CLASSIFICATION BIAS FOR BEGINNERS

PyLadies - 2021

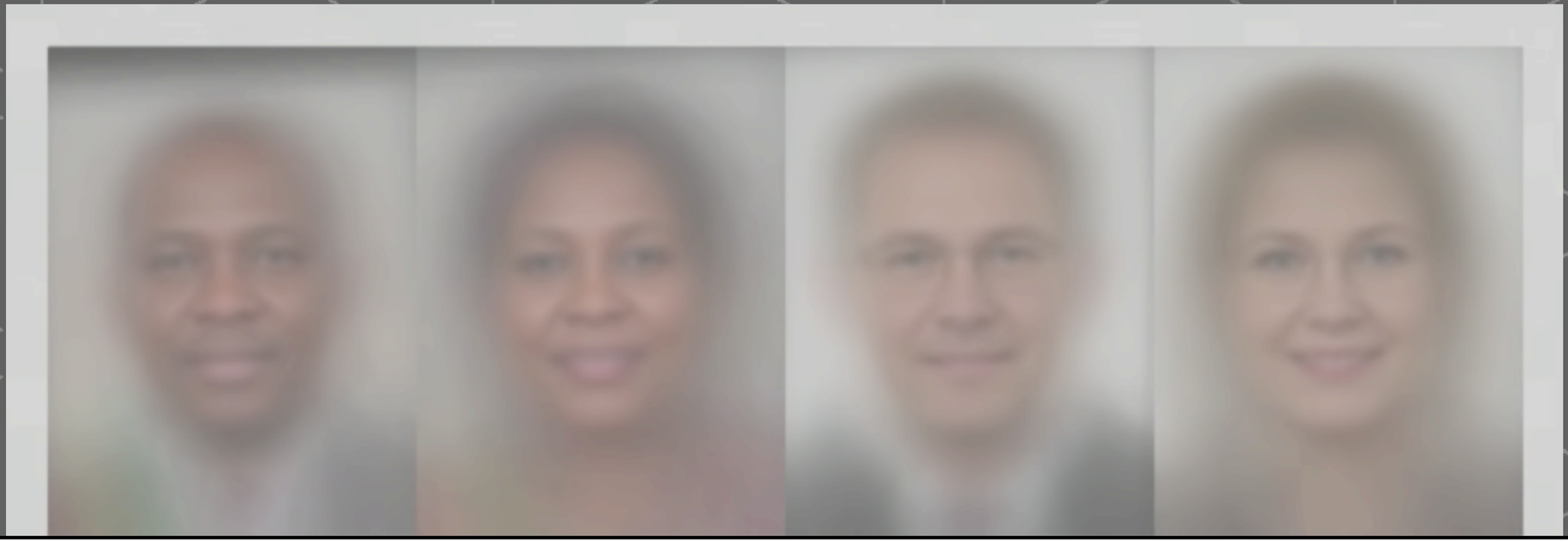
EMMA BEAUXIS-AUSSALET

e.m.a.l.beauxisaussalet@vu.nl
@emma_beauxis

eccentric, eccentric person, flake, oddball, geek: *a person with an unusual or odd personality*

- person, individual, someone, somebody, mortal, soul > anomaly, unusual person > eccentric, eccentric person, flake, oddball, geek

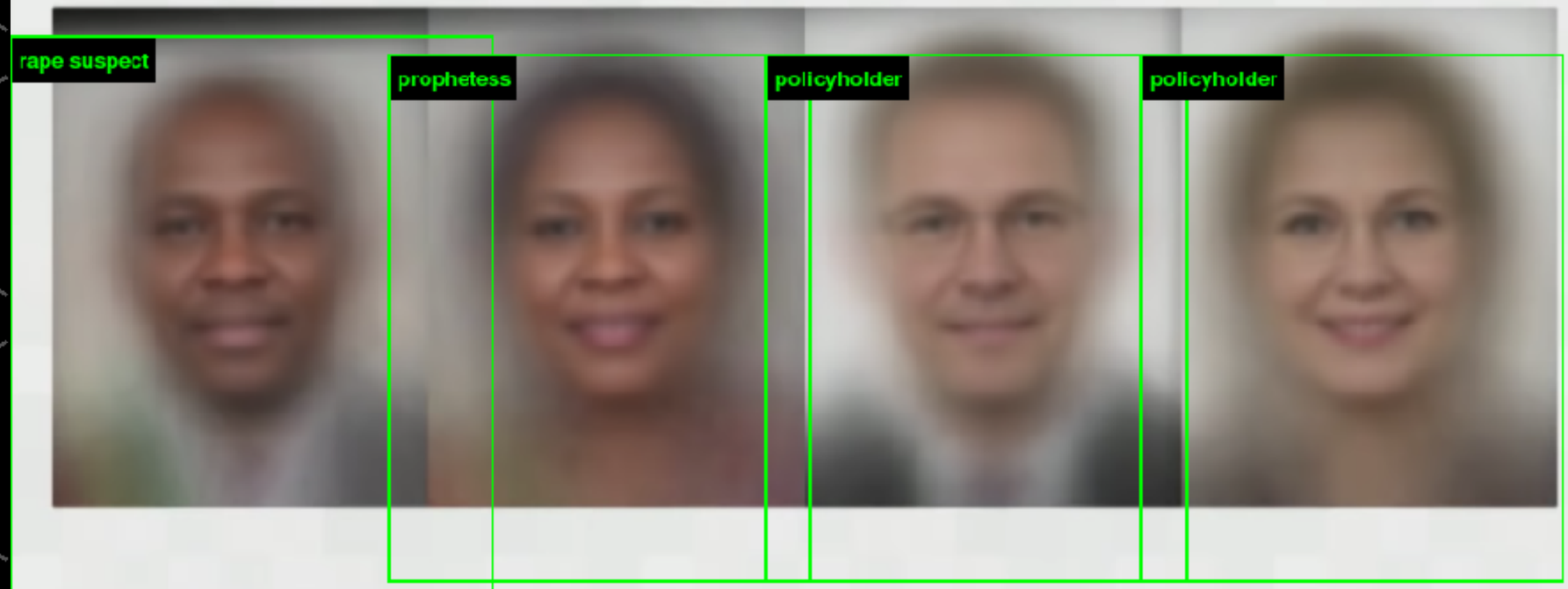
https://www.excavating.ai/



BOTH RANDOM ERROR AND SYSTEMATIC BIAS CAN HAVE SERIOUS IMPLICATIONS

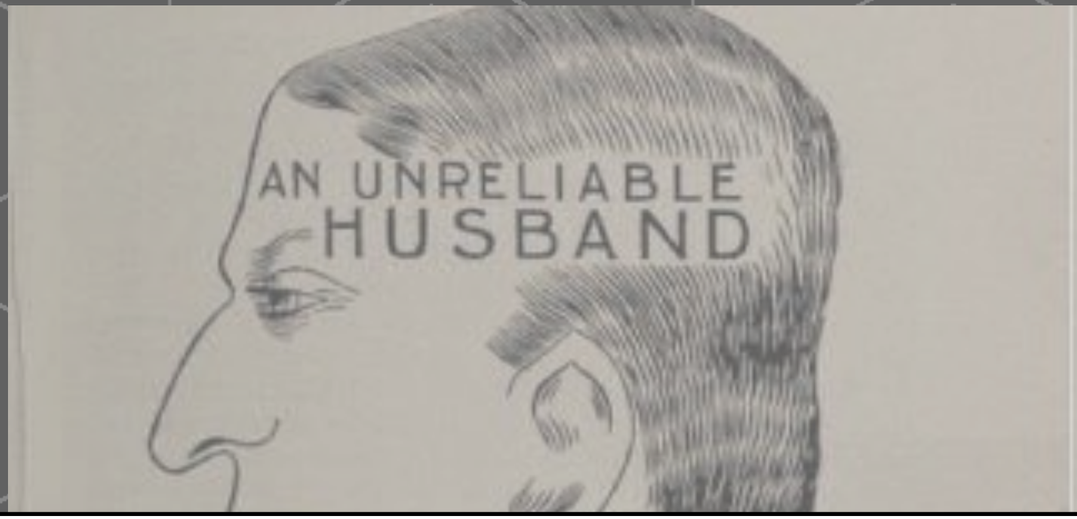
Microsoft	94.0%	79.2%	100%	98.3%	20.8%
FACE++	99.3%	65.5%	99.2%	94.0%	33.8%
IBM	88.0%	65.3%	99.7%	92.9%	34.4%

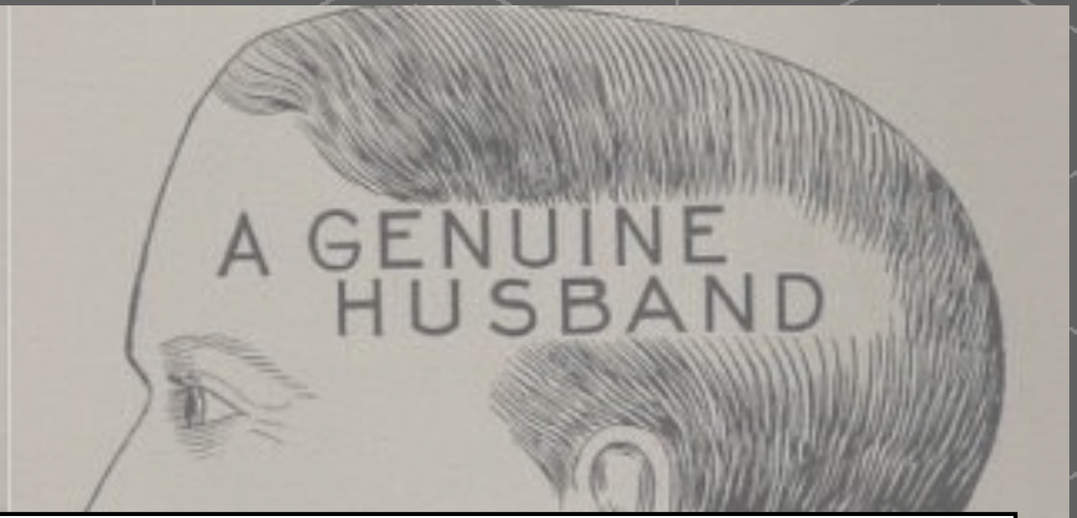
https://royalsociety.org/science-events-and-lectures/2018/07/you-and-ai-equality/



