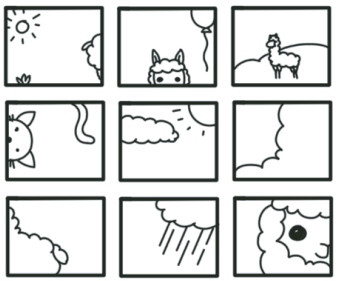
What is a llama?

Which pictures in this dataset have a llama?
How do you know?



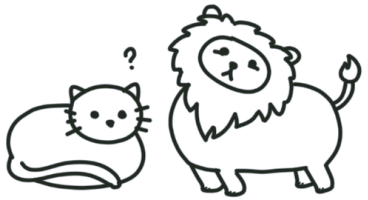
STEM Concept: **Dataset**

A dataset is a collection of data.
Data can be anything, like:

- photos you take with a camera,
- messages you send to other people,
- pictures created by artists,
- books written by authors,
- songs by musicians,
- videos or movies that people create,
- and anything that can be counted or measured!

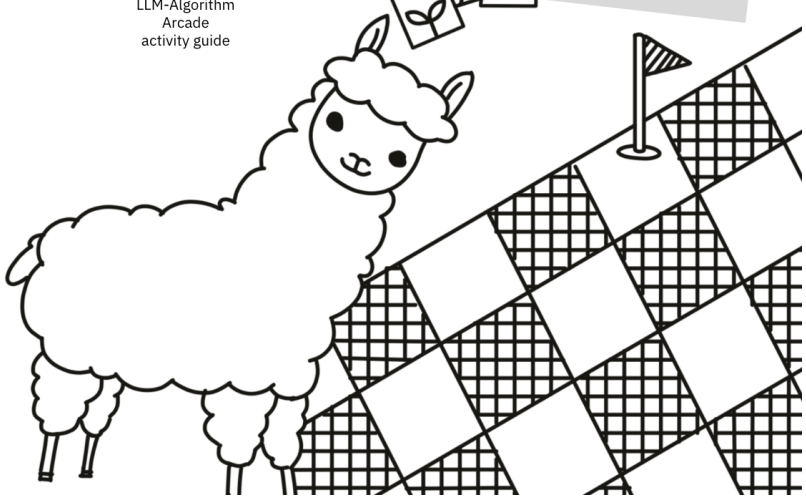
Learning step-by-step

Computer scientists create rules for machines to try to do things that people are good at. When they add many rules together and use them step-by-step, they create an **algorithm.**



When you draw a cat step-by-step, what parts do you draw first? What parts would you not draw?

LLM-Algorithm Arcade



Share your creations on Instagram and tag us @heyplix
plix.mit.edu

LLM-Algorithm Arcade activity guide

Are you sure? How sure?

Computer scientists use ***datasets*** of pictures, words, sounds, and videos with rules (***algorithms***) to teach machines to make a guess (***prediction***) about new data. This is called **machine learning.**

- Machine learning can make good guesses about
- if there are cats in pictures,
- or if you have messages from strangers,
- or if you might like a video on social media.

STEM Concept:
Confidence Score

Machine learning uses confidence scores from 0% to ***almost*** 100% to say how sure it is that the guess is correct.

What is the next word?

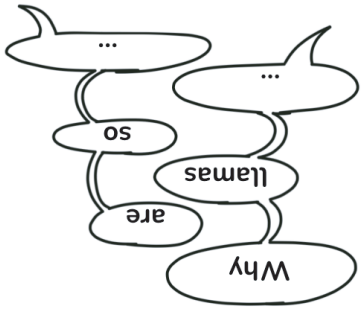
Once upon a _____
Rock, paper, _____
Far, far _____

STEM Concept: **Large language model (LLM)**

Large language models take datasets of words like books and messages and counts how many times two words show up together. It can make a guess about what word would come next!

LLMs are a type of machine learning.

Build a sentence together!



Play: With a friend, continue the sentence for as long as you can!

STEM Concept: **Word Vectorization**

LLMs turn pairs of words into numbers. This is called **word vectorization**.

Bigger numbers means the words go together more often. Artificial intelligence uses these numbers to build sentences one word at a time.