

Zhorai - Elementary School - Grades 3-5

Essential Question

How can an AI learn through interaction?

Objectives

Students will be able to

- Interact with a conversational agent
- Identify what AI can learn
- Brainstorm implications of bias in AI

Connection to Common Core

ISTE standards for CS Educators

- Standard 1. Computational Thinking - Learner (Understand the software component of computing systems)
- Standard 2. Computational Thinker - Facilitator (Educators engage students in identifying problems that can be solved using computational thinking)
- Standard 6. Equity and Inclusion - Advocate (Address a diverse range of ethical, cultural, and social perspectives on computing)

CSTA K-12 CS Standards (Grade 3-5)

- 1B-IC-18 - Discuss computing technologies that have changed the world, and express how those technologies influence, and are influenced by, cultural practices.
- 1B-CS-01 - Describe how internal and external parts of computing devices function to form a system.

Digital Literacy and CS (Massachusetts only)

- how computing affects society (for example, privacy and the security of information)

Tools and Materials

- Laptop (3-5 students per device per facilitator)
- Pencils (1 per student)
- Pre-activity assessment
- Post-activity assessment
- Worksheet
- Teacher resources

Summary

This small-group, learning activity focuses on learning how an AI can learn through teaching a conversational agent. As a demonstration, this lesson allows the teacher and students to explore and experiment with teachable AIs using a voice-based input and a browser-based UI that displays the mental model of the AI. The teacher will explain what to teach Zhorai, show what Zhorai has learned, and facilitate student discussion on how Zhorai can be biased. The

lesson will conclude by asking students to imagine what happens if AIs are biased in the real world, and how that would impact society.

Agenda

- Warm-up (10 minutes)
- Module 1: Understanding Knowledge Representation in Machines (10 minutes)
- Module 2: Teaching New Knowledge to Machines (10 minutes)
- Module 3: Witnessing Machine Learning Classification (10 minutes)
- Module 4: AI & Ethics Discussion (15 minutes)
- Cool down (5 minutes)

Assessment

Demonstration of Learning

- Students can interact naturally with Zhorai
- Students can understand how Zhorai organizes its knowledge (i.e., knowledge representation)
- Students teach Zhorai new knowledge

Classroom Discussion

- Student-generated ideas of what to teach Zhorai (i.e., what facts to feed to Zhorai's training model)

Oral Responses

- Students can explain how Zhorai makes decisions in Module 2
- Students can explain how AI (like Zhorai) can have bias and real-life implications of bias

Written Responses

- Students complete pre-activity worksheet to demonstrate any existing knowledge about AI and ML
- Students complete post-activity worksheet to demonstrate mastery of learning activity concepts

Teaching Guide

Warm-up (10 minutes)

Teacher begins class by assessing students what they know about learning in AI.

Potential discussion questions:

- How do you think Alexa or Siri works?
- Can Alexa learn new things?
- Can we teach Alexa new things?

Teacher then uses computer to introduce Zhorai. Zhorai has a built-in self-introduction module where Zhorai introduces what planet they're from (i.e., small talk) and students have a chance to build rapport with Zhorai the fictional character. The following script is an example of what this dialog might look like:

Moderator: Today we're going to talk to an alien from another planet! It's name is Zhorai, and it's actually a computer program, or *conversational agent* that MIT students created. Would someone like to say hi to Zhorai?

(The moderator chooses a student, who comes up to the computer and clicks record.)

Student: Hi!

Zhorai: Hi there! What's your name?

Student: My name is Erin.

Zhorai: Erin, what a nice name. Where are you from?

Student: I'm from Boston.

Zhorai: Interesting! I'm from planet Igbruhmmelkin. I've never heard of Boston before. Can you tell me more?

(A button appears with the words, "Find out what Zhorai knows about your planet!", which takes the class to the next module.)

