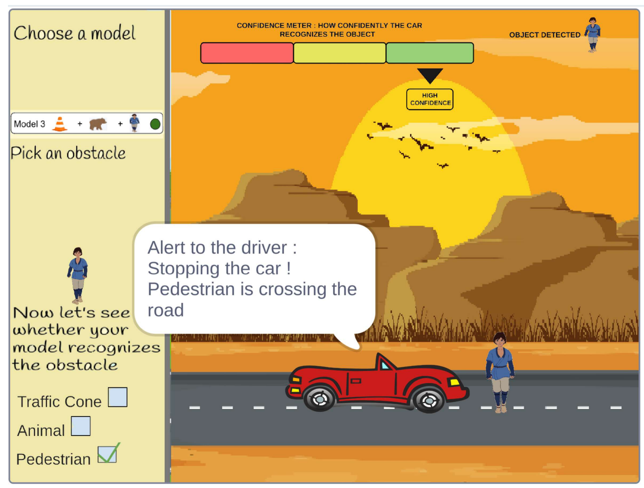

Fig. 3. Example illustrating our self-driving car tool using Model 3 trained on all obstacles and a pedestrian as the chosen obstacle. The car accurately identifies a pedestrian, stopping to let them cross the road.

educators and students to build classifiers, add images, and easily train/test new models. While this initial study didn't involve direct student interaction with the Scratch code, implementing our tool's user interface in Scratch ensures accessibility for non-expert programmers, teachers, and middle-school students. This allows them to update the tool code, explore different car behaviors, and understand AI, covering aspects like data collection, model creation, ethics, and societal impacts. Future iterations will focus on seamlessly integrating the tool into classrooms.

User Interface Design: The primary objective in designing the user interface was to provide a seamless and focused experience, enabling students to concentrate on learning about AI without the distraction of navigating multiple pages or platforms. Each simulation has clear navigation cues and narrative text to guide users intuitively and provide detailed descriptions of events. The interface allows students to interact with three distinct AI models and three types of obstacles using straightforward, easy-to-understand visual blocks, as illustrated on the left side of Figure 1. Model blocks are visually representative, displaying model numbers and obstacle images for easy comprehension. For example, the Model 1 block features a traffic cone, indicating its specialization in recognizing traffic cones. In contrast, the Model 3 block includes images of traffic cones, animals, and pedestrians, demonstrating its capacity to understand a variety of road obstacles. After model selection, students are prompted to choose dynamic obstacles like traffic cones, moving animals, and pedestrians.

On the right side of the interface, we designed visual demonstrations that simulate the car's reactions to obstacles based on the selected model and obstacle type. In the bottom-left corner, a feedback section provides immediate responses to the car's obstacle detection capabilities—displaying a red cross for incorrect detection and a green checkmark for successful detection (as shown in Figures 2 and 3). This feedback enables students to assess the model's performance quantitatively. To further enhance students' understanding of obstacle recognition, we integrated two components at the top of the interface: a confidence meter and an object-detected display. The confidence meter categorizes detection confidence into three levels: low (0-50), medium (50-80), and high (above 80), providing a visual representation of the model's confidence in recognizing the selected obstacle. Additionally, the "object detected" feature displays an image of the obstacle recognized by the car, helping students understand potential misinterpretations by the model. For instance, if the model incorrectly identifies a pedestrian as a traffic cone, the "object detected" feature would display a traffic cone, and the "confidence meter" would show how confident the model is about that prediction. This helps students understand and analyze the AI's decision-making process.

Self-Driving Car Simulation: In the simulation, our self-driving car is programmed to react differently to different obstacles. For traffic cones, the car deviates to avoid them; for animals on the roadside, it cautiously moves forward; and for pedestrians crossing the road, it stops promptly to ensure safety. Although typically, a car might react similarly to animals and pedestrians by slowing or stopping, our simulation distinguishes these scenarios to showcase the AI model's ability to differentiate between obstacles. We simulate an animal walking along the side of the road and a pedestrian crossing the road, prompting the car to slow down for the animal (anticipating it might enter the road) and to stop for the pedestrian until they have safely crossed. We avoided implementing more complex scenarios, such as varying the car's reaction when another vehicle is behind or running over cones to avoid obstacles. These could add significant complexity and potentially confuse middle-school students. Therefore, our tool focuses on simpler scenarios to help students intuitively understand machine learning concepts through a straightforward interface.