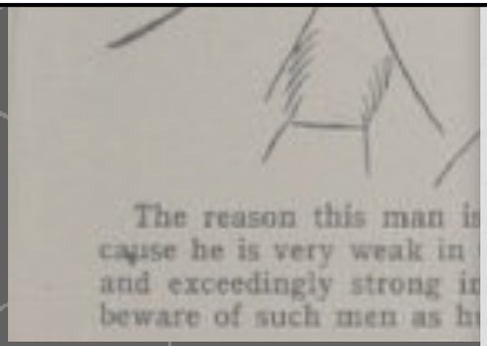
Gender Classifier	Darker Male	Darker Female	Lighter Male	Lighter Female	Largest Gap
Microsoft	94.0%	79.2%	100%	98.3%	20.8%
FACE++	99.3%	65.5%	99.2%	94.0%	33.8%
IBM	88.0%	65.3%	99.7%	92.9%	34.4%

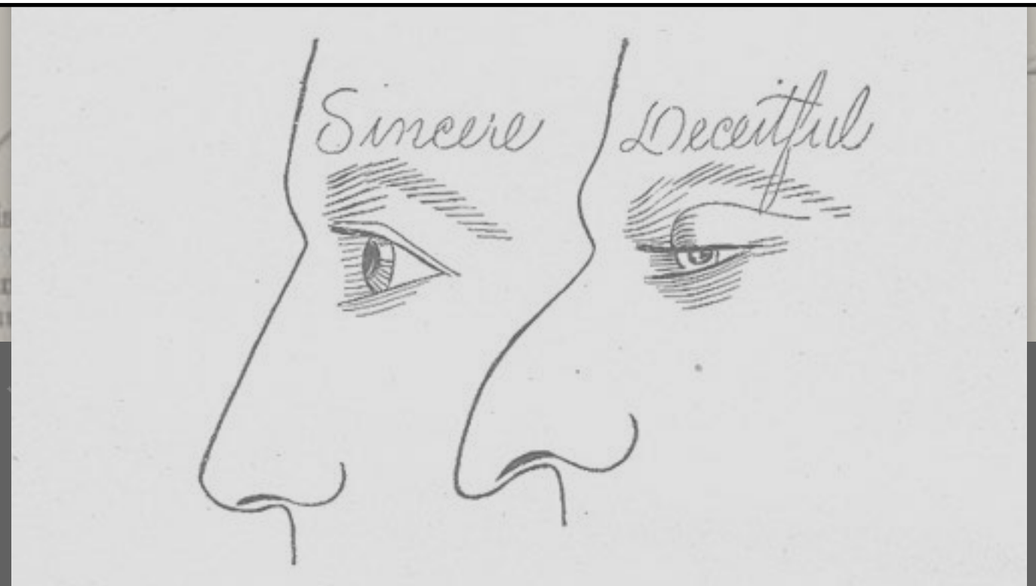
https://royalsociety.org/science-events-and-lectures/2018/07/you-and-ai-equality/



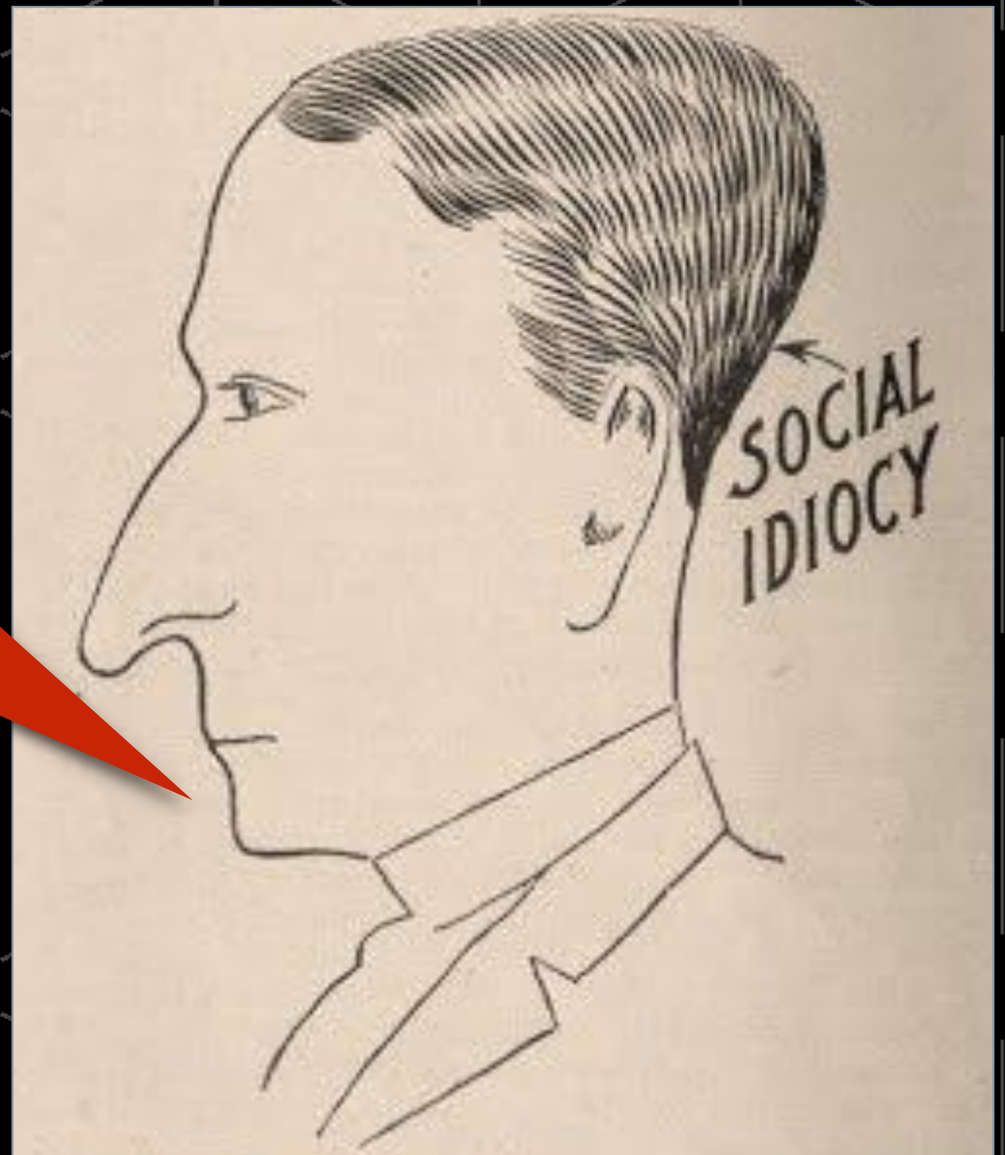


WE MUST NOT REPEAT PREJUDICES AT LARGE SCALE





I'm just
an engineer...



One may be strong intellectually and socially idiotic, as distinctly represented by this head and face.

We want all our readers to fix this head and face in their minds. It is a positive standard of constitutional laziness.

LAZINESS.

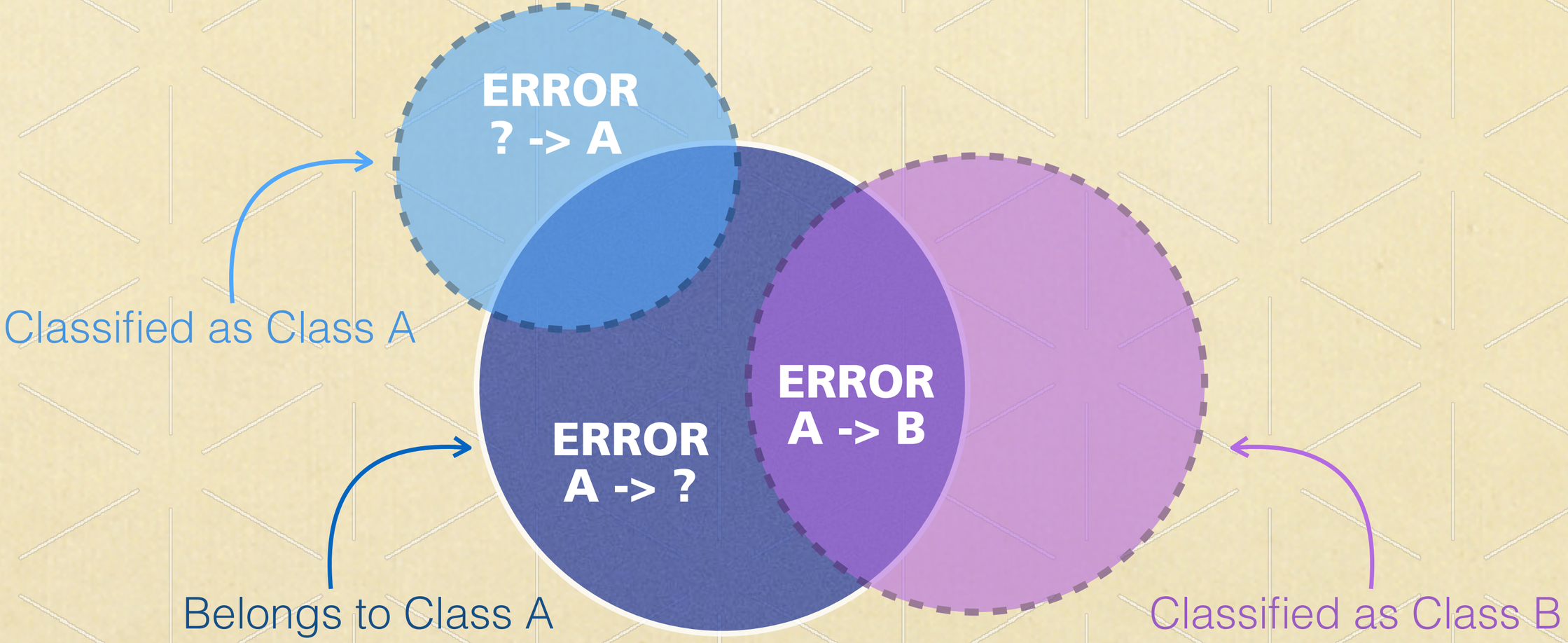
Genuine laziness is the result of deficient Destructiveness, Combativeness, Approbativeness, Acquisitive-

I'm not a nerd !