Module 1

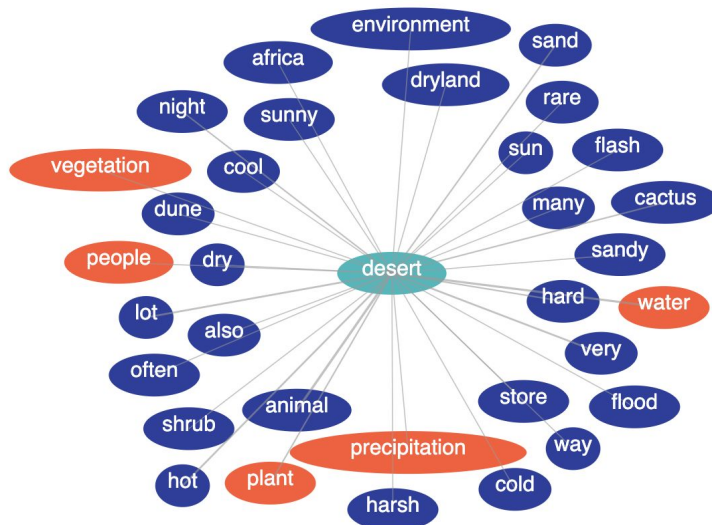
1. After the students have a chance to interact with Zhorai and familiarize with the controls (record and listen buttons), the teacher then states, "This is Zhorai's first time on Earth and Zhorai made a lot of observations about Earth's ecosystems. Does everyone know what an ecosystem is? (Zhorai knows about the desert, ocean, grassland, forest, and tundra). Let's look inside Zhorai's brain and see how it thinks!"
2. The teacher guides students into making observations about the ecosystem mindmaps. The teacher records student asking, "Zhorai, what do you know about deserts?"
3. After each ecosystem, the teacher shows the students what Zhorai's mental representation of each ecosystem is (i.e., the mindmaps).
4. The teacher shows them the sentences that Zhorai's model was trained on and asks, "What do you notice about these sentences and how it connects to Zhorai's mindmap for [ecosystem]?"

The following dialog illustrates what Module 1 might look like:

Moderator: Zhorai already knows a little bit about planet Earth. Let's ask it about what it knows. Who wants to ask Zhorai what it knows about the desert?

Student: What do you know about deserts?

Zhorai: Here's what I know about deserts. They're fascinating! (A mindmap about deserts appears, as shown below. This mindmap was created using the sentences that appear below the mindmap.)



The desert mindmap.

Zhorai heard these things about deserts from other Earthlings:

The desert has lots of sand
 The desert is very dry
 Deserts have cactus
 The desert is very hot and sunny
 Deserts don't have much water
 Deserts have few people
 Deserts have flash floods

A few of the sentences that created the desert mindmap.

Moderator: Why don't we ask Zhorai what it knows about oceans?

Student: What is an ocean?

Zhorai: Oh yes, oceans are very interesting. Here's what I know. (Another mindmap and list of sentences appears.)

Moderator: What other ecosystem can we ask Zhorai about?

Student: What do you know about plains?

Zhorai: Hmmm, I haven't heard about that ecosystem before, but I know about rainforests.

The class discusses why Zhorai knows only about some ecosystems. This is because Zhorai was only given data about specific ecosystems. In other words, Zhorai's *machine learning model* was only *trained on* particular data.

Module 2

1. The teacher introduces the second module by asking students, "Let's think about an animal that lives in one of these ecosystems. How might Zhorai know where this animal lives without us telling Zhorai explicitly?"

2. Let's try it! Teacher guides students into describing animals without saying the animal name (i.e., Taboo).
 - a. Here are some example sentences that enable Zhorai to classify the animal correctly. Notice how descriptive they are. Try to have students talk about the climate and vegetation that the animal lives around.
 - i. Bees fly around from plant to plant spreading pollen
 - ii. Birds fly from tree to tree
 - iii. Camels live in hot dry places with lots of sand and they don't need much water
 - iv. Cows graze in fields and eat lots of grass
 - v. Dolphins swim around in the sea and blow water out of their blowholes
 - vi. Fish are found in wet places with seaweed
 - vii. Polar bears live near icebergs in the arctic
 - viii. Whales are the largest animal in the sea and they blow water out of their blowhole
3. Afterwards, teacher shows the students the mindmap that Zhorai builds for the animal they discuss.