The self-driving simulation initiates when a student selects one of the three models and one of the three types of obstacles. This enables students to experiment with different scenarios to evaluate the model's robustness and ability to mitigate potential hazards on the road. The simulation outcomes include both successful and unsuccessful scenarios. For instance, if a student selects Model 1, the car uses Model 1's capabilities. Since Model 1 is trained solely on traffic cone images, it misinterprets all obstacles it encounters as traffic cones. When a pedestrian is chosen as an obstacle, as illustrated in Figure 2, the model erroneously identifies the pedestrian as a traffic cone. As a result, the car responds by changing lanes to avoid what it perceives as a traffic cone, which results in a simulated accident since the pedestrian is crossing the road. This trial-and-error process of model selection enables students to step into the role of developers, making informed decisions based on the training data of each model. This hands-on approach helps them grasp the technical aspects of model development. Additionally, by allowing students to select obstacles, the simulation engages them from an end-user's perspective, enabling them to directly experience and assess the model's real-world applications and consequences.

Conversely, the car performs optimally in an ideal scenario where a student selects Model 3, trained to recognize traffic cones, animals, and pedestrians. For instance, when a pedestrian obstacle is chosen, Model 3's training on pedestrian recognition ensures accurate detection. As depicted in Figure 3, the interface visually simulates that the car correctly identifies a pedestrian and stops accordingly to allow safe crossing. By illustrating both positive and negative outcomes through experimentation with different models and obstacles, we aim to give students a deeper understanding of AI's significant impact on human lives. This approach highlights the crucial role of developers in designing robust models that can accurately interpret real-world situations and mitigate potential risks.

B. Implementation

Implementing the ML Models: As mentioned earlier, we utilized the Machine Learning for Kids (ML4Kids) platform to develop three distinct image recognition models for our self-driving car tool, designed to recognize specific obstacles: traffic cones, animals, and pedestrians. As illustrated in Figure 4, Model 1 is trained solely on traffic cones, and Model 2 includes both traffic cones and animal images. In contrast, Model 3 is trained on images of all three obstacles: traffic cones, animals, and pedestrians. We intentionally varied the training data across the models while utilizing the same underlying image classification neural network [30], [31] provided by ML4Kids. This approach maintains consistency in the neural network architecture across the models. This enables students to concentrate on comprehending the influence of various training datasets rather than delving into the complexities of machine learning algorithms. Existing educational literature

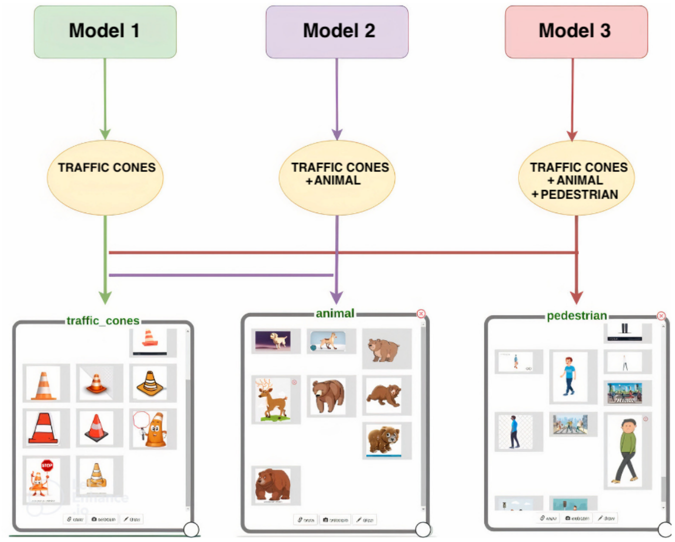

Fig. 4. Examples of images categorized for training the three image recognition models.

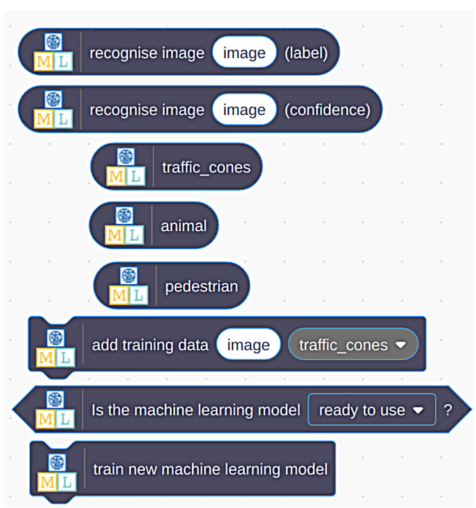

Fig. 5. Example Scratch blocks exported from ML4Kids for Model 3, trained with all obstacles - traffic cones, animals, and pedestrians.

