

Lesson 15: Data and Machine Learning

Overview

Question of the Day: How can machines "learn"?

In this lesson, students are introduced to the concepts of Artificial Intelligence and Machine Learning using the AI for Oceans widget. First students classify objects as either "fish" or "not fish" to attempt to remove trash from the ocean. Then, students will need to expand their training data set to include other sea creatures that belong in the water. In the second part of the activity, students will choose their own labels to apply to images of randomly generated fish. This training data is used for a machine learning model that should then be able to label new images on its own.

Purpose

In previous lessons, students have seen how we can use data to make decisions. We've also seen that data can be collected about us constantly, leading to a larger amount of data to analyze - more than a human can handle! This tutorial is designed to quickly introduce students to machine learning, a type of artificial intelligence that can be used to make decisions about large amounts of data. Students will explore how training data is used to enable a machine learning model to classify new data.

Agenda

Warm Up (5 minutes)

Journal

Activity (35 minutes)

Wrap Up (5 minutes)

Journal

Objectives

Students will be able to:

- Reason about how human bias plays a role in machine learning.
- Train and test a machine learning model.

Preparation

- Review and complete the online tutorial yourself. If you are not going to use AI for Oceans, explore the other options listed below.

Links

Heads Up! Please make a copy of any documents you plan to share with students.

For the teachers

- **CSD Unit 5 - Data & Society** - Slides

For the students

- **Video: Impact on Society** - Video ([Download](#))
- **Video: Machine Learning** - Video ([Download](#))
- **Video: Training Data and Bias** - Video ([Download](#))

Teaching Guide

Warm Up (5 minutes)

Journal

Prompt: Has a computer ever made a recommendation for you? How do you think it learned how to do

this?

Discussion Goal

Goal: Answers may vary and may depend on prior experiences students have with recommendation systems or other types of artificial intelligence. Try to steer the discussion towards conversations around the role that humans play in machine learning. It's ok if the discussion here is short - you are setting the stage for the upcoming activity.

Discuss: Have students brainstorm silently on their own, then have them share with neighbors, and finally have them share out with the room.

Remarks

Today we're going to be learning more about Machine Learning and its impacts.

Question of the Day: How can machines "learn"?

Activity (35 minutes)

Video: Play the video "What is Machine Learning".

1

Video: What is Machine Learning?

Teaching Tip

Alternatives to AI For Oceans: AI for Oceans was originally developed as an Hour of Code activity that can be completed by students with any device available. We have modified it for its usage here. Depending on your classroom situation, you might opt to replace the activity with:

- **Teachable Machines** - *Teachable Machine is a web-based tool that makes creating machine learning models fast, easy, and accessible to everyone. Teachable Machine is flexible - use files or capture examples live. It's respectful of the way you work. You can even choose to use it entirely on-device, without any webcam or microphone data leaving your computer.*
 - If your classroom's devices have cameras, Teachable Machines offers an engaging way to create training sets. Encourage students to teach the machine to represent rock, paper, or scissors with hand gestures. What are some possible ways for bias to enter in?
- **Machine Learning for Kids** - *This free tool introduces machine learning by providing hands-on experiences for training machine learning systems and building things with them. It provides an easy-to-use guided environment for training machine learning models to recognise text, numbers, images, or sounds.*
 - Machine Learning for Kids is a great option if your students want to work with text samples. Teach the machine to recognize words or passages that are happy or sad. Lots to play around with here!

Remarks

Machine learning refers to a computer that can recognize patterns and make decisions without being explicitly programmed. In this activity you're going to supply the data to train your own machine learning model. Imagine an ocean that contains creatures like fish, but also contains trash dumped by humans. What if we could train a computer to tell the difference and then use that technology to help

clean the ocean?

Do This: Direct students to Levels 2-5 on Code Studio. Students should spend around five minutes total on these levels. Prompt their thinking with the "Consider" on the slide. To program A.I., use the buttons to label an image as either "fish" or "not fish". Each image and label becomes part of the data used to train A.I. to do it on its own. Once trained, A.I. will attempt to label 100 new images on its own, then present a selection that it determined have the highest probability of being "fish" based on its training. Students who consistently label things correctly should see an ocean full of different types of sea creatures, without much (or any) other objects.