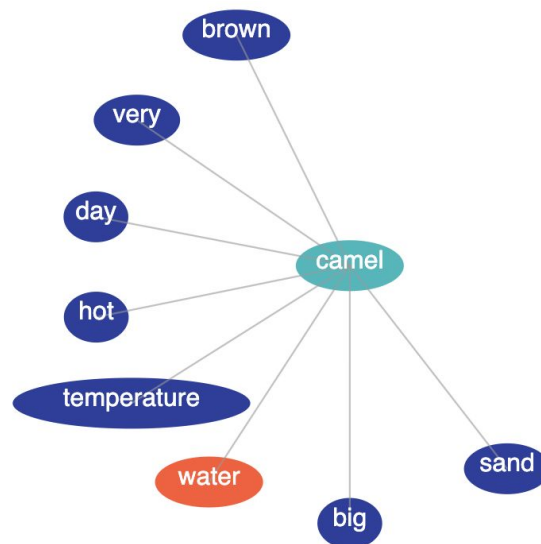
Moderator: Zhorai knows a little bit about ecosystems, but doesn't know about what animals live in each ecosystem. Let's teach Zhorai about these animals.

(In this module, there is a prompt on-screen that says, "Zhorai would like to know about *camels*. Could you teach it about them?".)

Student: Camels are big and brown. They walk on the sand all day. Camels live in hot temperatures and they drink very little water.

Zhorai: Wow, camels sound really interesting! Let me think for a bit and then I'll show you my thoughts.

(A mindmap about camels appears on screen, as shown below.)



An example mindmap for camels.

(A new prompt appears on the screen stating, "Zhorai would like to know about *fish*. Could you teach it about them?".)

Student: Fish live in the water. They are very small and swim all day.

Zhorai: Fish sound fascinating! Now I want to visit earth and all of it's life! I'll show you what I understand after I think for a little while.

The students then discuss how well they think Zhorai understands each animal.

Module 3

1. Teacher introduces third module by asking students, "Now, let's see how Zhorai makes decisions on which ecosystem [animal] lives in!"
2. Teacher shows the students the scatterplot Zhorai builds for comparing how close an animal is to the ecosystem representation. The students should notice that the Zhorai chooses the ecosystem that is the closest in proximity to the animal's datapoint.

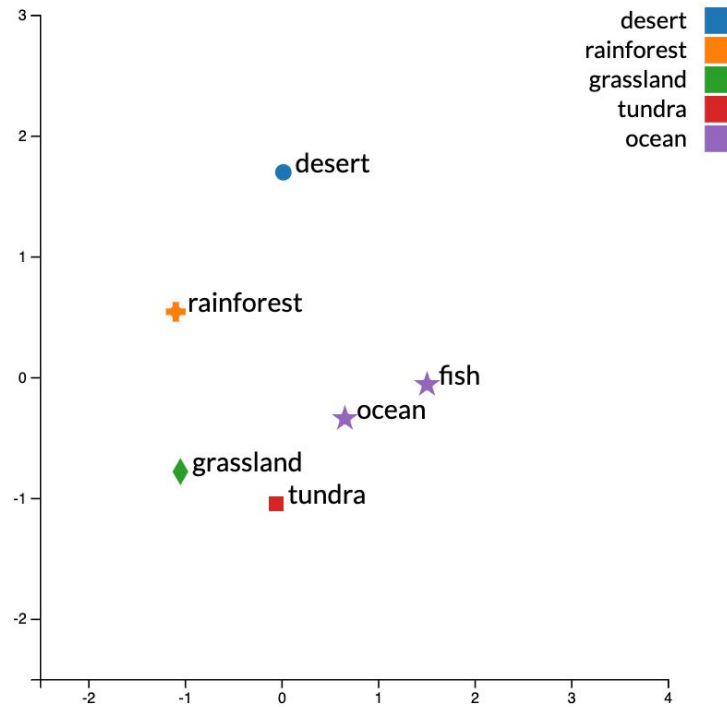
The following script illustrates example dialog for Module 3:

Moderator: Now that Zhorai knows about ecosystems and animals, let's see if it can guess which ecosystems certain animals belong to. Who would like to ask Zhorai to guess?

Student: Where are fish from?

Zhorai: Oh yeah, fish sound interesting. Let me think about where they might be from.

Zhorai: Based on what I know about Earth, I would guess fish live in oceans. (A scatter-plot appears on screen with fish and oceans close together, as shown below. The mindmap about fish from Module 2 also appears on screen.)



An example scatter-plot for fish.

Moderator: Let's think about this for a second. Zhorai is a computer program, so it uses numbers to think. This scatter plot shows what Zhorai understands about ecosystems and fish. You can see that there are numbers on the sides of the graph. Based on this graph, why do you think Zhorai thinks fish come from oceans?

Student: Because fish are really close to the ocean on the graph!

Moderator: Exactly! The distance (which can be represented by a number) between fish and the ocean ecosystem is the smallest. Let's see if Zhorai can guess where another animal comes from.

Student: Guess where polar bears are from.

Zhorai: I'll think about polar bears for a bit and let you know!

Zhorai: Hmm, I'm not sure! I haven't heard much about that animal.

Moderator: Since we didn't teach Zhorai about polar bears, it doesn't have any idea where they are from.

Each classification will be accompanied by a discussion about how Zhorai represents knowledge and why it has succeeded or failed at the classification task. Zhorai may fail at the classification task if it is not given enough descriptors to place the animal correctly in the *embedding space* (which is the high dimensional space that is simplified and illustrated by the scatter-plot). Zhorai needs a sufficient amount of description words about animals to discern where the animals might live.

Module 4

1. Teacher can discuss how Zhorai works:

a. How does Zhorai learn and perceive the world?

- i. Zhorai learns through being given information. It does not know anything without being given information. For example, Zhorai only knows about ecosystems because it was previously given information from the internet about ecosystems. Since Zhorai was not given information specifically about animals, it only knows about animals if it is told about them.
- ii. Zhorai can perceive the world by using a microphone. The microphone converts the sound waves to numbers, the numbers to phonemes, the phonemes to strings of letters and words, and the words to sentences. The sentences are then *parsed* or sectioned into parts and can be represented by a mindmap.

b. How did Zhorai learn English?

- i. Zhorai learned English in a similar way to how it learned about ecosystems: It was given many examples of English words being spoken. Through the audio files and transcriptions of the words being spoken, it learned how to replicate the audio into realistic-sounding speech. The transcriptions are usually manually created by hand, so the computer is very dependent on humans to learn how to speak.

c. Is it a similar way to how it learned about ecosystems?

- i. Yes, it is similar because the system required a lot of data (speech examples, in this case) to learn how to speak. The system also required a lot of ecosystem data to learn about Earth.

d. How did Zhorai hear us?

- i. Zhorai heard us through a microphone, which converts sound waves to digital signals and numbers.

e. What does the computer “see” or perceive when it gets information from a microphone?

- i. The computer “sees” an array of numbers. You can imagine this as a graph with time on the x-axis and pitch (or sound) on the y-axis. The resulting graph looks like a “wave”. This is what we call a “sound-wave”.

2. Teacher can discuss how Zhorai learns:

a. What else could we teach Zhorai?