ERROR

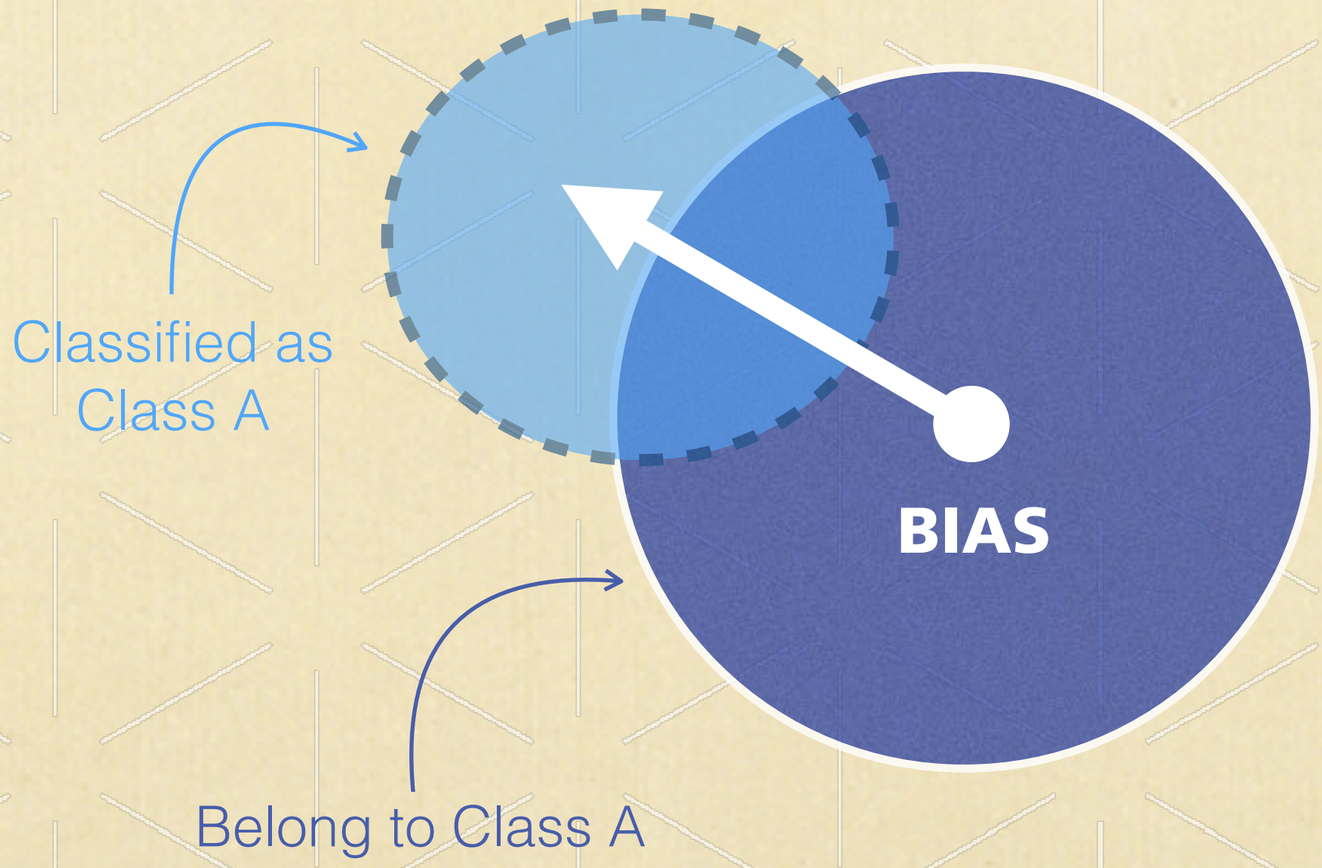
What belongs to Class A is classified as Class B.

It is **directional**: objects <u>from</u> Class A are misclassified <u>into</u> Class B.



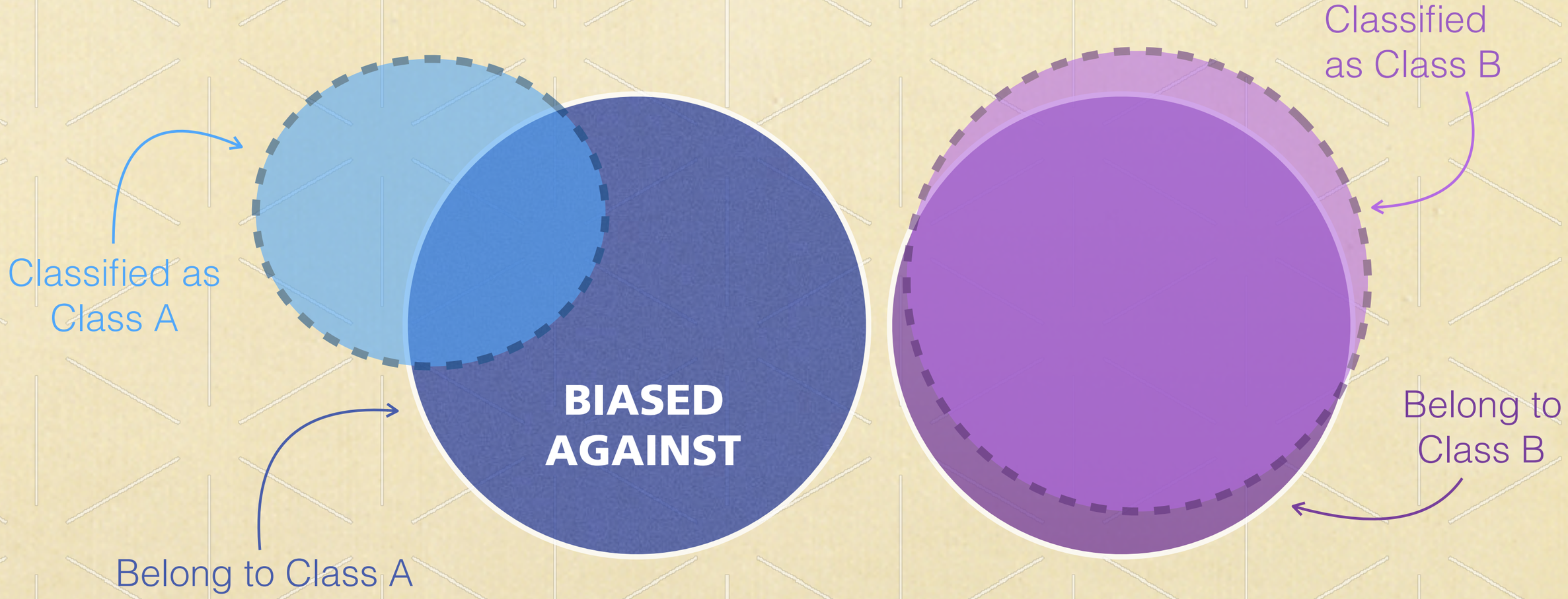
BIAS (1)

Errors systematically occur **between prediction and truth**.



BIAS (2)

Errors systematically differ **between classes**.



BIAS (3)

Errors systematically differ **between groups**.

BIAS (4)

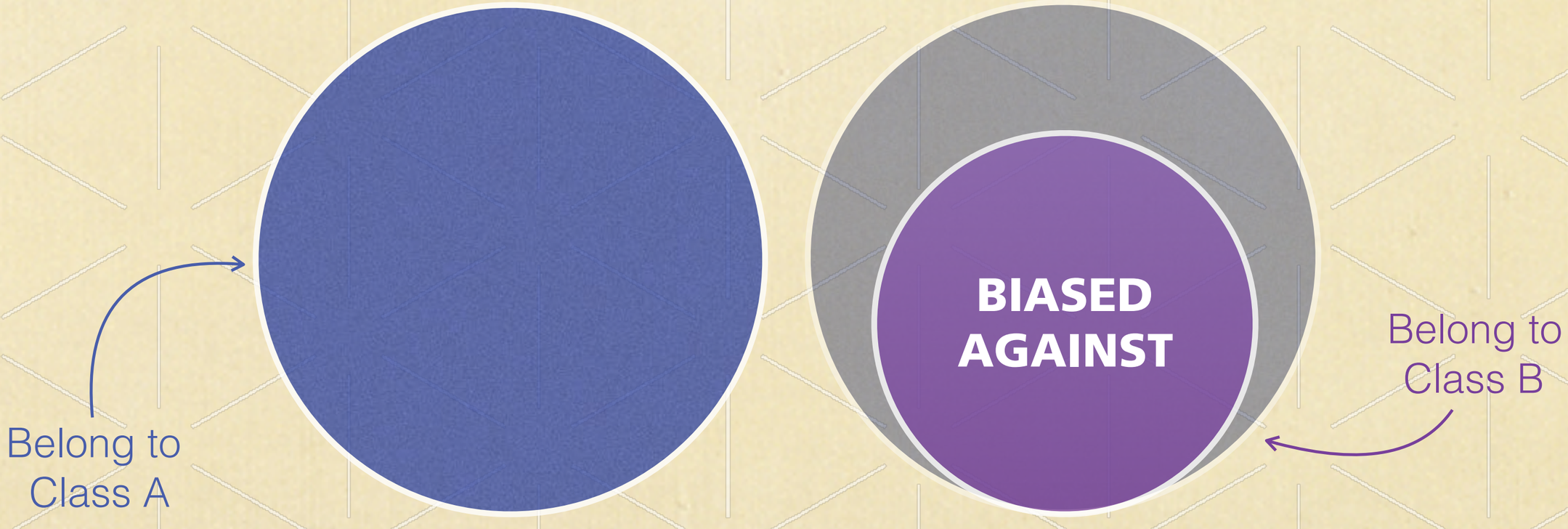
Groups are **under-represented**.