[21] suggests that such complexities are better suited for high-school students and above.

To enhance the robustness of our models, we ensured the inclusion of a diverse and high-quality image dataset during the training phase, as illustrated in Figure 4. This methodology was designed to foster an unbiased and balanced training environment. For validation, we leveraged the testing feature of the ML4Kids platform, which evaluates the models by providing confidence scores and predicted labels for provided test images. We conducted multiple testing iterations for each model using a variety of traffic cones, animals, and pedestrian images. This allowed us to systematically assess the model's robustness and accuracy through the consistency of confidence scores and the correctness of the predicted labels across different test scenarios.

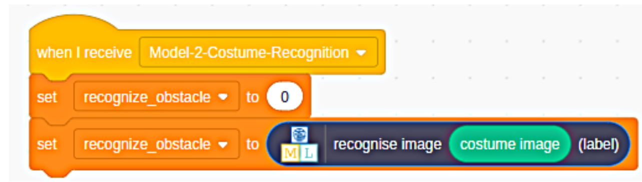

Fig. 6. Example Scratch code in our self-driving car tool to employ Model 2 for image detection.

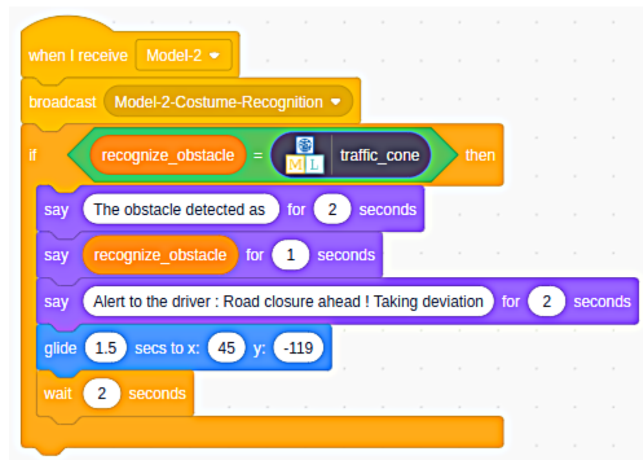

Fig. 7. Example Scratch code that moves the car to another lane when it predicts a traffic cone obstacle.

After ensuring the models functioned as anticipated and produced robust outputs, we exported them to the Scratch platform as interactive visual blocks using ML4Kids. An example of these blocks for Model 3, which can recognize traffic cones, animals, and pedestrians, is shown in Figure 5. Similarly, Model 1's visual blocks are labeled solely with a 'traffic cone.' Model 2's blocks are labeled with 'traffic cone' and 'animal.' Essentially, each model is represented by a distinct set of visual blocks on Scratch, corresponding to its specific image recognition capabilities.

Developing the Self-Driving Car Behavior: Using the exported Scratch blocks, we developed an interactive educational tool to enhance the learning experience. Integrating ML4Kids models into Scratch provided a consistent set of blocks for each model, as illustrated in Figure 5. In our tool, users select a model from the interface, and the corresponding blocks are automatically utilized for demonstrations. The tool represents images as Scratch sprites strategically positioned to activate specific block codes. These sprites interact with Scratch's in-built 'motion' and 'look' blocks, enabling dynamic movement and visual changes. This setup allows for interactive and seamless engagement with the visual elements, making the learning process intuitive and engaging.

To illustrate the interface construction in Scratch, we provide an example scenario that applies to all our simulations. Consider a student selecting Model 2 and selecting a traffic cone as the obstacle. The activation of this choice triggers specific code blocks shown in Figures 6 and 7. Initially, the