- i. Zhorai is currently programmed to learn only about ecosystems and animals, but it could be programmed to learn about anything, like common conversational agents, such as Siri and Alexa. We could teach these conversational agents about anything.

b. Can we teach it about different cultures?

- i. Sure, why not?

c. How would that be different?

- i. We would be talking about people and their cultures as opposed to animals and their environments. Then, biases would be much more apparent.
 - d. Would Zhorai know whether what we teach it is correct?**
 - i. No, Zhorai would not know whether we teach it is correct unless we told it (or programmed it to know) that it is not correct. For example, we could teach it that polar bears are tiny animals that like to eat rocks or we could teach it that Canadians are people who drink maple syrup for breakfast.
 - e. How would you feel if Zhorai learned something untrue about your culture?**
 - i. For example, how might you feel if Zhorai learned that all Americans are mean and nasty people?
 - ii. Some artificial intelligence agents have *biases*, just like humans do, since they learned from human-created data. It is important for the developers of these systems to minimize these biases to ensure all people are treated equally by the systems.
3. Teacher can discuss implications of Zhorai's mistakes:
- a. Why were there animals whose ecosystem Zhorai could not guess?**
 - i. Maybe some animals weren't explained thoroughly
 - ii. Maybe some animals live in more than one ecosystem
 - iii. Maybe Zhorai does not know enough about a certain ecosystem
 - b. Is Zhorai's knowledge of the world biased?**
 - i. Yes. First of all, Zhorai only knows about the ecosystems that were taught to it. It cannot classify animals to ecosystems that were not taught to it.
 - ii. In addition, there is bias between the ecosystems that it knows. It is possible that Zhorai knows a lot more about one ecosystem than the other. In that case, if an animal isn't explained thoroughly, it is more probable that Zhorai will guess one ecosystem over another. If the amount of knowledge about each ecosystem isn't the same, then they don't all have a fair shot at being the prediction when the prediction is unclear.
 - iii. Lastly, Zhorai's knowledge of animals is biased to what has been said to it. For example, if students only talk about the what the animal eats, then Zhorai will be able to classify better ecosystems with clear vegetation and animals with a clear diet.
 - c. How can we improve this?**
 - i. By training Zhorai with much more varied data, we can minimize Zhorai's biases.
 - d. Why is it important?**
 - i. It might not seem so important that Zhorai is biased towards ecosystems or animals, but if we generalize this to AI used in our day to day life, we can understand how biased systems can affect society.
 - e. How can we cause Zhorai to make mistakes?**

- i. If we teach Zhorai incorrect details about animals, it will classify the animals incorrectly. For example if we say that fish live in trees, Zhorai might guess that fish live in rainforests.
4. Teachers discuss with students the ethical implications of what Zhorai knows and how Zhorai learns.
 - a. In continuation to the previous questions, explain how biased systems affect society. Explain why it is important that they are trained with data that represents all of society and not only parts so that they can work just as well for everyone.

Cool down (5 minutes)

Teacher ends class by assessing students what they have learned in the last hour.

- Use post-learning activity assessment worksheet.
- Other possible discussion questions:
 - How is Alexa different from Zhorai?
 - Alexa was trained and continues to be trained on corpuses of all different kinds of areas so that Alexa know not only about ecosystems, but about many concepts that humans discuss.
 - How many hours do you think people needed to talk to Alexa to get as good as it is now?
 - This is an open ended question, but Alexa has been trained on corpuses with billions of sentences.
 - How many people do you think talked to Alexa? Who do you think talked to Alexa?
 - In 2018, 39 million Americans owned Alexa. Alexa is regularly used by more than one user in a home. Also, these are only the American users.
 - All sorts of people talk to Alexa about all sorts of things. This is how Alexa knows so much about so many subjects.

External Resources

[Synthesizing speech](#)

[Machine learning and how computers learn to understand speech](#)

[How does Alexa work?](#)

[What is a semantic parser?](#)

[What is a word embedding?](#)