AI Curriculum

Teaching Material for the AI Curriculum for Elementary and Secondary Schools
Computer Science at Secondary Schools

01 Machine Learning Deck

Dataset

Introductory Discussion

People sort things into categories
to make the world more understandable.

CC
Midjourney

Alien Detective Agency

Searching for the origins of the little foundling.

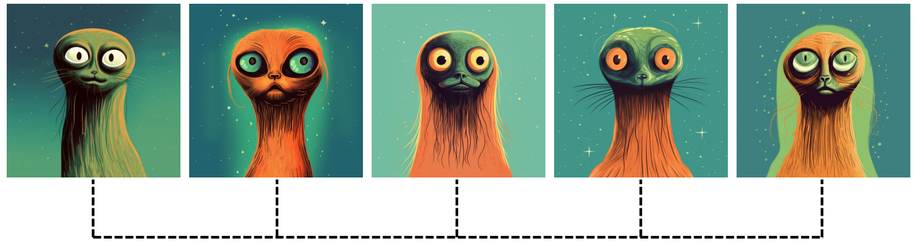
The Fluff Family

Earl family

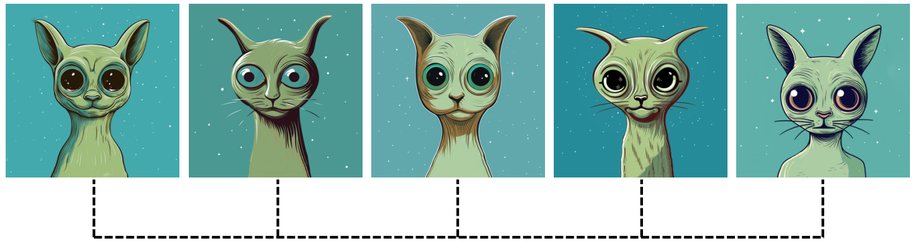
Dataset

A dataset is a collection of data used by people to train AI programs.

It can include images (but also audio, video, spreadsheets, and more). In our case, it consisted of portraits of aliens. If we want to teach an AI program to recognize different families, we first need to show it examples of individual family members and sort them into categories.



Fluffs



Earls

The program looks for different patterns on its own. Simplified, these could be things like ears, beards, fur, colors, and so on. So it might have found fur as a common feature in the Fluff family, and ears in the Earl family.

When we later present the program with new images it hasn't seen during training, it looks for similarities to what it learned before. Based on those, it decides which category the new image belongs to.

42%

probability that it's a Fluff, because fur seems to be the dominant feature

52%

probability that it's an Earl, because ears seem to be the dominant feature

See you again in the next lesson!

But this time, we'll let artificial intelligence decide where the foundling belongs!