

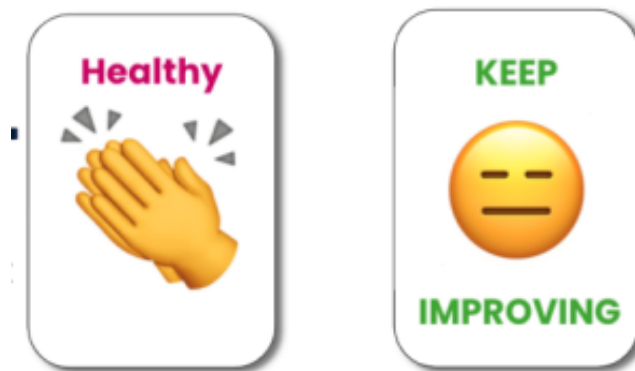
TECHNOVATION BEGINNERS: JUDGE THE DATASETS

This is how we would judge the datasets. There is no wrong answer!

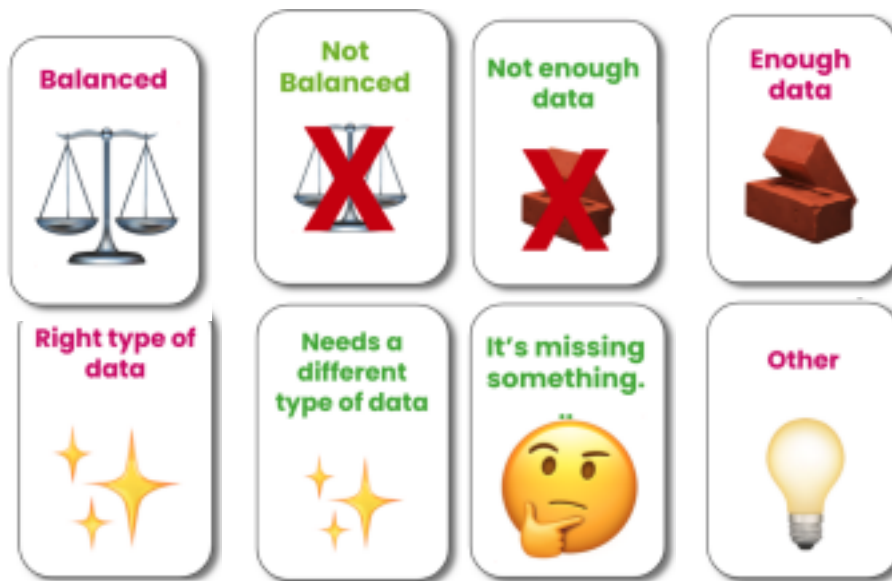
Goal: Judge if these datasets would make an AI model that could make good predictions.

Directions: For each dataset, you'll:

- Give a judgement: 🙌 **Healthy** or 😐 **Keep Improving**. Circle your choice.



- Then, pick 2-3 cards to describe the data by circling them on the page.



This dataset seeks to **identify birds by pictures of how they look.**

Healthy

Duck

20

Hawk

7

Sparrow

7

KEEP

IMPROVING

Not Balanced

Not enough data

Right type of data

While the images are the right type, there are many more (20) images of ducks than the other birds, and there could be more images!

Balanced

Enough data

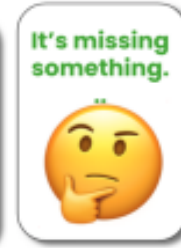
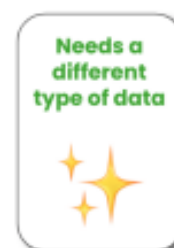
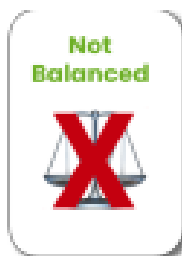
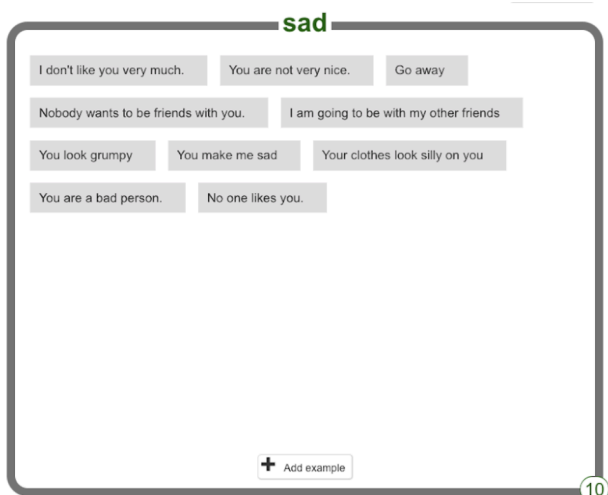
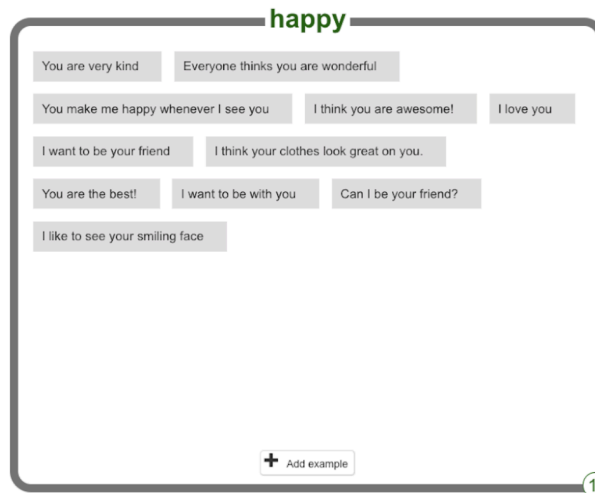
Needs a different type of data

It's missing something.

Other

This dataset seeks to **identify happy or sad sentences**

The sentences are balanced because there are the same number of sentences but 10 is not really enough to call it healthy. More sentences would make it healthier!



This dataset seeks to **tell you to throw away or recycle.**

Healthy

KEEP

IMPROVING

This dataset is hard to figure out because there are no labels! It probably could use more examples, and we can't tell if it's balanced or not.

Not
Balanced

Not enough
data

Right type of
data

It's missing
something.

Balanced

Enough
data

Needs a
different
type of data

Other