

Lesson 1: AI for Oceans

45 minutes

Overview

In this lesson, learners of all ages get an introductory experience with coding and computer science in a safe, supportive environment. 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

This tutorial is designed to quickly introduce students to **machine learning**, a type of **artificial intelligence**. Students will explore how **training data** is used to enable a machine learning model to classify new data. Students should have a positive experience during the tutorial and more importantly should be motivated to keep learning computer science.

Agenda

Warm Up (5 minutes)

Build Excitement!

Activity (40 minutes)

AI For Oceans

Celebrate and Keep Going!

Extended Learning

Digging into the Information View

Example 1: Training based on one trait (body shape)

Example 2: Using the Information view to improve the model

Objectives

Students will be able to:

- Discuss the role artificial intelligence plays in their lives.
- Reason about how human bias plays a role in machine learning.
- Train and test a machine learning model.

Preparation

One Week Before Your Hour

of Code

- Review the [Hour of Code Educator Guide](#) in order to begin to plan your Hour of Code event.
- [Register your Hour of Code event.](#)
- Review and complete the online tutorial yourself.
- Be sure to test it first before asking your students to complete it. Check your technology and decide if you need to troubleshoot anything in advance of your Hour of Code.

One Day Before Your Hour of

Code

- Each student who completes the activity should receive a certificate. Print one [certificate](#) for each student in advance to make this easier at the end of your Hour of Code.

Links

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

For the students

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

Vocabulary

- **Machine Learning** - How computers recognize patterns and make decisions without being explicitly programmed

Teaching Guide

Warm Up (5 minutes)

Build Excitement!

Motivate: Explain to students the goals of today's activity. They are going to start using a new tool that will let them train a real machine learning model, a form of artificial intelligence.

General Support: As a teacher your role is primarily to support students as they make their way through the tutorial. Here are a few tips that should help students regardless of the level they're working on.

- **Checking Correctness:** This tutorial will not tell students whether they completed the level correctly. It is possible to skip through the different parts of the activity quickly. Encourage students to watch the videos, read the instructions, and try different things along the way. At any time, they can share their findings with you or a classmate.
- **Collaborate with Neighbors:** Encourage students to check in with a neighbor to discuss what they are experiencing. Since this tutorial includes videos and students may be wearing headphones it can get easy to "go into a bubble". Help break those barriers by actively pairing students.
- **Read the Instructions:** The instructions usually provide helpful information about what is happening behind the scenes.
- **Go back and try different things:** If students finish quickly, encourage them to go back to "Train More". In the last part of the activity, students can also go back and choose a "New Word". More training data tends to make the machine learning model more accurate and consistent. Students can also learn by purposefully training their model incorrectly, or not training it at all. **Video: [AI: Machine Learning](#)** The first level of this activity is a video that gives important context around artificial intelligence and machine learning. Watch it as a class and debrief afterwards to help students build connections to the content.

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**

Quick Share-out: Where have you seen or experienced artificial intelligence in your lives? Examples from the video include:

- email filters
- auto-complete text
- video recommendation systems
- voice recognition
- translation apps
- digital assistants
- image recognition

Prompt: Based on what you saw in the video, what is machine learning?

Discuss: Beginning in small groups then moving to whole class, students share their responses. The goal here is to get students acquainted with the world of artificial intelligence, reflect on the prevalence of artificial intelligence in our lives and think broadly about potential future innovations.

Say: *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?*

Activity (40 minutes)

AI For Oceans

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.

Levels 2-4 - Train AI to Clean the Ocean

Students can work through levels 2-4 on their own or with a partner. 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

Train an AI to Clean the Ocean

2

3

4

Level 5 - Training Data & Bias

Quick Share-out: How well did A.I. do? How do you think it decided what to include in the ocean?

5

AI: Training Data & Bias - Video

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

Discuss: In small groups, students share their responses. Circulate the room and listen to student ideas. This can be followed with full class discussion, or students can jump right back into the tutorial. The goal is to 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".