Female
belonging to Class A

Male
belonging to Class A

BIAS (5)

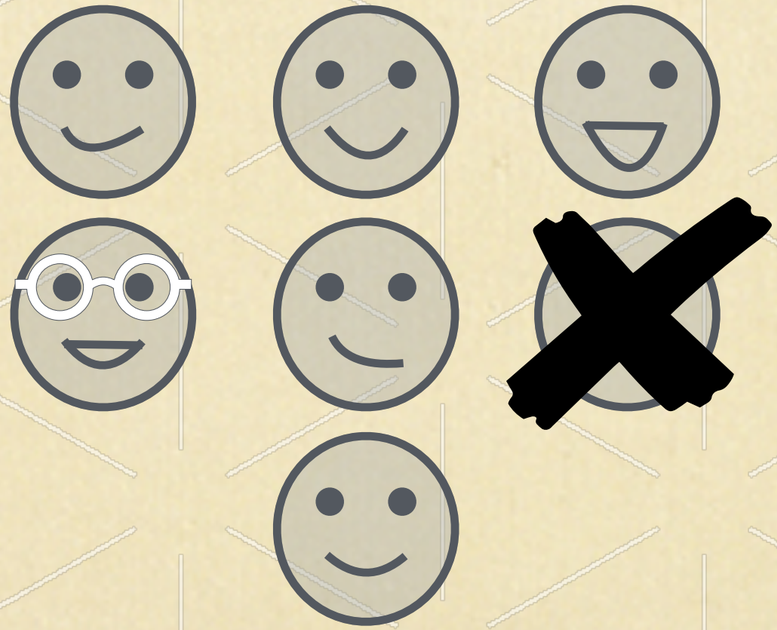
Classes are **under-represented**.



MEASURING THE ERRORS

Test sets contain examples of correct classifications, used to measure the errors.

