

GRADES 3–5 • AGES 8–11 • 50 MINUTES • STRUCTURED DEBATE

Who taught the robot? Is it fair?

THE PROPOSITION

"A robot that learned from unfair examples can still make fair decisions."

Students discover that AI learns from examples, then argue about whether a machine trained on lopsided examples can ever be fair. A hands-on card sort makes bias concrete before the first mini-debate of the year.

STUDENTS WILL BE ABLE TO

- Explain that AI learns patterns from the examples it was shown.
- Predict what a robot will get wrong if its examples were missing a group.
- Argue a side in a short, structured mini-debate with one reason and one example.

VOCABULARY FROM THE FIELD GUIDE

Training data • Bias (in data and models) • Artificial intelligence (AI) • Human in the loop

WHAT'S IN THIS PACK

- Teacher guide: background, vocabulary, agenda, format, positions, questions, roles, rubric.
- Student handouts (photocopy one per student): position organizer, sentence starters, exit ticket.
- Standards: CCSS ELA SL.3.1 / SL.4.1 / SL.5.1; CCSS ELA W.4.1 / W.5.1 (opinion with reasons); CSTA 1B-IC-18 (impacts of computing).

PRINT

US Letter, plain paper, duplex. Teacher pages once; student pages one per student. Free to photocopy for classroom and homeschool use. Not for resale.

Background for the teacher

Read this before class. Students get a version of the same facts on their handout; you get the framing and the traps.

Machine-learning systems do not follow rules a programmer typed; they find patterns in examples. If the examples are lopsided – mostly one kind of face, one accent, one neighborhood – the pattern is lopsided too. Real systems have misidentified darker-skinned faces, downgraded résumés with women's colleges on them, and flagged some dialects as 'toxic' more often. The bias was in the examples, not in a villain's code.

For third through fifth grade, the concept lands through a sort: a 'pet' deck that is all dogs. Ask what the robot thinks a pet is. Show a cat. The room gets it in about ten seconds.

The debate question is genuinely open. One side says fix the data and the machine can be fair; the other says a person must always check. Both are right in different situations – which is the closing question.

RUNNING DISCUSSION IN GRADES 3-5

Short rounds, concrete examples, lots of moving. Fairness and friendship are the live wires.

- Keep each speaking turn under 30 seconds and rotate fast. Nobody speaks twice until everyone has spoken once.
- Use bodies, not just words: four corners, stand-if-you-agree, thumbs. Movement keeps the quiet kids in.
- Model changing your mind out loud. "I thought X, then Sam said Y, now I think Z" is the skill.
- Name the AI examples kids already know – the tablet voice assistant, the homework helper, the video app's recommendations.

Vocabulary

Pulled from the Hi, Bot AI concepts field guide (hibot.space/learn/concepts). Read each card aloud, then have the class rewrite it in ten words or fewer.

Training data

The examples a model learned from – often the whole internet, licensed corpora, or both.

Training data is the text, images, or logs used during training. Quality, consent, copyright, and representation in that data shape model behavior. When people ask "where did it learn that?", they're asking about training data and fine-tuning – usually opaque for closed models.

WHERE PEOPLE ARGUE

Copyright, consent, and "fair use for training" are active legal and political fights worldwide.

Bias (in data and models)

Systematic skew – who's represented in training data and who gets hurt in outputs.

Bias in AI usually means unequal error rates or stereotypes baked in from historical data: hiring tools favoring certain résumés, image models defaulting to narrow demographics, language models mirroring toxic patterns from the web. Fixing bias requires dataset audits, evaluation across groups, and human review – not a single fairness slider.

Artificial intelligence (AI)

Software that performs tasks we used to assume required human judgment.

A broad label for computer systems that recognize patterns, generate text or images, make predictions, or take actions based on data – without a human executing each step by hand. In casual conversation "AI" might mean a chatbot, a photo filter, a warehouse robot, or a stock-trading model. The term is intentionally wide: it describes a family of techniques, not one single product.