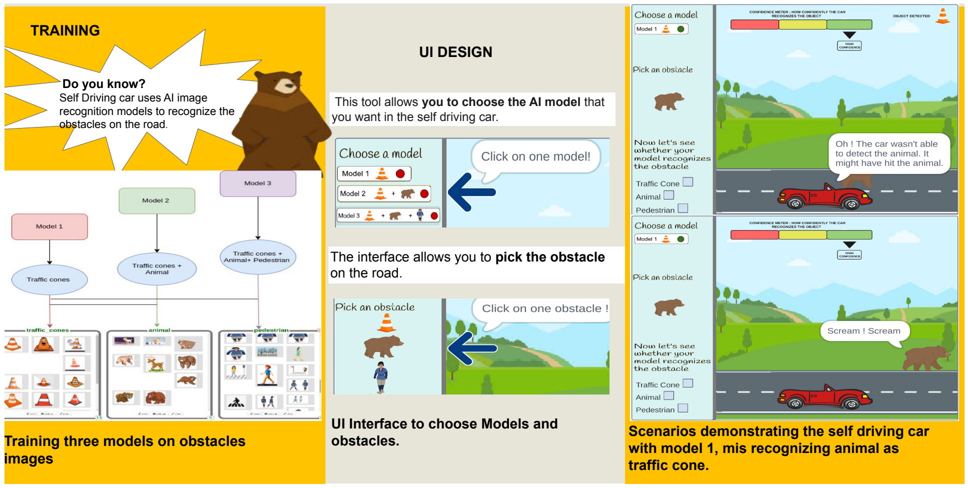

Fig. 8. The trifold poster used during the AI exhibit to orient students towards training data, AI models, and the self-driving car simulation tool.

code in Figure 6 uses Model 2 to recognize the traffic cone image. Upon successful identification, the code in Figure 7 starts the self-driving car simulation. If the detected obstacle is a traffic cone, the corresponding 'if' condition in the code evaluates as true, leading to a series of programmed responses: the simulation alerts the student to the detected traffic cone, initiates a lane change due to the supposed road closure, completes the lane switch, and then continues moving forward. Additionally, the code is designed to adapt the car's speed when an animal is detected nearby or to stop and proceed when a pedestrian is crossing the road.

IV. User Study Design

This tool, developed in the Spring 2023 semester as part of a special-topics course at our institution, is one of five AI exhibits [32]–[34]. We evaluated it in partnership with a K–8 public school in a gateway city in the Northeast U.S. This section details the participant recruitment, assent process, and user study design.

Participant Recruitment: We conducted a two-day educational study involving 26 middle school students from grades 6, 7, and 8 at K-8 school in Northeast region in United States. This study was approved by our Institutional Review Board (IRB). Students from grades 7 and 8 participated on the first day, while the second day included students from grades 6 and 7. Each session lasted approximately one hour. Specifically, the participant group consisted of seven students from grade 6, twelve from grade 7, and seven from grade 8. Initially, these students indicated unfamiliarity with specific terms such as 'model,' 'training,' and 'AI impacts.' While they were generally aware of the term Artificial Intelligence, they lacked a detailed understanding of how it worked. Consent was obtained from all participants and their parents, adhering to the ethical standards of our institution's ethics review board.

Tool Orientation: Recognizing that middle-school students may not be familiar with terms like 'training data,' 'model,' and 'predictions,' we initiated each session with a 5-minute introduction to these key concepts. We described 'training data' as a collection of information organized and labeled based on similarities and differences, including text and images. For instance, traffic cones are categorized and labeled together, while all animals can be classified and labeled in another category. We defined a 'model' as an intelligent computer program that uses mathematical techniques to learn from this training data. We demonstrated how the model learns to distinguish between different categories of images is similar to how humans learn new skills. We explained 'predictions' as the process of testing the model with familiar and new images to assess its accuracy and effectiveness. To make these concepts relatable, we drew an analogy to human learning, noting that both AI and humans start with no prior knowledge and improve through continuous exposure to new information. Moreover, to clarify the operation of self-driving cars, we likened AI and human vision: just as humans use their eyes to spot objects and their brains to recognize them, self-driving cars employ cameras to capture images, with the model as the brain to identify obstacles. During the session, we utilized a trifold poster (shown in Figure 8) to deepen understanding and engage the students. The left section of the poster explained different AI models and their training data, while the center and right sections detailed the tool's operations.

User Study: We began each session by first obtaining students' assent for participation. After the assent process and orientation explaining the key concepts and the tool's operations, we proceeded to outline the expectations for students during their interaction with the tool. Each class session was scheduled to last between 45 and 60 minutes, wherein students engaged with the AI tool through two separate terminals. Each student was allocated approximately 10 minutes to explore model selections and navigate through obstacles on our visual interface, allowing them to experiment with various combinations. Due to time constraints, allowing students to interact with the code was considered impractical. Further exploration of students' knowledge gained through code interaction is left for future work. Following each session, we administered a survey questionnaire where anonymous student responses were collected to assess their comprehension and insights gained during the experience. This systematic approach enabled us to gauge the effectiveness of our AI tool in conveying fundamental AI-related concepts.