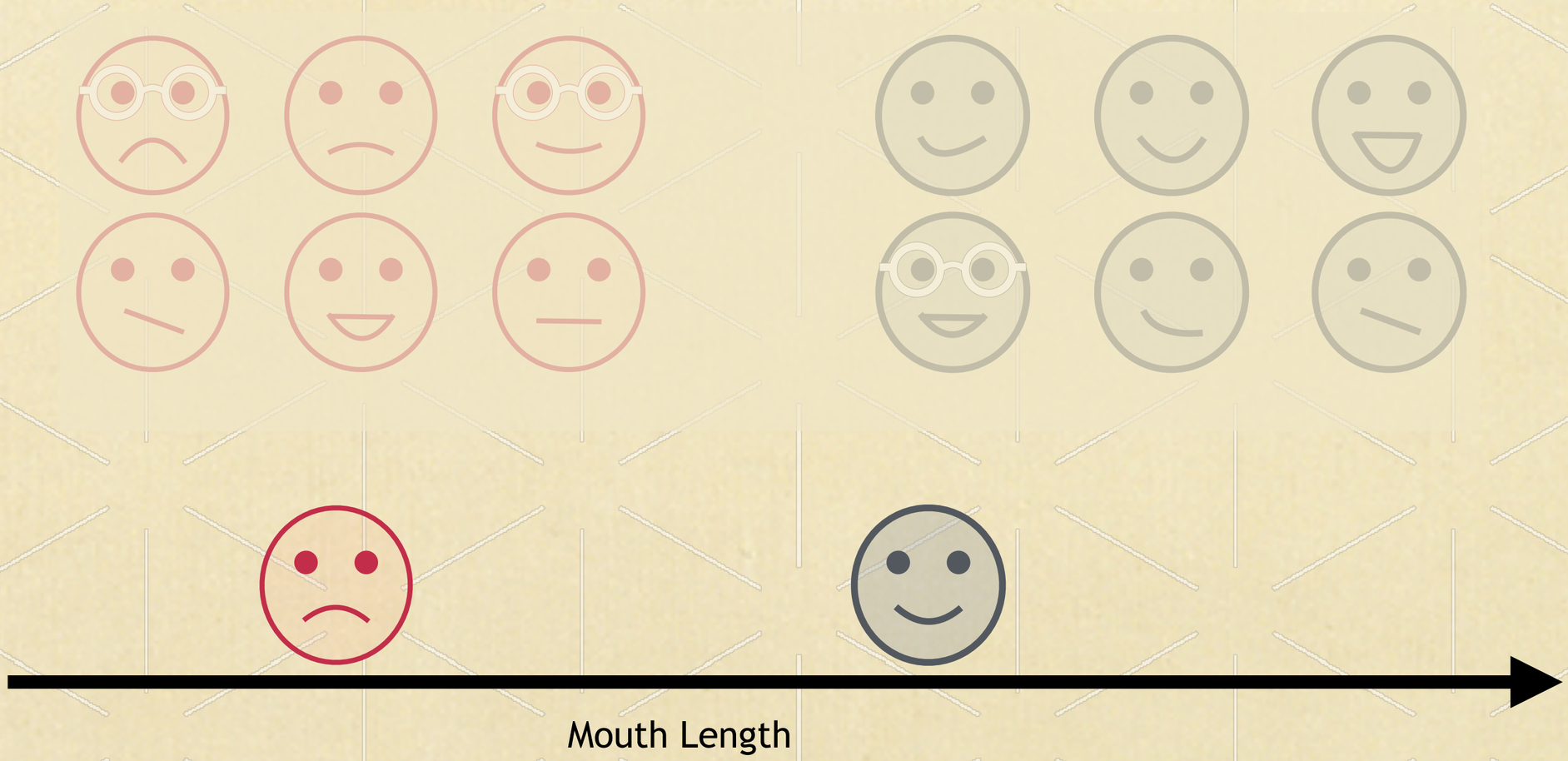
Examples of
Sad Customers



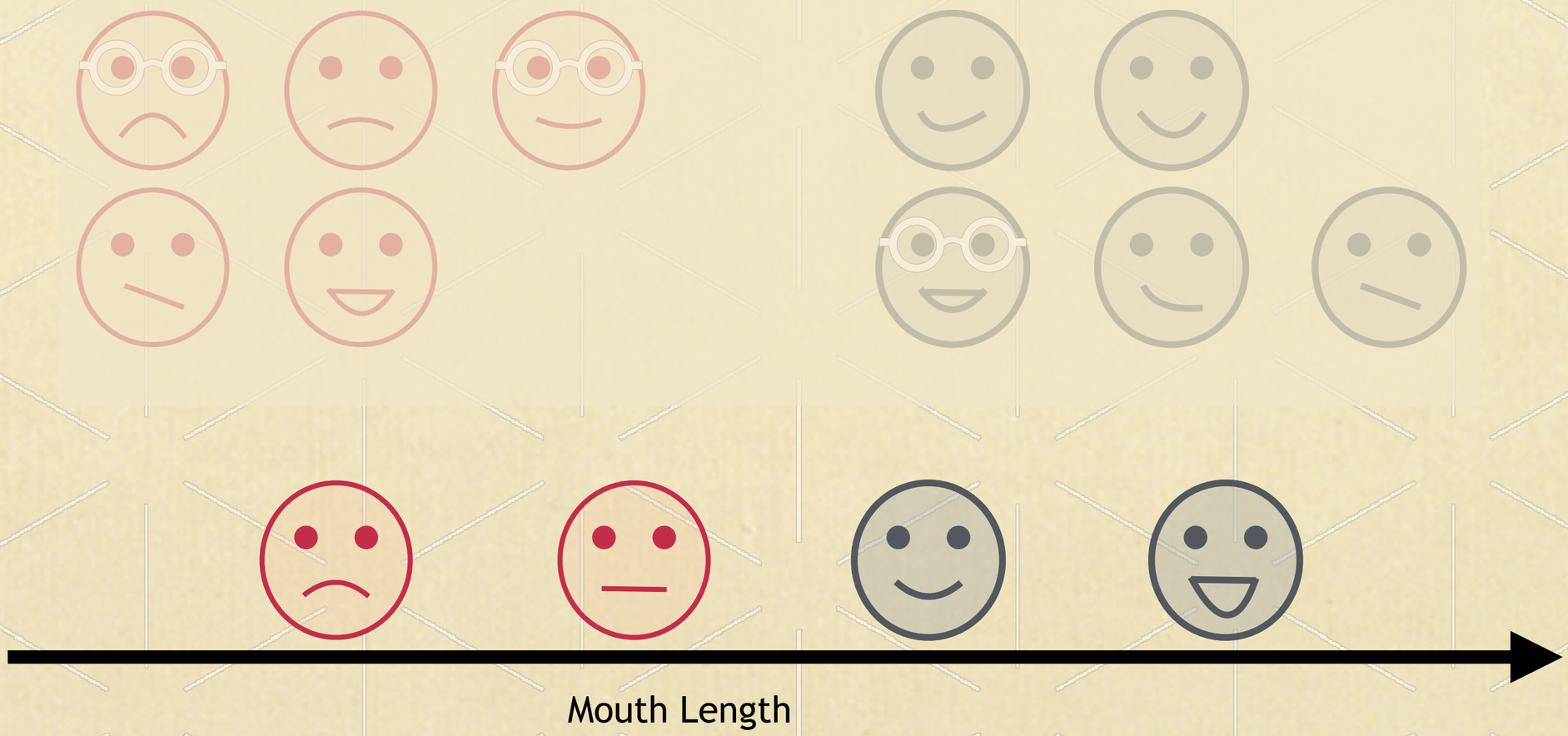
Examples of
Happy Customers

1-FEATURE CLASSIFIER

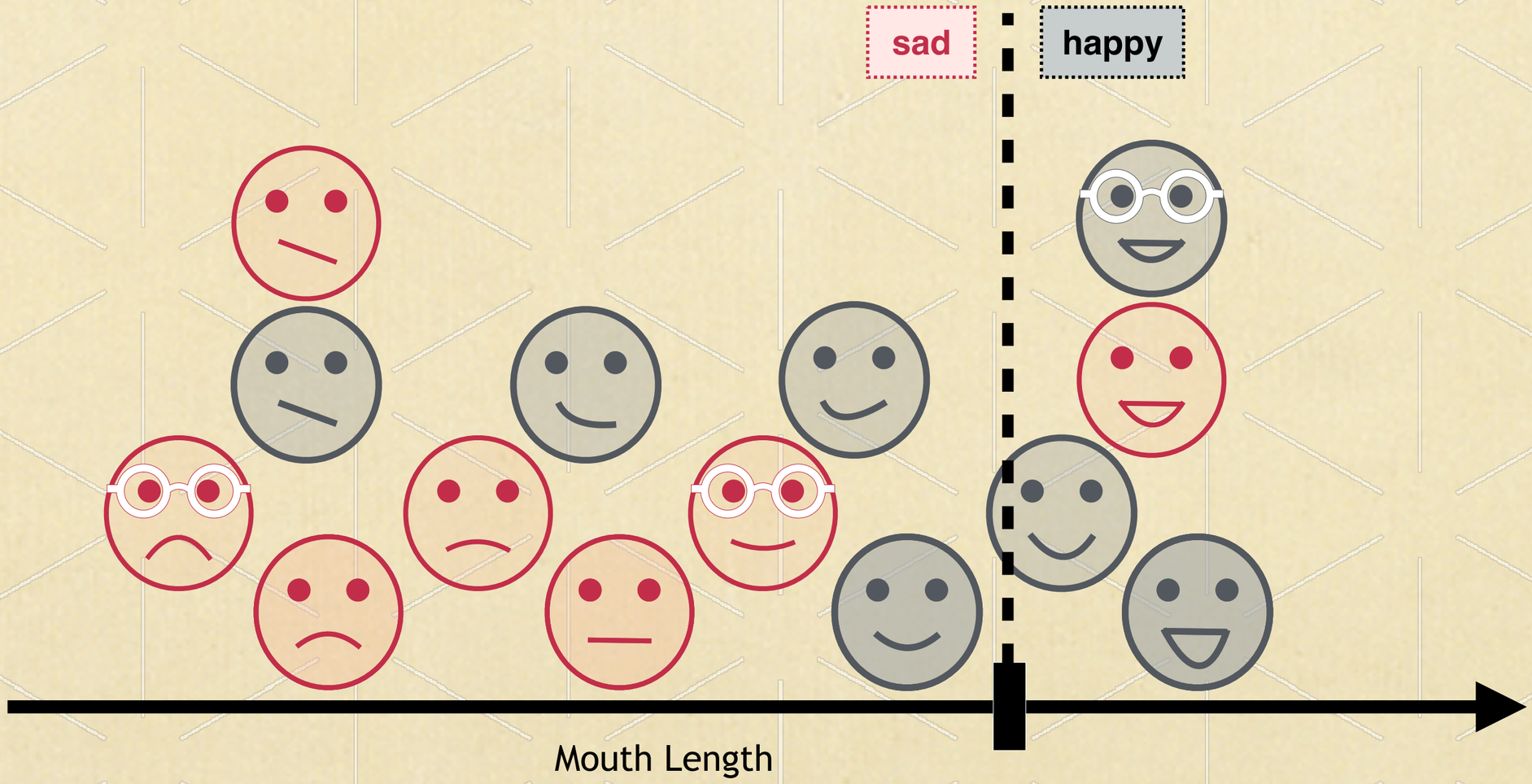
1-FEATURE CLASSIFIER



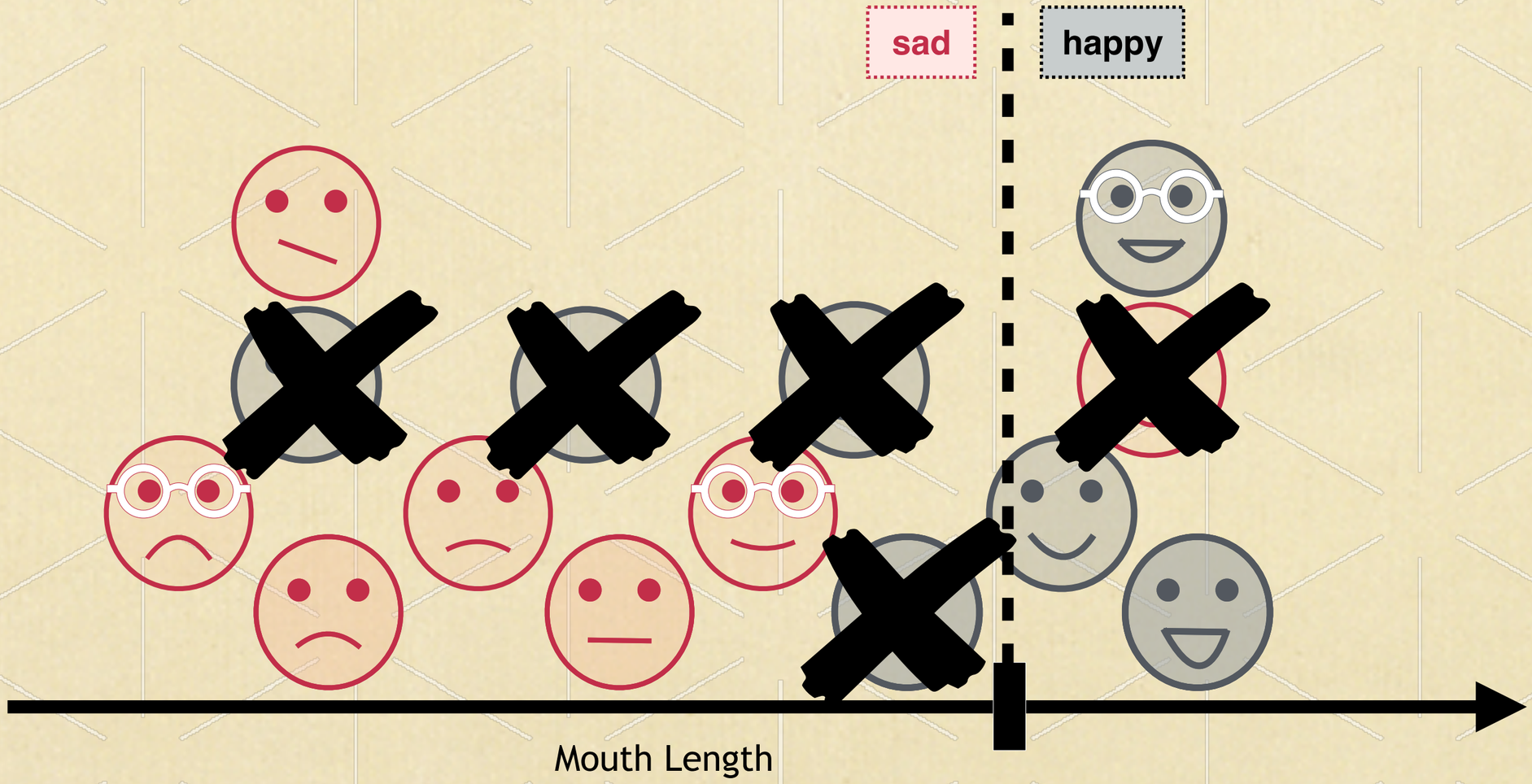
1-FEATURE CLASSIFIER