Levels 6 - Using Training Data

In the second half of the activity, students will teach A.I. about a word of their 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 students will end up with different results based on their training data. Some students may even intentionally train A.I. incorrectly to see what happens. If students are reflecting on how machine learning works, it should be encouraged!

6

Using Training Data



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.

Level 7 - Impacts on Society

Say: *Artificial intelligence systems learn from the data we give it, but that data could be opinion-based or biased.*

7

AI: Impact on Society - Video

Say: Think back to the examples of artificial intelligence we discussed at the beginning. Think of a time where machine learning might have gotten something wrong in the real world? (For example, voice recognition fails to understand you.)

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

Discuss: Beginning in small groups then moving to whole class, students share their responses. The goal of this discussion is to bring students back to the context of artificial intelligence in the real world.

Say: Some ways to address bias are to use a larger or more inclusive set of training data. In the final part of the activity you're going to teach A.I. a word that could be interpreted in different ways.

Level 8 - Teach A.I. a New Word

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 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.

Open question: What is a problem you care about that artificial intelligence could help to solve?



Teach AI a New Word

Celebrate and Keep Going!

Print certificates for students to celebrate their achievement.

At the end of the lesson encourage students to head to [Code.org/learn](https://code.org/learn) where they can get ideas for more ways to keep their adventure with computer science going.

Teaching Tip



Professional Learning Opportunity

Are you interested in teaching your students more about AI? Consider engaging in our self-paced professional learning module, **[Preparing to Teach AI and Machine Learning](#)**! This module will give you all the tools, skills, and knowledge you need to teach your students our AI and Machine Learning unit.

Extended Learning

Help Classify Animals at Mountain Zebra National Park

Snapshot Safari has placed hundreds of hidden cameras across southern Africa, capturing millions of images of beautiful and rare animals. Students can help protect the endangered Cape Mountain Zebra by classifying the different animals in these images. You can read about the project [here](#) or click below to give it a try!

[Snapshot Mountain Zebra - Zooniverse](#)

Digging into the Information View

On the final results screen, a visualization of A.I.'s machine learning model can be viewed by clicking



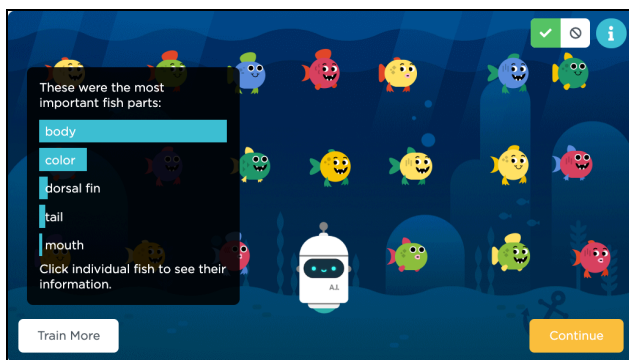
Example 1: Training based on one trait (body shape)

As an example, let's look at a student who is training A.I. to recognize "circular" fish by only selecting fish with by looking at a single trait:

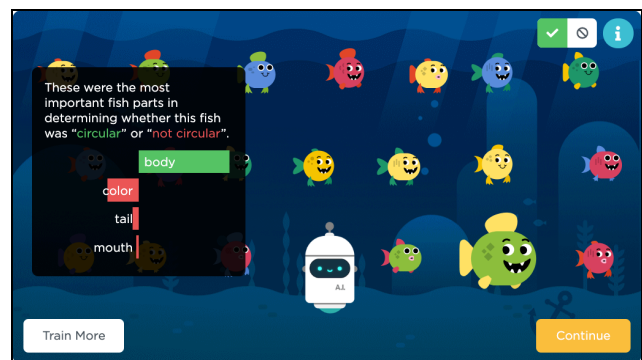
"Circular" body shapes:

After about 30 sets of training data, we can see that A.I. is able to classify the fish perfectly. It has successfully recognized a pattern: All of the fish the student selected had one of the same two body shapes.