V. Student Experiences

During our sessions, we noted significant enthusiasm and active comprehension from the middle school student participants. When we introduced our self-driving car tool, they readily made connections to terms like ‘automatic car’ and ‘Tesla.’ Our discussions primarily revolved around the concepts of models and training. One inquisitive student insightfully compared model training to practicing a skill, remarking, “the more they practice, the better they become.” This analogy resonated with the other students, demonstrating their deep engagement and understanding.

During their interactions with the tool, students were particularly excited by its visually appealing interface. They enjoyed navigating the car along the road, and watching pedestrians, and were especially entertained by elements like the ‘bear screaming and running.’ The visual interface significantly enhanced engagement, holding their interest throughout the session. We noted that the feedback section located at the bottom-left of the user interface did not capture much attention. Students were primarily focused on selecting models and obstacles and watching the simulation unfold. They seldom checked the feedback icons—green check or red cross—to see if their observations matched the expected outcomes. This might be because they relied on their own assessments of how the car behaved rather than on the provided feedback signs. In the future, as we introduce more complex scenarios, we expect the feedback UI to become more relevant and to help students verify their intuitions against expected behaviors. We plan to integrate the feedback more directly into the simulation window, making it easier for students to both experience the simulation and view the results simultaneously

While interacting with Model 1, students observed that the car performed poorly, particularly when encountering animals and pedestrians, which initially led to disappointment. They found it challenging to understand why Model 1 had these limitations. When they switched to Model 2, they encountered similar issues with pedestrian interaction. This experience helped them make a connection between the observed deficiencies and the model’s training, as one student pointed out, “Oh, because the model isn’t using pedestrians”. As the students moved on to Model 3, they observed substantial improvements in how the car navigated obstacles, clearly linking these advancements to the model’s comprehensive training. This exploration across all three models deepened the students’ understanding of each model’s unique capabilities. This hands-on experience encouraged some students to revisit Model 1 with varied obstacles, allowing them to analyze the outcomes more effectively.

TABLE I

Summary of the user study results showcasing participants’ understanding of key AI concepts and awareness of societal impacts associated with self-driving cars.

Student Insights	**Correct Responses**		
	Grade 6 (N=8)	**Grade 7 (N=12)**	**Grade 8 (N=7)**
Grasping Road Safety Ethics in Self-Driving Cars	7	11	6
Recognizing Impacts of Limited Model Training	6	11	6
Understanding Role of Data in Robust AI Model Development	8	11	7

During the study, the students fostered a collaborative atmosphere by actively engaging in discussions, sharing their insights, and drawing conclusions about the model behaviors. Notable moments arose during discussions on the benefits of AI in self-driving cars. The students made insightful comments, highlighting that “people with disabilities can use (the car)” and pointing out the safety feature where “(the car) can stop if there is an obstacle or if something unexpected happens.” These discussions highlighted the students’ ability to understand both the advantages and limitations of AI, appreciating its varied applications. It was truly gratifying to witness students embracing a more nuanced perspective of AI through this interactive experience.

VI. Evaluation Results

In this section, our goal is to assess students’ understanding of several key aspects, including the societal implications of AI, the role of data in training AI models, and the student’s ability to evaluate model performance. Our survey included four questions specifically designed to probe these areas. Each question was verified for readability using a Flesch-Kincaid readability calculator [35] to ensure they were age-appropriate for the student’s grade. The summarized survey results are presented in Table I.

Grasping Road Safety Ethics in Self-Driving Cars: We presented students with a multiple-choice question to evaluate their understanding of ethical behavior in self-driving cars when encountering pedestrians. The question was: *“What should a self-driving car do when it sees a pedestrian on the road?”* with options: *“Stop until the pedestrian crosses the road,” “Slow down for the pedestrian,”* and *“Keep going and don’t stop for the pedestrian”*. This assessment is crucial for understanding students’ awareness of responsible and ethical decision-making in the design of self-driving cars, often referred to in the literature as “Theory of Mind” [36]. Based on their knowledge that vehicles should stop for pedestrians, the students applied the same expectation to self-driving cars. Of 26 students, 23 (88.4%) chose that the self-driving car should stop upon detecting a pedestrian, while the remaining three students (one from each grade) thought the car should merely slow down. These responses indicate a firm grasp of road safety ethics among the students, underscoring the importance of instilling these values in AI technology.

