Machine Learning

LIN 313 Language and Computers

UT Austin Fall 2025

Administrivia (Monday, September 22)

- HW 1 Graded
- Corpus text due this Wednesday
- HW 2 due Monday
- 10/3 Discussion post
- AI use on the homework

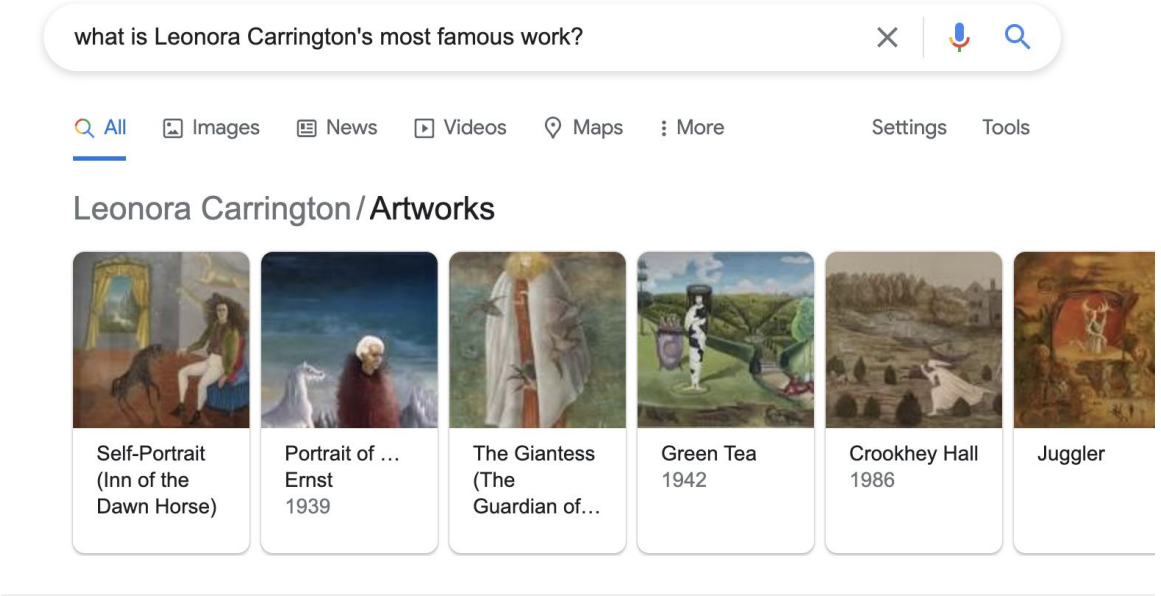
Objectives

- Types of machine learning
  - supervised & unsupervised
- Steps in building a classifier
- Categorization engines

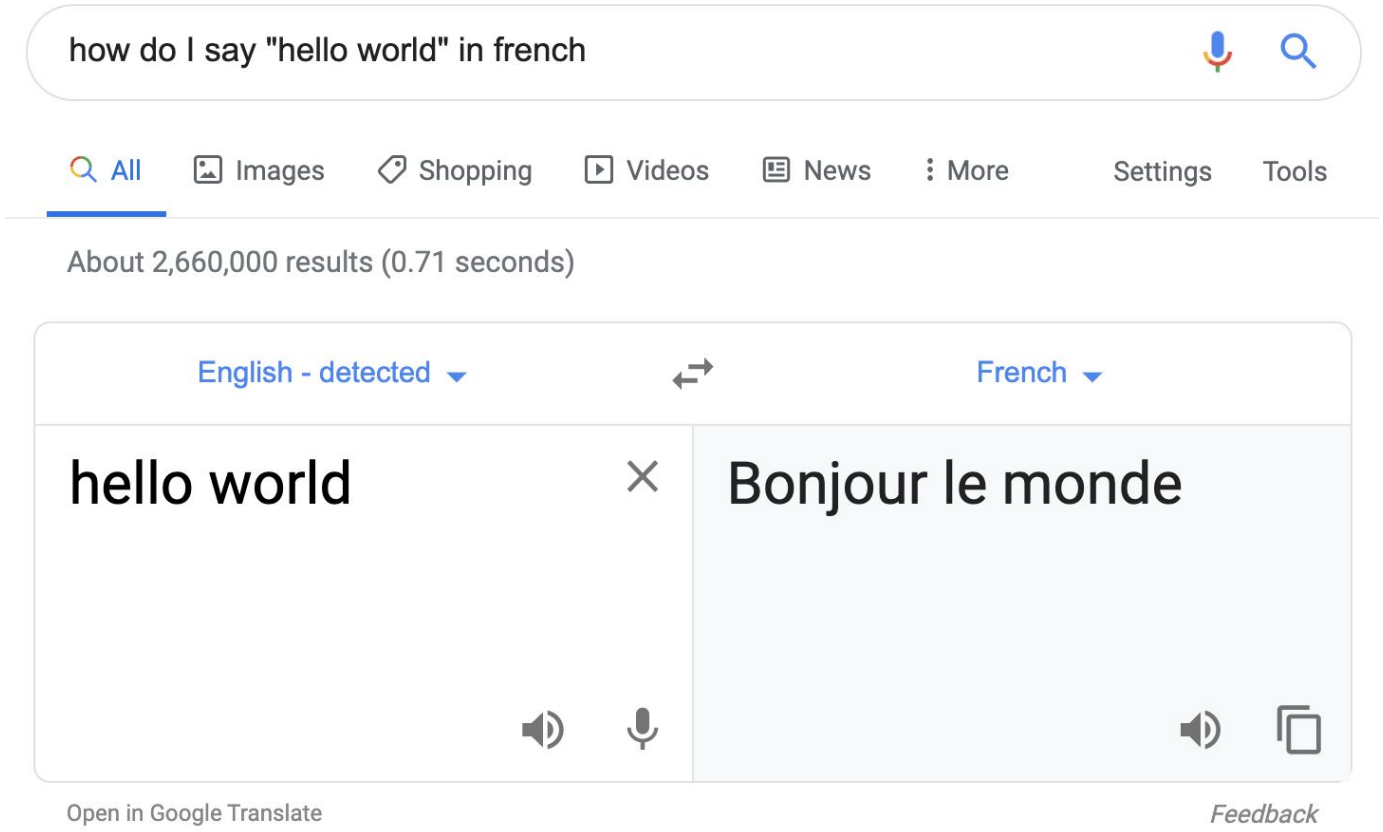
What are some (other) examples of ML?