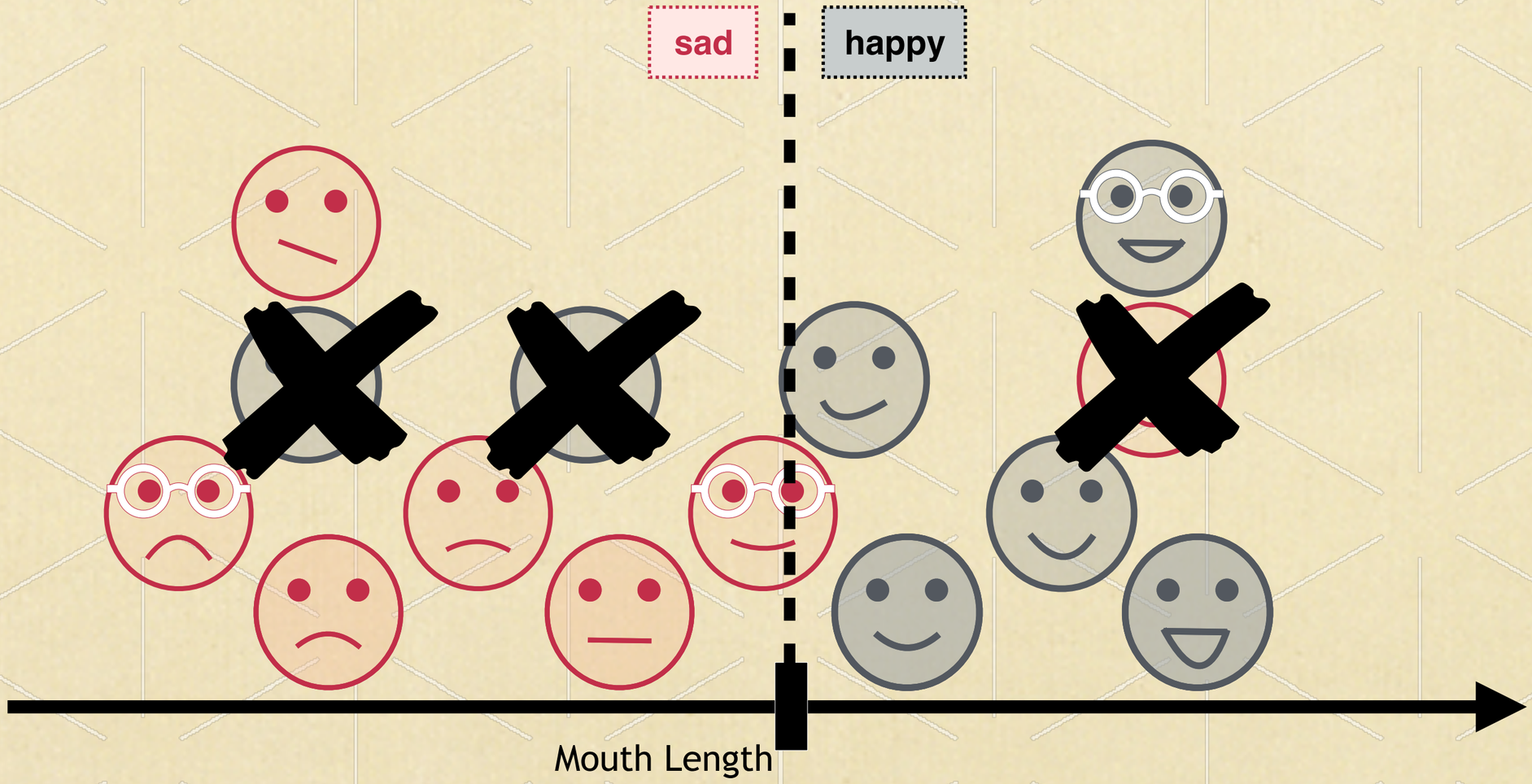
1-FEATURE CLASSIFIER



1-FEATURE CLASSIFIER

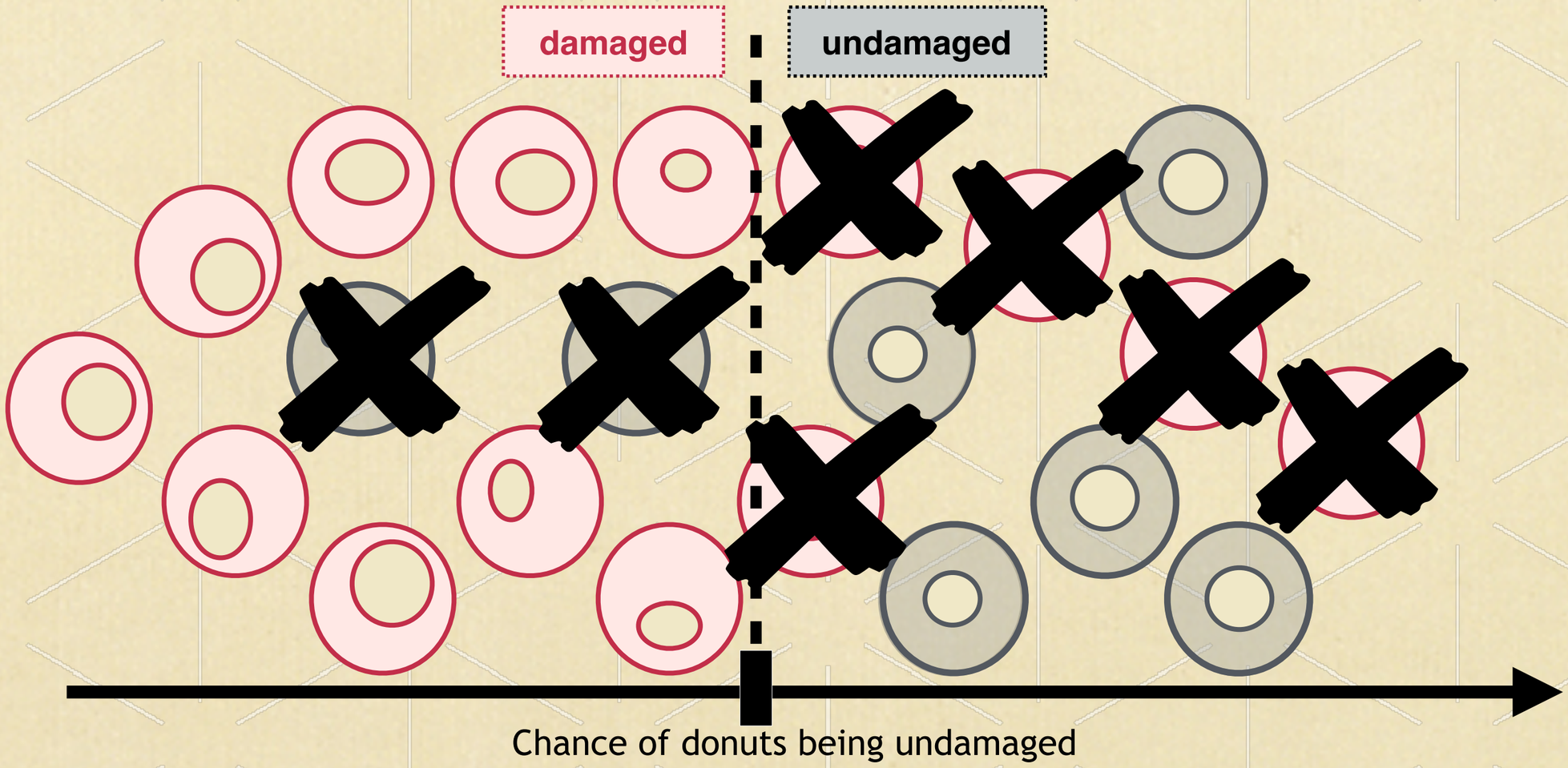


1-FEATURE CLASSIFIER



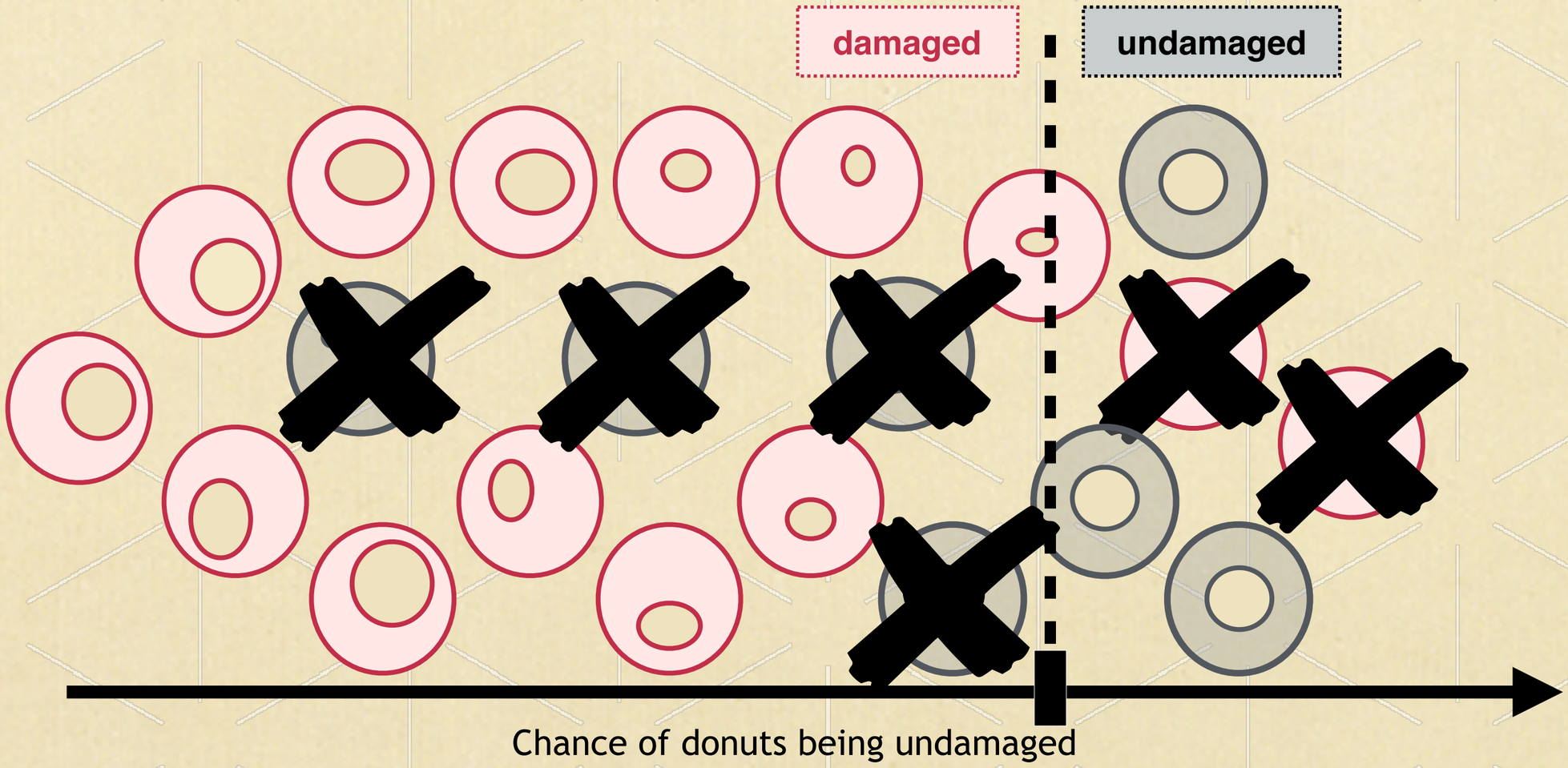
1-FEATURE CLASSIFIER

Tuning **decisions** depend on the use case, and on stakeholder requirements.