Recognizing Limitations and Impacts of Limited Model Training: We posed another question, *"Assume we taught a car only to recognize traffic cones. What do you think the car would actually see if this car saw a pedestrian walking by?"* with choices: *"Traffic cone," "Pedestrian,"* and *"Animal."* This question gauges students' understanding of model limitations when trained with limited data and subsequent implications. It assesses our objective of visually demonstrating societal impacts on end-users. The results indicate that 22 out of 26 students (84.6%) correctly identified that a model trained solely on traffic cones would recognize a pedestrian as a traffic cone. The results were consistent across grades 6 to 8, with six grade 6 students (out of eight), eleven grade 7 students (out of twelve), and six grade 8 students (out of seven) accurately answering the question. These results suggest middle school students comprehend that model limitations lead to obstacle misclassification, impacting end-users' safety.

Understanding Role of Data in Robust AI Model Development: We included a multiple-choice question to evaluate students' understanding of how training data affects a self-driving car's decision-making capabilities. The question asked, *"Select which option would help a self-driving car (i.e., model) make better decisions when it sees objects on the road?"* with choices like *"Adding more data to train the models," "Less training data should be fine,"* and *"The model works the same no matter what data it uses."* This question aimed to assess the students' awareness of the importance of training data and the developers' role in enhancing the model performance. It also assesses the efficacy of our objective of incorporating three distinct models within the interface. Results indicate that most participants, 25 out of 26 (96.2%), answered correctly. The results underscore that students understood the benefits of adding more training data, which leads to more responsible and robust AI behavior.

Students' Perspectives on AI's Positive and Negative Impacts in Self-Driving Cars: As part of an optional survey, we posed the open-ended question, *"Describe both good and bad things about using AI/self-driving cars."* Remarkably, almost 23 out of 26 students chose to respond, highlighting their engagement. This question aimed to assess the students' understanding of both the positive and negative outcomes of AI in self-driving cars, gauging their ability to grasp a nuanced perspective of AI's impact. Additionally, it sought to ascertain whether the tool facilitated their ability to connect these concepts to real-world experiences. Nearly all students reflected on the positive impact of self-driving cars while also recognizing the importance of effective model training to achieve these positive outcomes. Responses such as *"one positive impact of using AI in these scenarios is that it could lead to fewer accidents if trained properly, and one negative impact is that it could lead to accidents if not trained properly"* and *"positive is that neither a pedestrian nor animal get hurt as well as the driver. The negative is that if it is only programmed for a traffic cone, it could hit a pedestrian"* revealed a nuanced understanding of the advantages and drawbacks of self-driving cars, a crucial aspect in today's AI landscape.

VII. Reflections on First Grade Students

We acknowledge that the understanding of our self-driving car tool might be limited for first-grade students. Therefore, we have excluded their responses from the primary section of our study. However, we are separately analyzing their feedback to gain valuable insights. We hope that our insights will then inform and guide future research endeavors, specifically aimed at developing effective AI education tools for young learners.

Our study involved 12 first-grade students who participated in the AI exhibit. With this group, we employed a distinct approach compared to the middle-school group. To introduce the topic, we engaged the children in a playful physical exercise. They assumed the role of a car and navigated around a chair used as an obstacle. Afterward, we inquired about their strategies for maneuvering around the chair. The majority of them mentioned they relied on their ability to see the chair. This simple analogy mirrored the car's use of cameras and sensors to identify obstacles on the road. Subsequently, we showcased our tool, presenting scenarios where the car successfully detected obstacles and instances where it struggled. Our focus was exclusively on explaining the concept of self-driving cars and how they perceive obstacles. By incorporating hands-on activities and relatable comparisons, we aimed to effectively engage and educate the first-grade students.