The "Algorithms" (a.k.a recommendation systems)

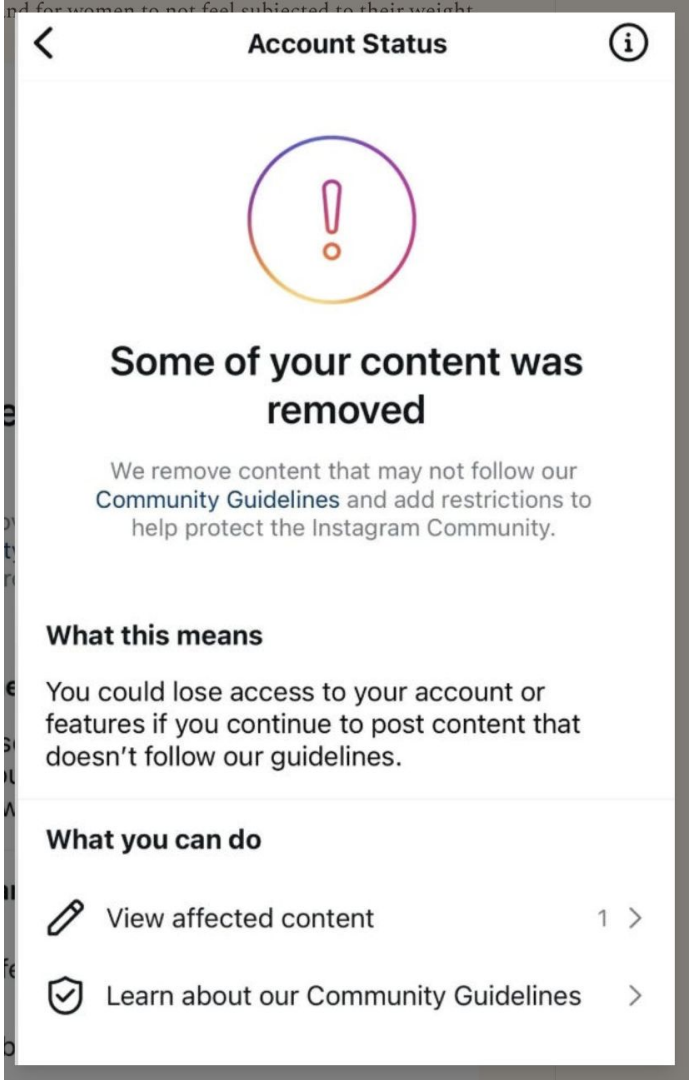
Search


what is Leonora Carrington's most famous work?
All
Images
News
Videos
Maps
More
Settings
Tools
Leonora Carrington / Artworks
Self-Portrait (Inn of the Dawn Horse)
Portrait of … Ernst
1939
The Giantess (The Guardian of…
Green Tea
1942
Crookhey Hall
1986
Juggler
https://www.theartstory.org › ... › Leonora Carrington
Leonora Carrington Artworks & Famous Paintings | TheArtStory
Jan 25, 2015 — Progression of Art · The Meal of Lord Candlestick · Portrait of Max Ernst · Self-Portrait · The Giantess (The Guardian of the Egg) · Ulu's Pants · Bird ...