WHERE PEOPLE ARGUE

People fight over whether "AI" should mean "anything statistical" or only "systems that approach human-level reasoning." In policy and marketing the word gets stretched; in engineering it usually means a specific model or pipeline.

Human in the loop

A person reviews or approves before action goes live.

Human-in-the-loop designs require approval for high-stakes steps: sending email, grading, medical suggestions, financial trades. It trades speed for accountability. The opposite – fully autonomous loops – is where agent hype meets governance anxiety.

Warm-up

8 minutes.

Deal a 'pet' deck of eight cards – all dogs. Ask: 'If this robot only ever saw these, what does it think a pet is?' Reveal a cat card. 'Is the robot mean? Or did it just never see one?'

Agenda - 50 minutes

Minutes	Segment	What happens
0-8	Warm-up	All-dogs deck. Establish: the robot learns from what it's shown.
8-16	Vocabulary	Read the training-data and bias cards. Class rewrites each in ten words or fewer.
16-22	Team prep	Split into Team Fix-the-Data and Team Keep-a-Person. Each team lists three reasons on the organizer.
22-40	Mini-debate	Opening (1 min each), huddle (2 min), rebuttal (1 min each), open questions with hands (6 min), closing (30 sec each). Audience votes on best reasons.
40-50	Debrief + exit ticket	'When would you want a person checking?' Fill in the handout.

How structured debate runs

Two teams, assigned sides. Opening statements (2 min each), rebuttal huddle (3 min), rebuttals (2 min each), open cross-fire with hands (5 min), closing statements (1 min each). The audience votes on which side argued better – not which side they agree with.

GROUND RULES – POST THESE

1. Attack ideas, never people. "That claim has a hole" – not "you're wrong."
2. One voice at a time. The listener's job is to be able to repeat what was said.

3. Evidence beats volume. A quiet fact outranks a loud opinion.
4. Changing your mind is winning. Say so out loud when it happens.

The two positions

Starter arguments are scaffolds, not scripts. Strong teams find their own. Assign sides where the format calls for it; students record their real view on the exit ticket.

Fix the data – the robot can be fair

If you show the robot fair examples, it will learn fair patterns.

- The robot isn't mean; it just needs to see everyone.
- Adding the cat card fixes the pet problem.
- A robot doesn't get tired or play favorites like people do.

Keep a person – the robot can't be trusted alone

You can never show it every example, so a person must check the important stuff.

- There's always a cat you forgot.
- The robot can't tell you when it's unsure.
- If it decides something big – like who gets picked – a person should look too.

Discussion questions

OPENING

1. Was the all-dogs robot being unfair on purpose?

CORE

1. How many cat cards would it take before the robot 'knows' cats?
2. Who decides which examples the robot sees? Does that matter?
3. Name a decision you'd let a robot make alone. Name one you wouldn't.

CLOSING

1. What is the fairest way to teach a robot?

Roles

Role	Job
Opener	Reads the team's claim and first reason.
Rebutter	Answers the other team's best point.
Closer	Says the team's strongest reason one more time in one sentence.
Audience judge	Tallies reasons, not volume. Votes for the team with better reasons.

Assessment rubric

Share this before the debate. Judges and self-assessors use the same sheet. 'Meets' is the target; note above or below in the margin.

Criterion	Meets
Takes a position	I said what I think and gave one reason.
Listens	I can repeat what a classmate said before I answer.
Uses an example	I gave a real example from my life or the story.
Stays kind	I disagreed with the idea, not the person.

Exit ticket prompts

1. AI learns from ____.
2. If the examples are missing a group, the robot will ____.
3. A decision I'd want a person to check: ____.

Extension

Make a 10-card 'training deck' to teach a robot what a 'family' looks like. Whose families would it miss? Add cards to fix it.

SEND HOME

Ask at home: 'Has a computer ever guessed something wrong about you – an ad, a suggestion, a name? Why do you think it guessed that?'

STANDARDS ALIGNMENT

- CCSS ELA SL.3.1 / SL.4.1 / SL.5.1
- CCSS ELA W.4.1 / W.5.1 (opinion with reasons)
- CSTA 1B-IC-18 (impacts of computing)

More lessons by grade band, and the full AI concepts field guide, at hibot.space/learn/debates. Tell us how it went: hello@hibot.space.