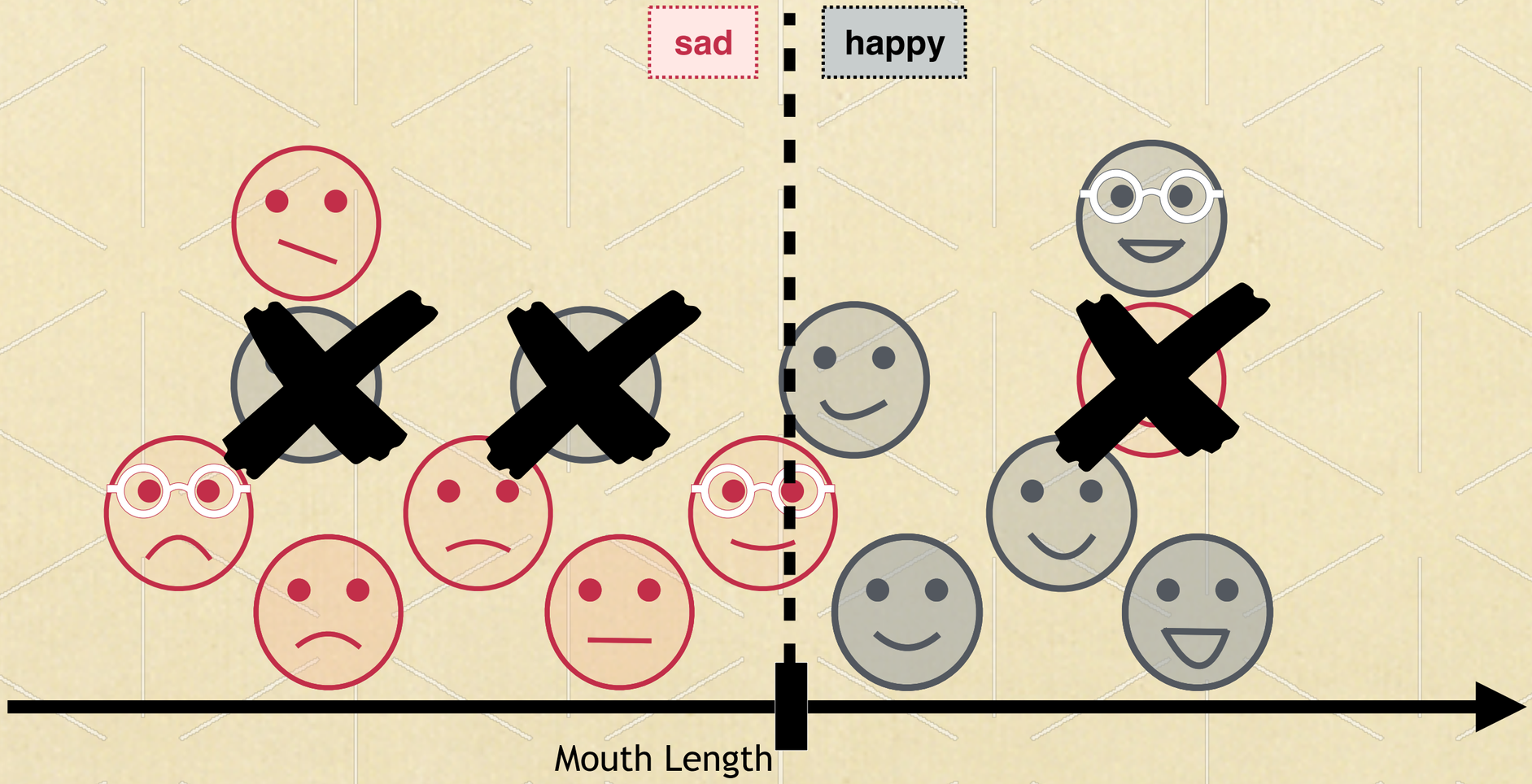
1-FEATURE CLASSIFIER

Tuning **decisions** depend on the use case, and on stakeholder requirements.



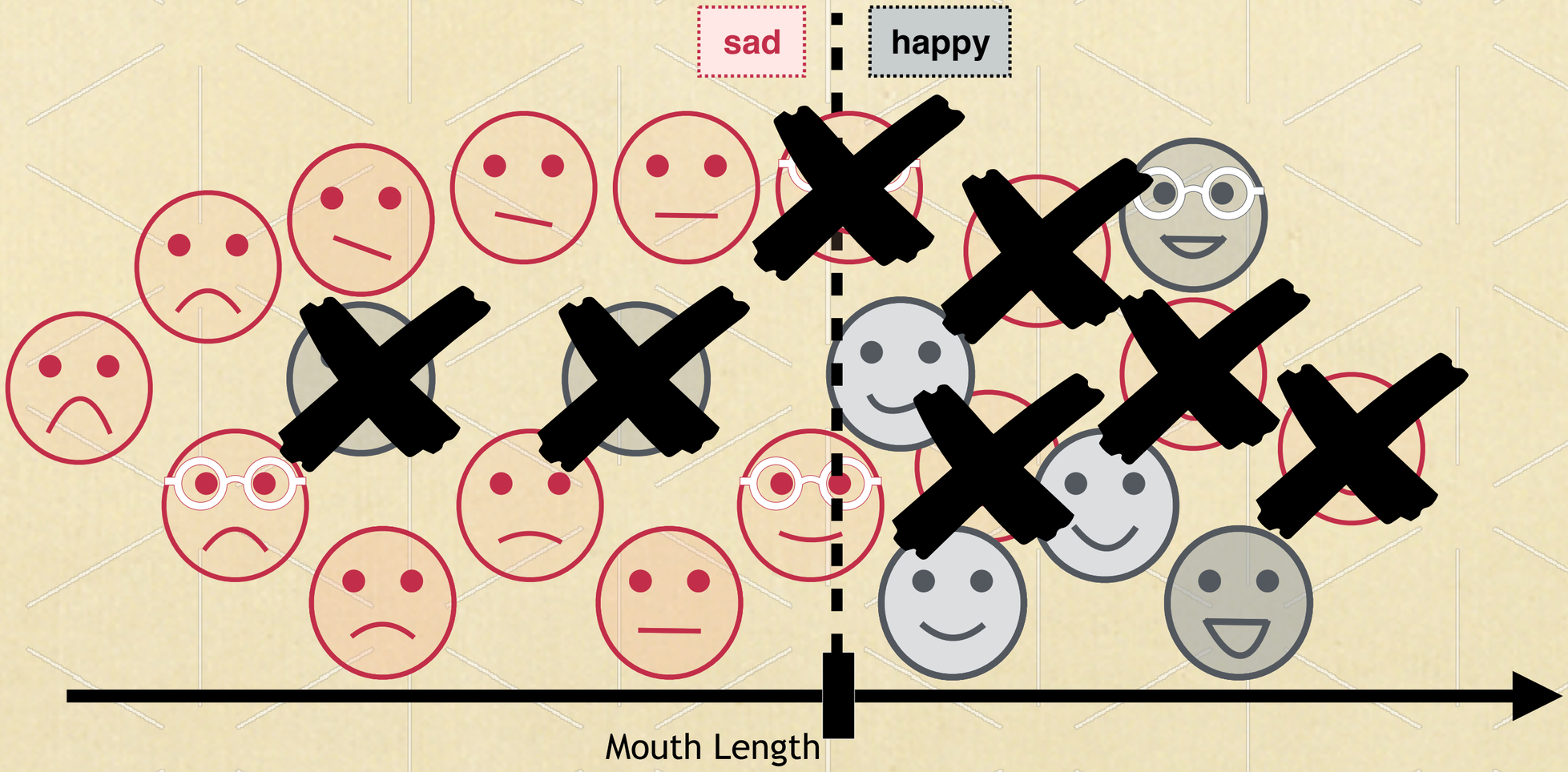
1-FEATURE CLASSIFIER

Class proportions may vary in real life.