Google Translate

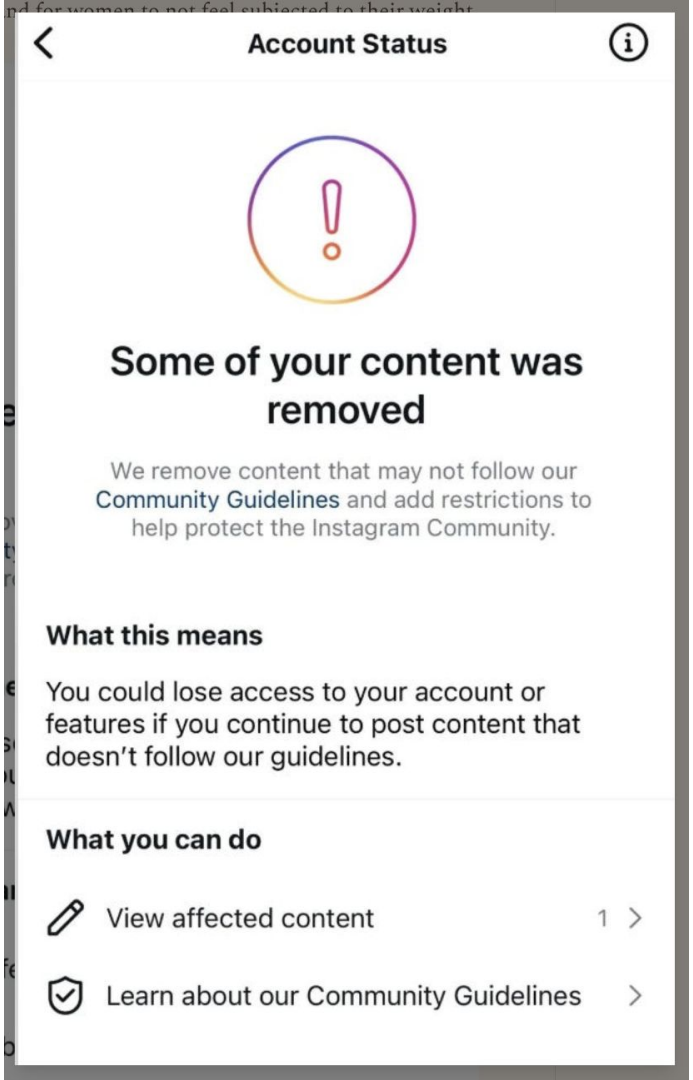


Content Moderation



Content Moderation

- the post in question was about body positivity
- it was flagged for 'encouraging self harm'
- https://idaliasalsamendi.substack.com/p/is-instagram-shadowbanning-you



How do they work?

Natural Language Processing (NLP) Systems

- INPUT: Takes in text
- OUTPUT: Spits out
  - Sentiment labels
  - Named entities (Pope Francis, Assata Shakur, Apple (sometimes))
  - Topics (sports, national legislation, country music, black and white photography)
  - Geo-coordinates
  - Syntactic Structures
  - Translations

What happens in between?

What is Machine Learning?

“It gives the computer the ability to learn without being explicitly programmed” (Arthur Samuel)

“A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E. “ (Tom Mitchell)

(We will need to unpack this)

How do Humans Learn

from Jenny Saffran's lab page, https://infantlearning.waisman.wisc.edu/research/ :

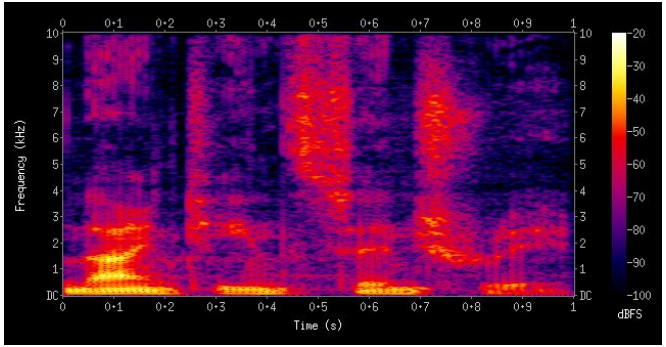
Imagine you are faced with the following challenge:

You must discover the structure of an immense system which contains tens of thousands of pieces, all generated from a small set of materials. These pieces, in turn, can be combined in an infinite number of ways. Only a subset of those infinite combinations is actually correct. However, just to make things even more difficult, this subset is itself infinite. Somehow you must rapidly converge on the internal structure of this system so that you can use it to communicate.

Oh, and you are a very young child.

This system, of course, is **human language**. Given its richness and complexity, it seems improbable that children could ever discern its structure. Nevertheless, they do, almost without exception. The process of acquiring such a system is unlikely to be any less complex than the system itself.

Unsupervised Learning in People: learning word boundaries



Statistical learning in language acquisition:

Imagine you are an 8-month-old infant, hearing a continuous stream of language coming at you. Before you can learn words, you first have to learn word boundaries.

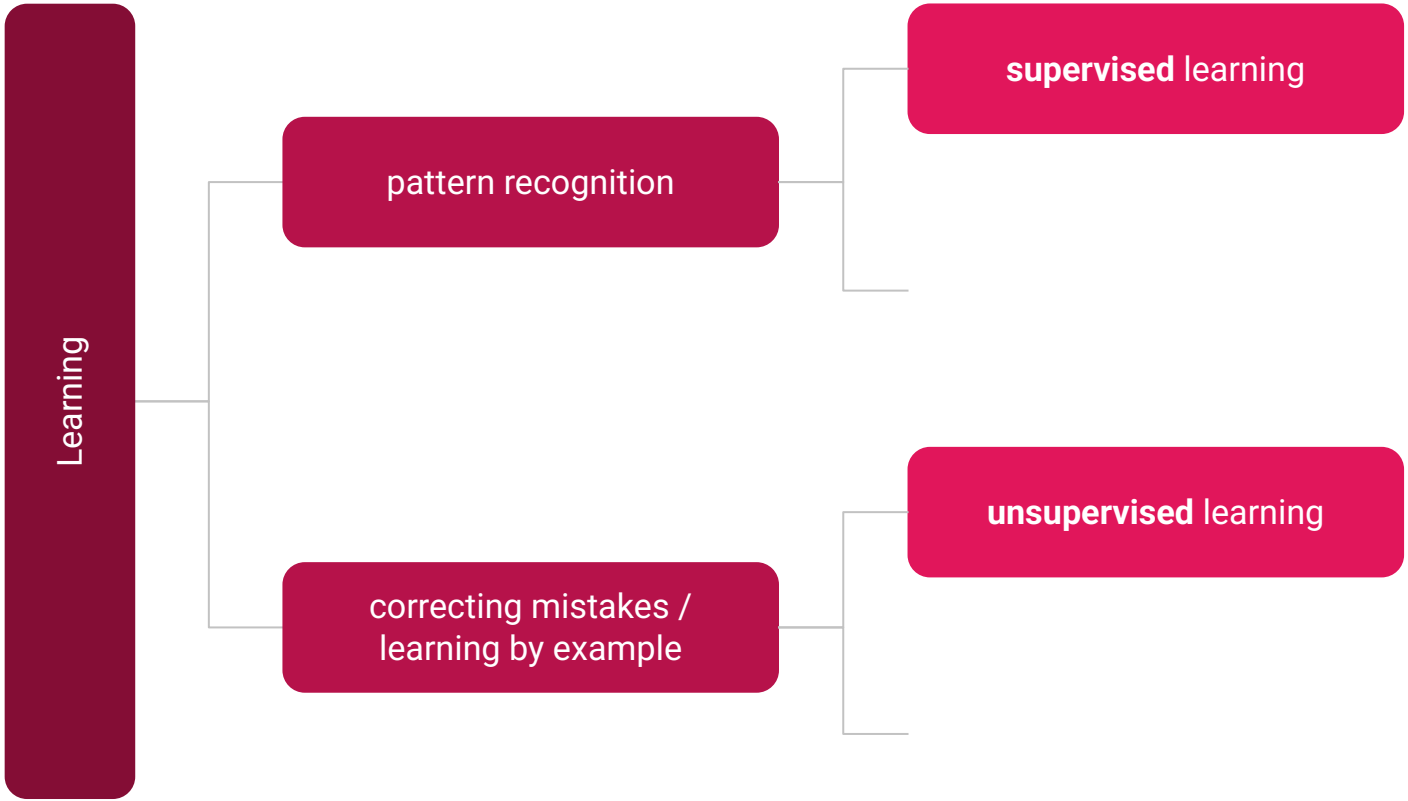
Here is a spectrogram of a male speaker saying “nineteenth century”: You cannot see where one word ends and the next one begins