To evaluate the first graders' grasp of self-driving cars, we devised a distinct set of questions in comparison to those used with the middle school group. We inquired about their preference for the best-performing car. Among the 12 students, 11 students deemed the Model 3 car as good. Their assessments mainly revolved around the car's capability to avoid collisions. For the subsequent question, we asked the students to identify the worst-performing car and explain their choice. In response, eight students provided input. Among them, five highlighted the Model 2 car as the worst, while three students singled out the Model 1 car. One student's response encapsulates this sentiment: *"good: because the third one could do all of it and bad: second car because it failed the animal."*

VIII. Conclusion and Future Work

The ascent of AI brings promise for positive industry transformations but also raises concerns about fairness, inclusivity, and safety, with instances of bias in self-driving technology. This paper introduces an interactive tool simulating self-driving cars, aiming to provide children with insights into AI from both developer and user perspectives. The tool enhances understanding of societal AI impacts and stresses the importance of equitable model development. Initial assessments with middle-school students suggest that learners can effectively grasp AI concepts and responsible usage through the tool, indicating potential for shaping well-informed AI interactions in younger generations. In the future, we aim to broaden the tool's applicability by involving a larger and more diverse group of student participants across various educational settings, with the goal of further enhancing its effectiveness and inclusivity in schools.

Acknowledgement

We would like to thank the teachers who assisted in setting up our exhibit at their K-8 school in a gateway city in the Northeast U.S. We also sincerely thank the children who participated in our study. This work was supported by the National Science Foundation under Grant IIS-2112633.

References

[1] M. S. M. Raub, "Bots, bias and big data: Artificial intelligence, algorithmic bias and disparate impact liability in hiring practices," 2018. [Online]. Available: https://api.semanticscholar.org/CorpusID:169863807

[2] Z. Obermeyer, B. Powers, C. Vogeli, and S. Mullainathan, "Dissecting racial bias in an algorithm used to manage the health of populations," *Science*, vol. 366, no. 6464, pp. 447–453, 2019.

[3] J. Buolamwini and I. D. Raji, "Actionable auditing: investigating the impact of publicly naming biased performance results of commercial ai products." Conference on Artificial Intelligence, Ethics, and Society, 2019.

[4] E. Ferrara, "Should chatgpt be biased? challenges and risks of bias in large language models," *arXiv preprint arXiv:2304.03738*, 2023.

[5] National Highway Traffic Safety Administration, "Summary Report: Standing General Order on Crash Reporting for Level 2 Advanced Driver Assistance Systems," https://www.nhtsa.gov/sites/nhtsa.gov/files/2022-06/ADAS-L2-SGO-Report-June-2022.pdf, June 2022.

[6] D. Touretzky, C. Gardner-McCune, F. Martin, and D. Seehorn, "Envisioning ai for k-12: What should every child know about ai?" *Proceedings of the AAAI Conference on Artificial Intelligence*, 7 2019. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/5053

[7] I. Lee, S. Ali, H. Zhang, D. DiPaola, and C. Breazeal, "Developing middle school students' ai literacy," in *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*, ser. SIGCSE '21, 2021. [Online]. Available: https://doi.org/10.1145/3408877.3432513

[8] D. Long and B. Magerko, "What is ai literacy? competencies and design considerations," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI '20, 2020. [Online]. Available: https://doi.org/10.1145/3313831.3376727

[9] N. Garrett, N. Beard, and C. Fiesler, "More than "if time allows": The role of ethics in ai education," in *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, ser. AIES '20, 2020. [Online]. Available: https://doi.org/10.1145/3375627.3375868

[10] P. Lin, J. Van Brummelen, G. Lukin, R. Williams, and C. Breazeal, "Zhorai: Designing a conversational agent for children to explore machine learning concepts," *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/7061

[11] D. DiPaola, B. H. Payne, and C. Breazeal, "Decoding design agendas: An ethical design activity for middle school students," in *Proceedings of the Interaction Design and Children Conference*, ser. IDC '20, 2020. [Online]. Available: https://doi.org/10.1145/3392063.3394396

[12] P. Rayavaram, "Self driving car tool," 2024, gitHub repository. [Online]. Available: https://github.com/PranathiRayavaram/AI-Education-SelfDrivingTool

[13] M. Research, "Self-driving cars market by component (radar, lidar, ultrasonic, & camera unit), vehicle (hatchback, coupe & sports car, sedan, suv), level of autonomy (l1, l2, l3, l4, l5), mobility type, ev and region - global forecast to 2030," MarketResearch.com, 2023. [Online]. Available: https://www.marketresearch.com/MarketsandMarkets-v3719/Self-driving-Cars-Component-Radar-30653755/

[14] W. Holmes and I. Tuomi, "State of the art and practice in ai in education," *European Journal of Education*, vol. 57, 2022. [Online]. Available: https://doi.org/10.1111/ejed.12533

[15] Massachusetts Institute of Technology, "MIT Moral Machine," 2023. [Online]. Available: https://www.moralmachine.net/

[16] S. Ali, V. Kumar, and C. Breazeal, "Ai audit: A card game to reflect on everyday ai systems." AAAI Press, 2023.