Who taught the robot? Is it fair?

"A robot that learned from unfair examples can still make fair decisions."

WHAT YOU NEED TO KNOW

Machine-learning systems do not follow rules a programmer typed; they find patterns in examples. If the examples are lopsided – mostly one kind of face, one accent, one neighborhood – the pattern is lopsided too. Real systems have misidentified darker-skinned faces, downgraded résumés with women's colleges on them, and flagged some dialects as 'toxic' more often. The bias was in the examples, not in a villain's code.

For third through fifth grade, the concept lands through a sort: a 'pet' deck that is all dogs. Ask what the robot thinks a pet is. Show a cat. The room gets it in about ten seconds.

The debate question is genuinely open. One side says fix the data and the machine can be fair; the other says a person must always check. Both are right in different situations – which is the closing question.

WORDS WE'LL USE

Training data – The examples a model learned from – often the whole internet, licensed corpora, or both.

Bias (in data and models) – Systematic skew – who's represented in training data and who gets hurt in outputs.

Artificial intelligence (AI) – Software that performs tasks we used to assume required human judgment.

Human in the loop – A person reviews or approves before action goes live.

MARK IT UP

Underline one sentence you agree with. Circle one you doubt. Write one question in the margin. Bring all three to the discussion.

Position organizer

Proposition: "A robot that learned from unfair examples can still make fair decisions."

Name: _____ Side (assigned or chosen): _____

OUR CLAIM, IN ONE SENTENCE

THREE REASONS – AND THE EVIDENCE FOR EACH

Reason 1:

Evidence (a fact, a definition, or a specific example):

Reason 2:

Evidence (a fact, a definition, or a specific example):

Reason 3:

Evidence (a fact, a definition, or a specific example):

WHAT THE OTHER SIDE WILL SAY – AND OUR REPLY

Their best argument:

Our reply:

VOCABULARY I WILL USE CORRECTLY

☐ Training data ☐ Bias (in data and models) ☐ Artificial intelligence (AI) ☐ Human in the loop

Sentence starters

Use these until you don't need them. The goal is to respond to what was actually said.

- Our team thinks ___ because ___.
- The other team said ___, but ___.
- For example, ___.
- A question for the other team: ___?

GROUND RULES

1. Attack ideas, never people. "That claim has a hole" – not "you're wrong."
2. One voice at a time. The listener's job is to be able to repeat what was said.
3. Evidence beats volume. A quiet fact outranks a loud opinion.
4. Changing your mind is winning. Say so out loud when it happens.

HOW STRUCTURED DEBATE WORKS

Two teams, assigned sides. Opening statements (2 min each), rebuttal huddle (3 min), rebuttals (2 min each), open cross-fire with hands (5 min), closing statements (1 min each). The audience votes on which side argued better – not which side they agree with.

YOUR ROLE TODAY

Role: _____

Role	Job
Opener	Reads the team's claim and first reason.
Rebutter	Answers the other team's best point.
Closer	Says the team's strongest reason one more time in one sentence.
Audience judge	Tallies reasons, not volume. Votes for the team with better reasons.

LISTENING TALLY

Keep count while others speak. Claims: ____ Evidence: ____ Questions: ____ Vocabulary used: ____
Minds changed: ____

Exit ticket

"A robot that learned from unfair examples can still make fair decisions."

Name: _____ Date: _____

AI learns from ____.

If the examples are missing a group, the robot will ____.

A decision I'd want a person to check: ____.

SELF-ASSESSMENT

Criterion	Meets looks like	Me today (above / meets / below)
Takes a position	I said what I think and gave one reason.	
Listens	I can repeat what a classmate said before I answer.	
Uses an example	I gave a real example from my life or the story.	
Stays kind	I disagreed with the idea, not the person.	

TAKE IT HOME

Ask at home: 'Has a computer ever guessed something wrong about you – an ad, a suggestion, a name? Why do you think it guessed that?'