The main idea on **statistical learning**: In a phrase like “pretty baby”,

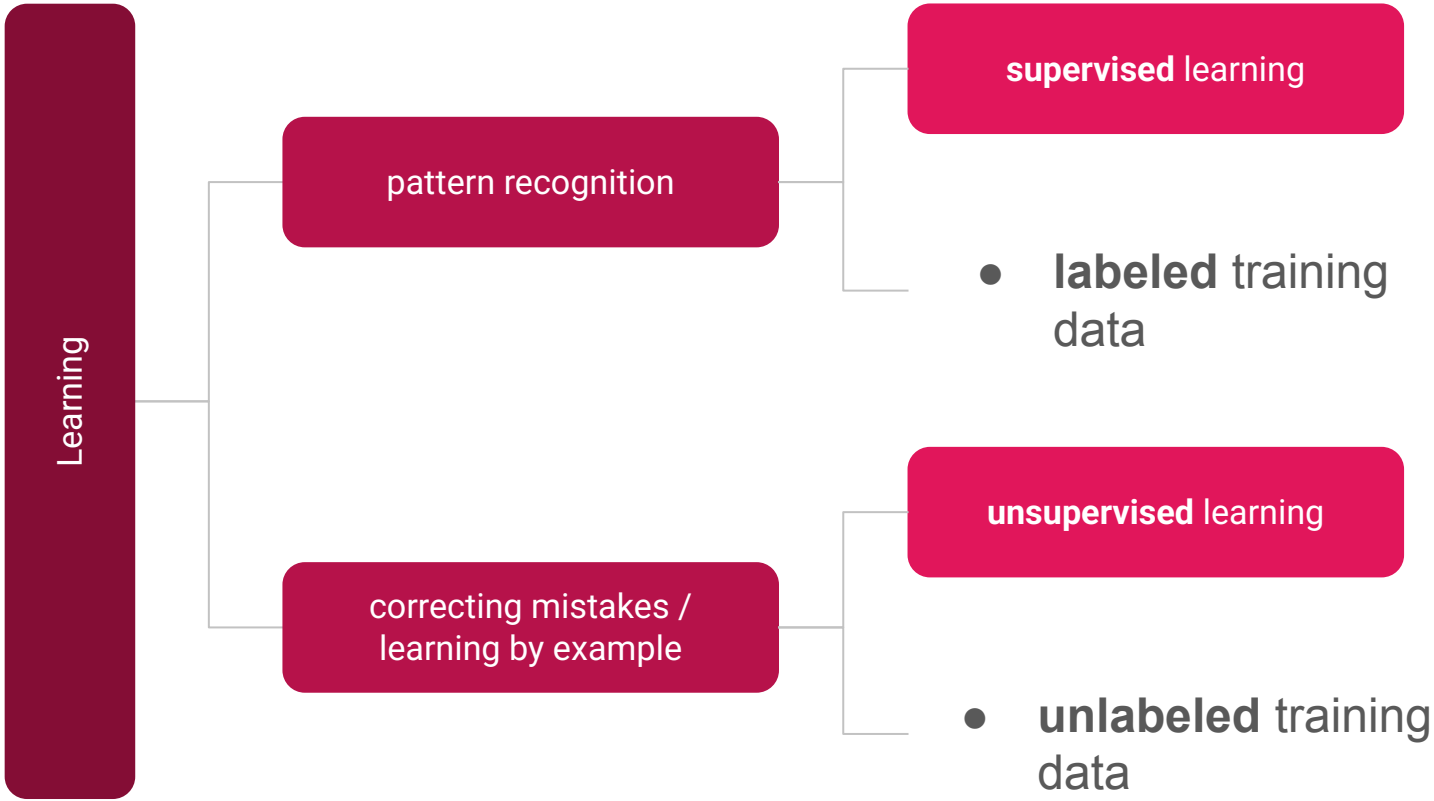
you are more likely to hear “pre” followed by “ti” than you are to hear “ti” followed by “ba”

Experiments by Jenny Saffron and Elissa Newport: Adults as well as infants can learn these regularities. This has been confirmed with experiments using pseudo-words.

