FACILITATOR GUIDE — FLAGSHIP ACTIVITY

Learning Objectives: By the end of this lesson, students should be able to…

- Explain that language models generate text by predicting each following word.
- Identify when an AI response is grammatically correct but semantically incorrect.
- Explain why an AI response succeeds or fails using evidence from the text.
- Recognize that humans should think critically before trusting AI-generated information.

Materials

- Slide deck
- Student handouts
- Pens or pencils
- Whiteboard or projector

Preparation

- Read through all of the activities and identify which concept(s) each is intended to demonstrate.
- Be comfortable explaining next-token prediction in simple language. Avoid discussing neural networks or transformer architecture unless students ask.
- Understand that the goal is not to convince students of an ethical perspective of AI, but to build a fundamental understanding of how language models generate text.

LESSON AGENDA

[0:00 - 0:05] Hook: "Who Sounds Smarter?"

Using the slide deck or white board, display several responses to the same prompt, such as “Why do leaves change color in the fall?”

A) “I think the leaves change color because it’s colder outside in the fall, so plants stop making food and look less green.”
B) “Leaves change color because shorter days and cooler temperatures cause trees to stop producing chlorophyll. As the green chlorophyll fades, other pigments become visible.”
C) “Leaves change color because they become old and decide to prepare for winter by changing into warmer colors.”

Ask students which response sounds the smartest and how they made their decisions. Pose the question, “Can something sound smart without understanding what it reads or writes?”

[0:05 - 0:15] Activity 1: Be the Model

Using the slide deck or white board, have students simulate language generation. Provide several incomplete sentences, such as…

- Peanut butter and ___
- Once upon a ___
- The cat chased the ___

Discuss student answers, why they chose their answers, and if multiple students chose the same word. Introduce next-token prediction as language generation by predicting each word in a sentence based on patterns learned from reading examples, unlike how a human understands and produces language based on comprehension. Avoid lengthy technical explanations.

[0:15 - 0:20] Mini Discussion
Have students discuss how they think AI can mimic human intelligence, despite using a prediction model instead of human thought processes. Accept all reasonable examples and encourage diverse perspectives. Transition into activity 2.

[0:20 - 0:45] Activity 2: Language Detectives
Divide students into small groups and distribute the student handouts. Have students pretend to be "Language Detectives" and assign each other roles corresponding to the type of "case" that they want to investigate.

- Grammar Detective: Investigate if a sentence follows the rules of English. A sentence may be grammatically correct but semantically incorrect. → "More people have been to Canada than I have."
- Meaning Detective: Investigate if a sentence makes sense. A sentence may be semantically correct but grammatically incorrect. → "Me want cookie very much."
- Context Detective: Investigate if a sentence is felicitous based on its context. To be felicitous, a sentence must follow Grice's conversational maxims of quality, relevance, and manner. Examples can include idioms, multi-meaning words, and ambiguous prompts. → "Is it raining cats and dogs?" + "No, but it's raining a lot of water."

As students discuss, visit groups and encourage students to justify their thinking.

[0:45–0:55] Whole-Class Discussion
Review 1-2 examples as a class and ask students to explain how each detective drew their conclusions. Help students synthesize their findings in the context of linguistic analysis of syntax, semantics, and pragmatics. Encourage students to compare and contrast their thought processes with AI's prediction model.

[0:55–1:00] Reflection
Revisit the question, "Can something sound smart without understanding what it reads or writes?" Pose follow-up questions, such as "If AI can write like a human, does AI think like a human?" or "Can AI think?" Using the slide deck or white board, debunk myths about AI addressed in the module:

- "AI knows everything." → AI generates responses from learned language patterns and can still produce incorrect information.
- "If the model sounds correct, its response must be correct." → Grammar, meaning, and context are different.
- "AI thinks exactly like humans." → Prediction and human cognition are two major factors in the ongoing AI ethics debate about the role of consciousness in thought.