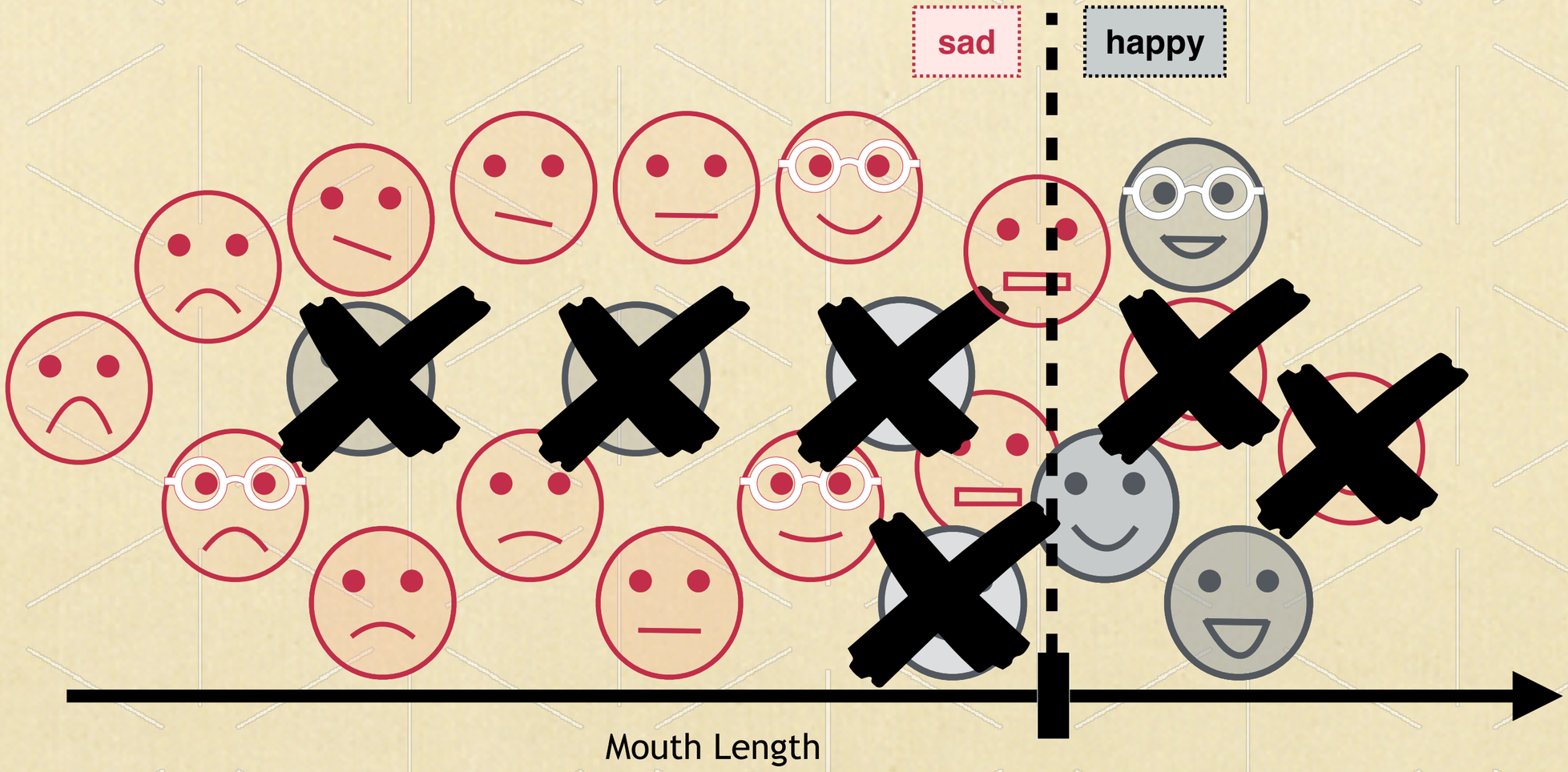
1-FEATURE CLASSIFIER

Class proportions may vary in real life.



1-FEATURE CLASSIFIER

Class proportions may vary in real life.



VARIANCE OF ERROR ESTIMATES

The test set is only a sample, used to estimate the errors in entire batches.

VARIANCE OF ERROR ESTIMATES

In practice, **the training set is fixed**.

Test set

"How many errors for this test set?"

Should be a random sample.

Training set

"Let's run the AI and count them."

May be a non-random sample.

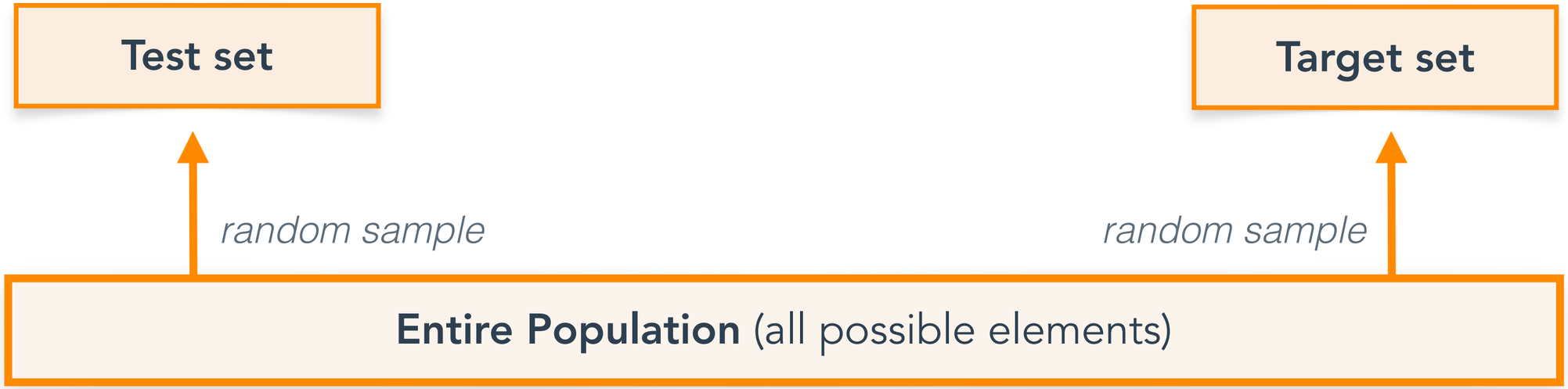
Target set

"So how many errors in this other set?"

May be a non-random sample.

VARIANCE OF ERROR ESTIMATES

- **Smaller samples** give estimates with **higher variance**.



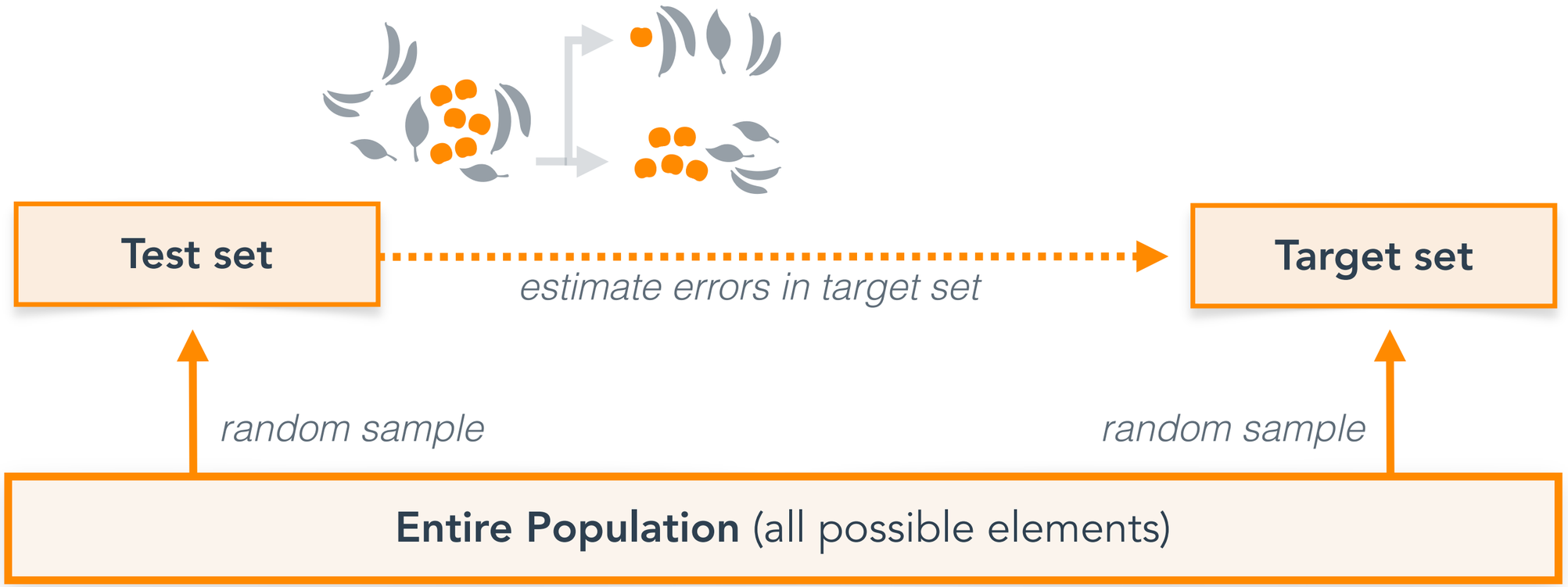
VARIANCE OF ERROR ESTIMATES

- **Smaller test or target sets** give estimates with **higher variance**.

[4] Beauxis-Aussalet & Hardman, Extended Methods to Handle Classification Bias (2017).

VARIANCE OF ERROR ESTIMATES

- **Class distribution** may vary… randomly or not.

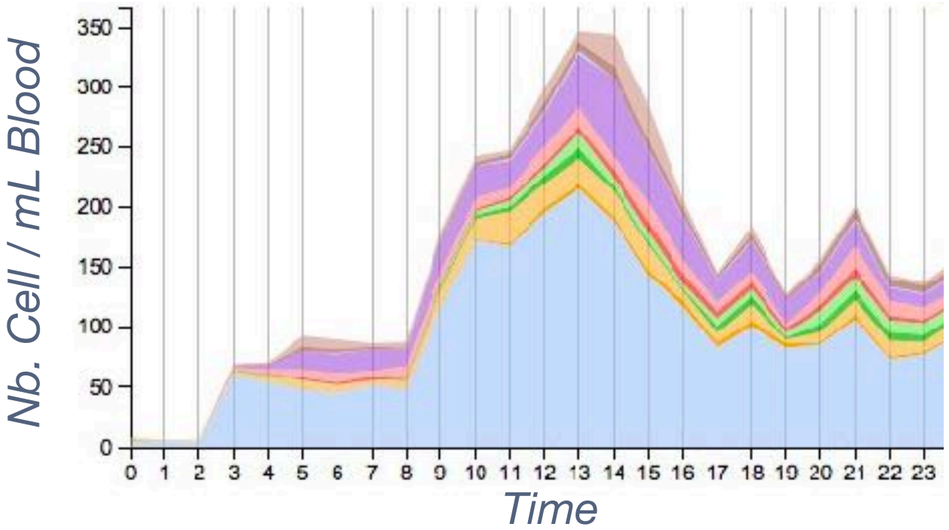


VARIANCE OF ERROR ESTIMATES

- **Class distribution** may vary… randomly or not.

NOW IN REAL LIFE….

Let’s have the AI check if you’re sick…

How many errors in the end-results?

Types of cells over time