What is Learning?



What is Learning?



Two Ways to Learn from Data

Supervised Learning

Learns from data with answers.

Unsupervised Learning

Finds hidden patterns in data without answers.

https://www.geeksforgeeks.org/machine-learning/supervised-unsupervised-learning/#

Learning with Labels

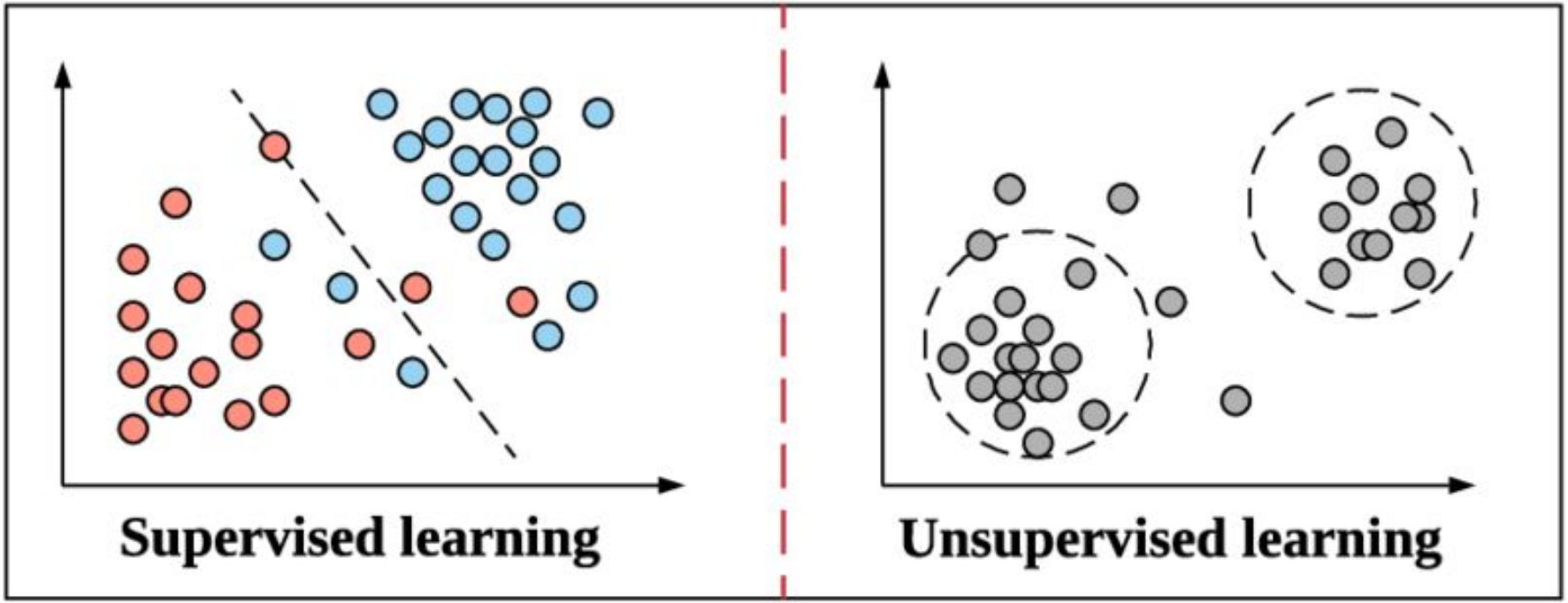
Machine learns by studying labeled fruits.

Recognizes new fruits by comparing features to what it learned.