2-4

AI for Oceans: Fish vs Not Fish

2

3

4



Content Corner

Every image in this part of the tutorial is fed into a neural network that has been pre-trained on a huge set of data called **ImageNet**. The database contains over 14 million hand-annotated images. ImageNet contains more than 20,000 categories with a typical category, such as "balloon" or "strawberry", consisting of several hundred images. When A.I. is scanning new images and making its own predictions in the tutorial, it is actually comparing the possible categories for the new image with the patterns it found in the training dataset.

Discuss: How well did A.I. do? How do you think it decided what to include in the ocean?

Video: Play the video "Training Data & Bias".

5

Video: Training Data and Bias

Prompt: How do you think your training data influenced the results that A.I. produced?



Discussion Goal

Goal: Get students to reflect on their experience so far. It is important at this point that they realize the labeling they are doing is actually programming the computer. The examples they show A.I. are the "training data".

Remarks

In the second half of the activity, you will teach A.I. about a word of your choosing by showing it examples of that type of fish. As before, A.I. doesn't start with any training data about these labels. Even though the words in this level are fairly objective, it's possible that you will end up with different results based on their training data. You might even intentionally train A.I. incorrectly to see what happens!



Content Corner

The fish in this tutorial are randomly generated based on some pre-defined components, including mouths, tails, eyes, scales, and fins, with a randomly chosen body color, shape, and size. Rather than looking at the actual image data, A.I. is now looking for patterns in these components based on how the student classifies each fish. It will be more likely to label a fish the same way the student would have if it has matching traits.

Do This: Direct students to Levels 6-7 on Code Studio. Students should spend around five minutes total on these levels. Prompt their thinking with the "Consider" on the slide. Here, as before, students will use training data to teach A.I. to recognize different types of fish. The words in this list are intentionally more subjective than what students will have seen so far. Encourage students to decide for themselves what makes a fish look "angry" or "fun". Two students may choose the same label and get a very different set of results based on which fish traits were their focus. Encourage students to discuss their findings with each other or go back and choose new words. Each student will rely on their own opinions to train A.I. which means that A.I. will learn with the same biases held by the students. As students begin to see the role their opinion is playing, ask them to reflect on whether this is good or bad, and how it might be addressed.



6-7

AI for Oceans: Types of Fish

6

7

Prompt: How could biased data result in problems for artificial intelligence? What are ways to address this?



Discussion Goal

Goal: At this point, students should have some preliminary thoughts on how biased data leads to problems for artificial intelligence. They may bring up that if the data sets are trained incorrectly, there will be incorrect or misinterpreted conclusions. It can be addressed through diverse training sets. The following video dives into this subject further.

Video: Play the video "How I'm fighting bias in algorithms" with Joy Buolamwini.

Prompts:

- What are other ways human bias could appear in Machine Learning?
- How can we try to avoid that bias?

Allow students to discuss these prompts with a neighbor, then share out ideas to the full class. This may be the first time students are considering issues of bias in technology, so it's okay to not arrive at any solid conclusions and leave with more questions.

Display Display the Problem Solving Process graphic with Empathy in the center.

Remarks

Machine learning have led to innovations in medicine, business, and science but information discovered in this way has been used to harm or exclude groups of individuals.

As we've seen, problems of bias are often created by the type or source of data being collected. Collecting more data does not mean that the bias is removed.

Programmers (that includes you!) should take action to reduce bias in the apps and websites we use. An important strategy for this is at the center of our Problem Solving Process: empathizing with others and making sure no groups are excluded from our work. Be on the lookout for bias, and be empathetic and inclusive to try and avoid it!

Review: Play the video "Impact on Society" which recaps the concepts discussed today.

Wrap Up (5 minutes)

Journal

Prompt: What is your biggest takeaway from today's lesson, either about machine learning or how bias appears in technology?

💡 Teaching Tip

You can share these stories with your class to help them see how AI will impact the future.

- **Food Waste Is a Serious Problem. AI Is Trying to Solve It**
- **AI tech can identify genetic disorders from a person's face**
- **How an AI Startup Designed a Drug Candidate in Just 46 Days**
- **MIT AI tool can predict breast cancer up to 5 years early**
- **The Army steps up its pace on self-driving cars**
- **San Francisco says it will use AI to reduce bias when charging people with crimes**
- **AI is helping scholars restore ancient Greek texts on stone tablets**

Discuss: Time may be running short at this point in the class. Encourage students to share with a neighbor or share out with the room.