"Circular" fish - training: 30



By clicking on the information button, we can see that A.I. is very confident that "body" is the most important component overall in the fish it has determined are "circular".



An individual fish can be inspected by clicking it. The information here shows that for this fish, A.I. determined that its body shape was a positive indicator that this was a "circular" fish.

"Not circular" fish - training: 30



By toggling the view, we can also analyze the fish A.I. determined are "not circular". Green bars pointing to the right represent a signal that the trait is "circular". Red bars pointing left signal that it's "not circular". A.I. decides how to classify each fish by weighing *all* of the components, even those that aren't shown in the list.*

What about the other traits?

You may have noticed that the other traits such as color seem to matter to A.I. even though the student was just focusing on body shape. Besides the obvious fact that A.I. can't literally read our minds, there is another reason A.I. might draw some of the wrong conclusions.

The fish in this tutorial are generated randomly. The fish the student labeled as "circular" happened to randomly have some of the same colors. If the set of training data didn't include fish that are green with round bodies, A.I. will be less confident about classifying one as "circular" even if that's what the student intended.

Example 2: Using the Information view to improve the model

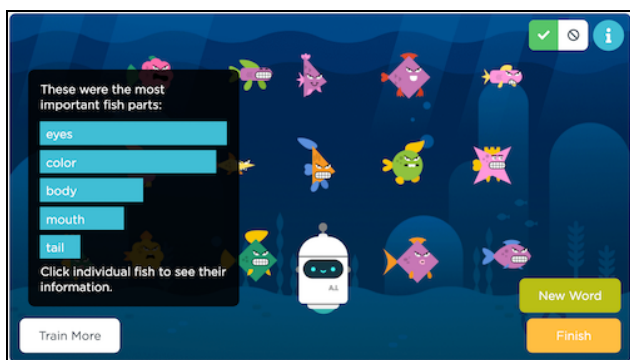
Adding more training data not only changes the results, but the data shown in the Info view. Sometimes the results graph shown can make it clearer why the results don't match expectations.

Let's consider a student who wants to train A.I. to recognize **"angry"** fish. They want to accomplish this by selecting any fish with eyebrows that point inwards and frowning mouths. They will label a fish as "angry" if it has a combination of these traits:

"Angry" eyes:  "Angry" mouths: 

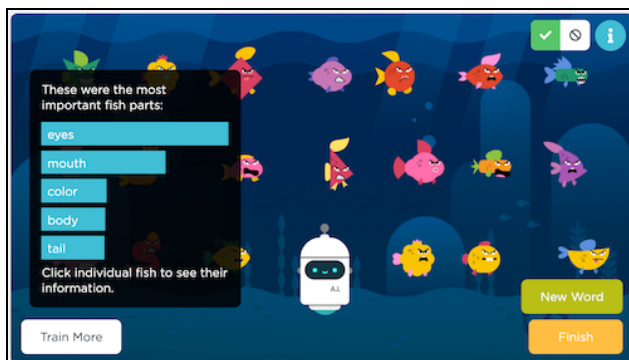
Can A.I. figure out what the student is doing without being explicitly programmed? It depends on how much training data the student provides and the quality of data in the set.

"Angry" fish - training: 30



Here we can see that A.I. picked up right away on the "eyes" being the most important trait, but closely followed by "color", not "mouth". There are a few purple fish in the left image that have mouths that are different from what the student was targeting. Perhaps many of the "angry" fish the student initially selected were purple? One way to improve the model would be to continue training, making sure to keep going until several purple "not angry" fish have been correctly labeled.

"Angry" fish - training: 200



Improved results after additional training.



This work is available under a [Creative Commons License \(CC BY-NC-SA 4.0\)](https://creativecommons.org/licenses/by-nc-sa/4.0/).

If you are interested in licensing Code.org materials for commercial purposes [contact us](#).