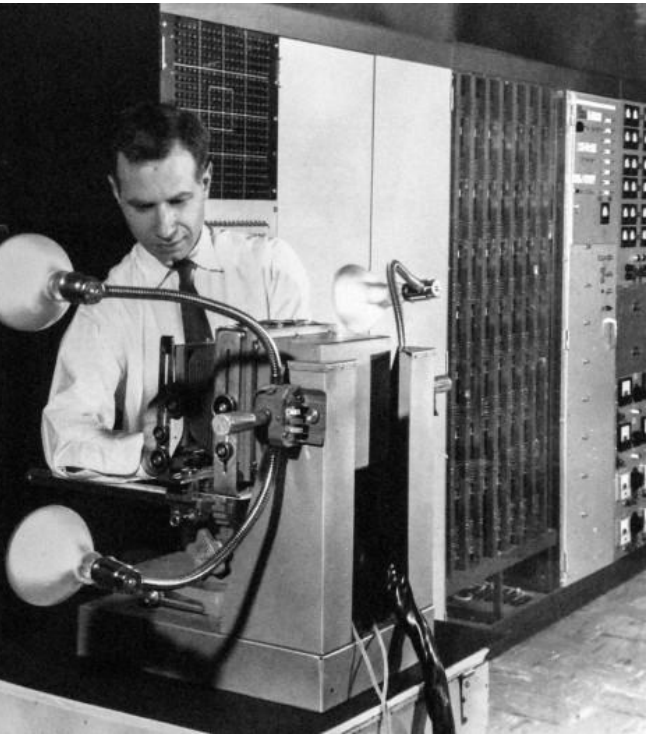
Learning Without Labels

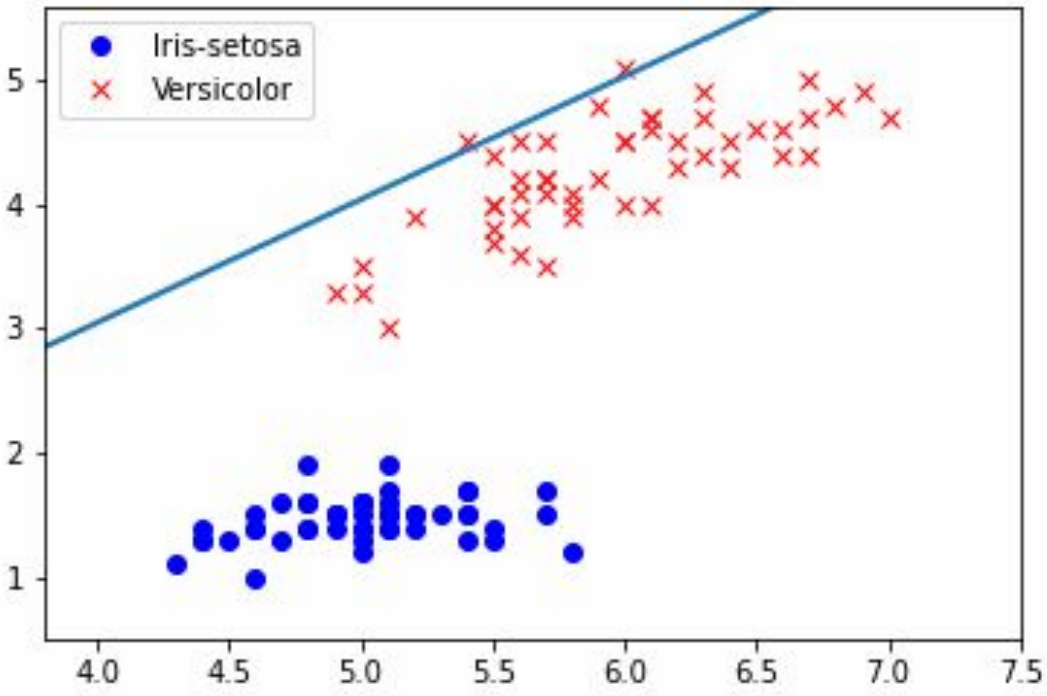
Machine groups unlabeled images by similar features

It discovers patterns without knowing the exact categories.