[17] H. Furey and F. G. Martin, "Introducing ethical thinking about autonomous vehicles into an ai course," in *AAAI Conference on Artificial Intelligence*, 2018. [Online]. Available: https://api.semanticscholar.org/CorpusID:19190942

[18] MIT Media Lab, "AI + Ethics Curriculum for Middle School," MIT Media Lab, 2023. [Online]. Available: https://www.media.mit.edu/projects/ai-ethics-for-middle-school/overview/

[19] A. Sabuncuoglu, "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*, ser. ITiCSE '20, 2020. [Online]. Available: https://doi.org/10.1145/3341525.3387364

[20] Code.org, "Code.org AI for Oceans Activity," Code.org, 2023. [Online]. Available: https://code.org/oceans

[21] Elements of AI, "Elements of AI," Elements of AI, 2023. [Online]. Available: https://www.elementsofai.com/

[22] Brown University, "An Interactive Introduction to Artificial Intelligence (AI)," Brown University, 2023. [Online]. Available: https://okai.brown.edu/index.html

[23] H. N. Pang, R. Parks, C. Breazeal, and H. Abelson, ""how can i code a.i. responsibly?": The effect of computational action on k-12 students learning and creating socially responsible a.i." *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, 2023. [Online]. Available: https://doi.org/10.1609/aaai.v37i13.26902

[24] Roireneolink, "Zumi AI Robot," Robolink, 2023. [Online]. Available: https://www.robolink.com/products/zumi

[25] Cognimates, "An AI education platform for building games, programming robots & training AI models," Cognimates, 2023. [Online]. Available: http://cognimates.me/home/

[26] Machine Learning for Kids, "Teach a computer to play a game," Machine Learning for Kids, 2023. [Online]. Available: https://machinelearningforkids.co.uk/

[27] Google, "Teachable Machine by Google," Google, 2023. [Online]. Available: https://teachablemachine.withgoogle.com/

[28] Apple, "Create ML by Apple," Apple, 2023. [Online]. Available: https://developer.apple.com/machine-learning/create-ml/

[29] Scratch Foundation, "Scratch," MIT, 2023. [Online]. Available: https://scratch.mit.edu/

[30] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4510–4520.

[31] IBM, "Taxinomitis - a machine learning tool for kids," 2024, accessed: 2024-04-12. [Online]. Available: https://github.com/IBM/taxinomitis/tree/master

[32] S. Vahedian Movahed, J. Dimino, A. Farrell, E. Irankhah, S. Ghosh, G. Jain, V. Mahipal, P. Rayavaram, I. T. Sanusi, E. Salas, K. Wolkowicz, S. Narain, and F. Martin, "Introducing children to ai and ml with five software exhibits," in *Extended Abstracts of the 2024 CHI Conference on Human Factors in Computing Systems*, ser. CHI EA '24. New York, NY, USA: Association for Computing Machinery, 2024. [Online]. Available: https://doi.org/10.1145/3613905.3650991

[33] F. Martin, S. Vahedian Movahed, J. Dimino, A. Farrell, E. Irankhah, S. Ghosh, G. Jain, V. Mahipal, P. Rayavaram, I. Temitayo Sanusi, E. Salas, K. Wolkowicz, and S. Narain, "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 V. 2*, ser. SIGCSE 2024. New York, NY, USA: Association for Computing Machinery, 2024, p. 1740–1741. [Online]. Available: https://doi.org/10.1145/3626253.3635512

[34] F. Martin, V. Mahipal, G. Jain, S. Ghosh, and I. T. Sanusi, "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*, ser. SIGCSE 2024. New York, NY, USA: Association for Computing Machinery, 2024, p. 792–798. [Online]. Available: https://doi.org/10.1145/3626252.3630804

[35] J. P. Kincaid, R. P. Fishburne Jr, R. L. Rogers, and B. S. Chissom, "Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel," 1975.

[36] S. Baron-Cohen, A. M. Leslie, and U. Frith, "The child's theory of mind," *Mind & Language*, vol. 1, no. 3, 1985.