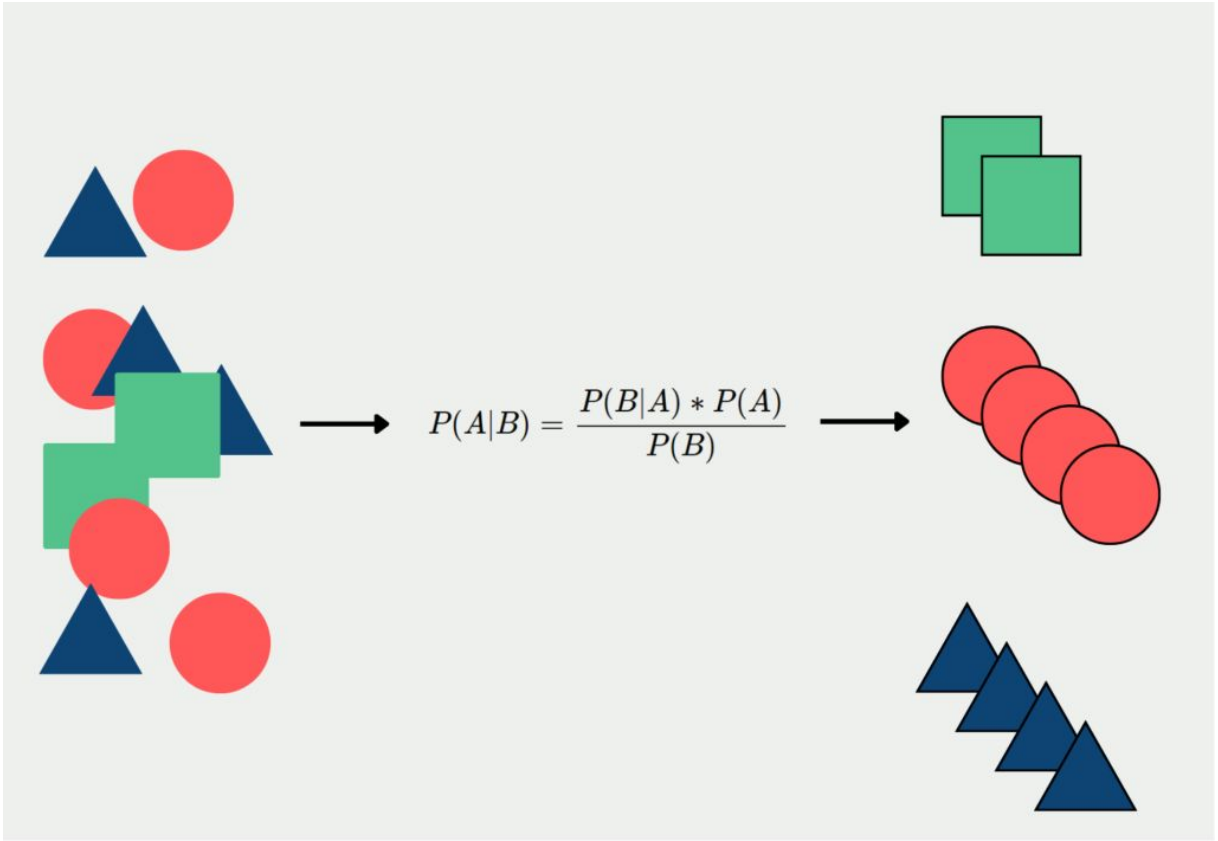

Supervised learning
Unsupervised learning

Supervised Learning: The Perceptron



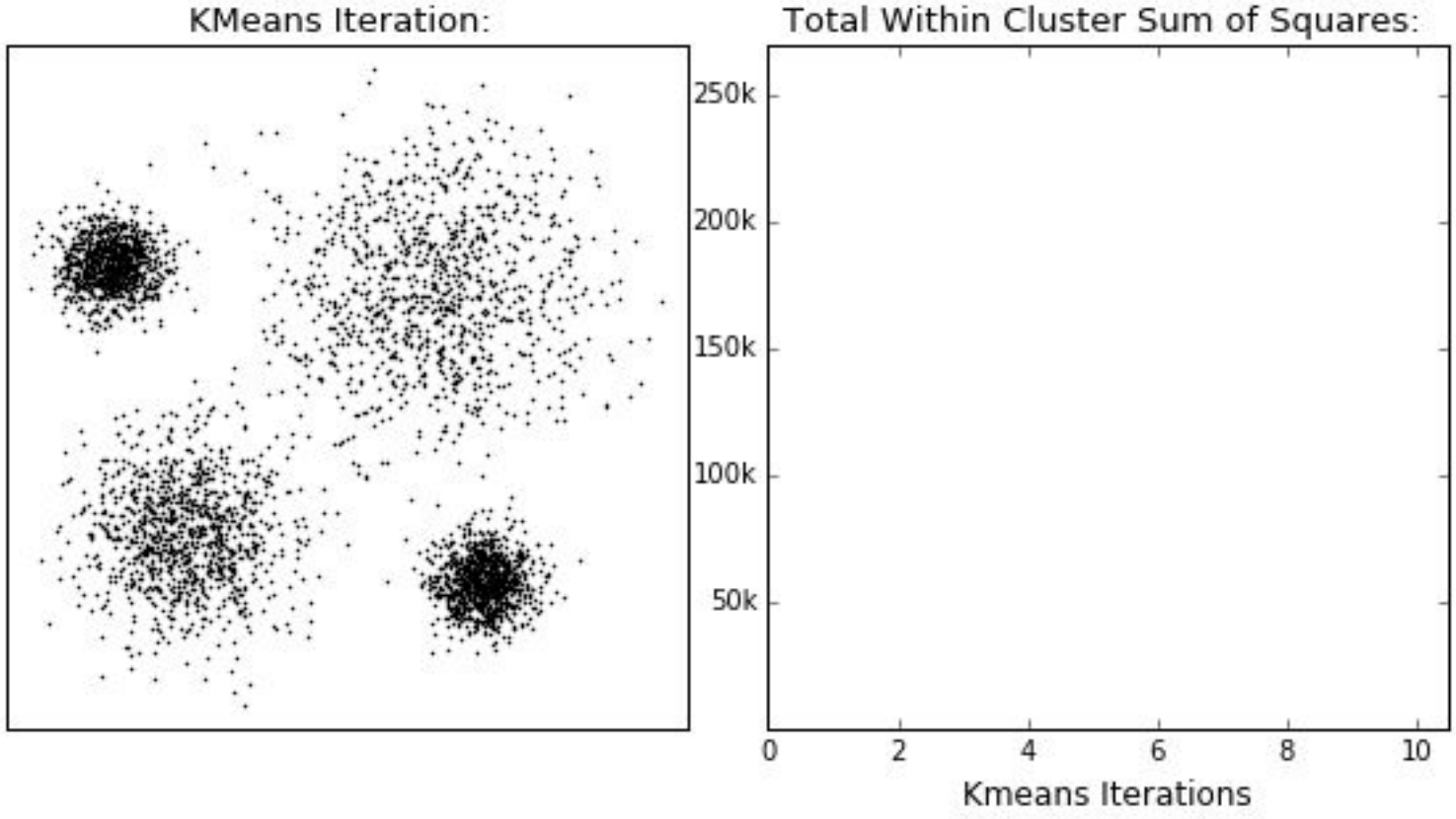


Supervised Learning: Naive Bayes

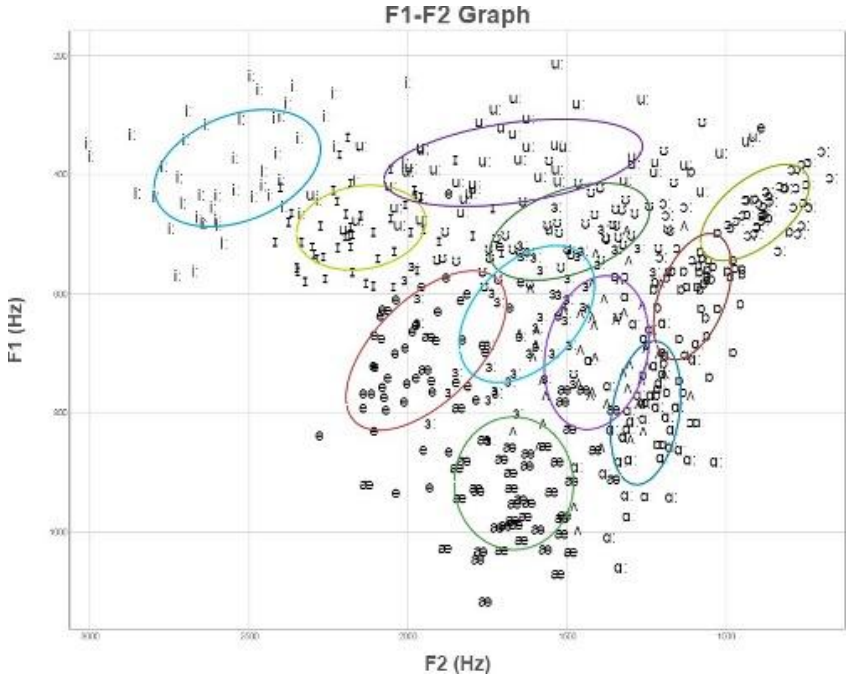


Unsupervised Learning: K-means clustering

(and a

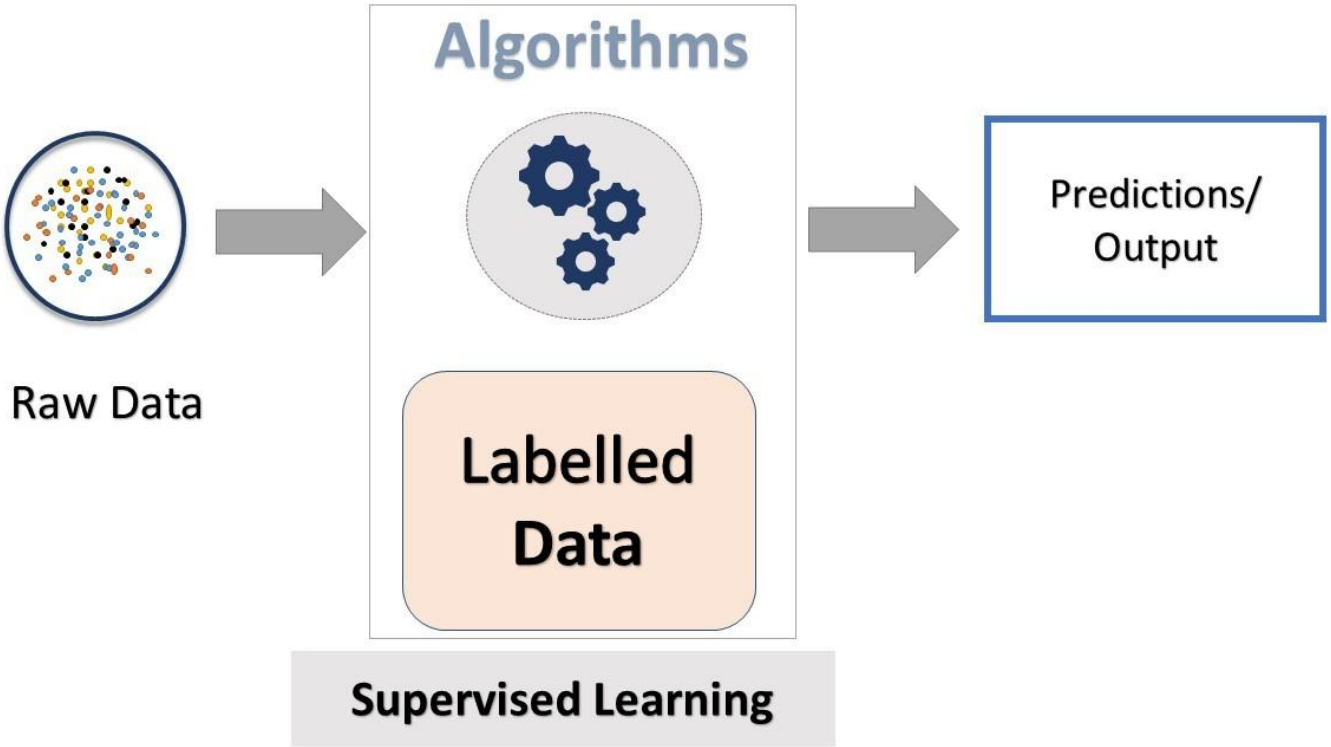


Unsupervised Learning: Clustering in Experimental phonetics



https://www.phon.ucl.ac.uk/courses/spsci/expphon/week5.php

Zooming in: Supervised Learning



Supervised Classification: Training data

In **supervised learning**, the learner observes data with **labels** (aka “**gold labels**” or "**ground truth labels**") and generalizes over the data to learn how labels apply to datapoints.

represent a data point as an **ordered pair (x, y)**

UserName	ScreenName	Location	TweetAt	OriginalTweet	Sentiment
3799	48751	London	16-03-2020	@MeNyrbie @Phil_Gahan @Chrisitv https://t.co/i...	Neutral
3800	48752	UK	16-03-2020	advice Talk to your neighbours family to excha...	Positive
3801	48753	Vagabonds	16-03-2020	Coronavirus Australia: Woolworths to give elde...	Positive
3802	48754	NaN	16-03-2020	My food stock is not the only one which is emp...	Positive
3803	48755	NaN	16-03-2020	Me, ready to go at supermarket during the #COV...	Extremely Negative

Supervised Classification: Training data

In **supervised learning**, the learner observes data with **labels** (aka “**gold labels**” or "**ground truth labels**") and generalizes over the data to learn how labels apply to datapoints.

represent a data point as an **ordered pair (x, y)**

x: data point

y: Gold Label

	UserName	ScreenName	Location	TweetAt	OriginalTweet	Sentiment
3799	48751	London	16-03-2020	@MeNyrbie @Phil_Gahan @Chrisitv https://t.co/i...	Neutral
3800	48752	UK	16-03-2020	advice Talk to your neighbours family to excha...	Positive
3801	48753	Vagabonds	16-03-2020	Coronavirus Australia: Woolworths to give elde...	Positive
3802	48754	NaN	16-03-2020	My food stock is not the only one which is emp...	Positive
3803	48755	NaN	16-03-2020	Me, ready to go at supermarket during the #COV...	Extremely Negative

Supervised Classification: features

How do we represent the (data, label) pairs with numbers??

Supervised Classification: features

How do we represent the (data, label) pairs with numbers??

Output is easy: 0 (false) or 1 (true)

Supervised Classification: features

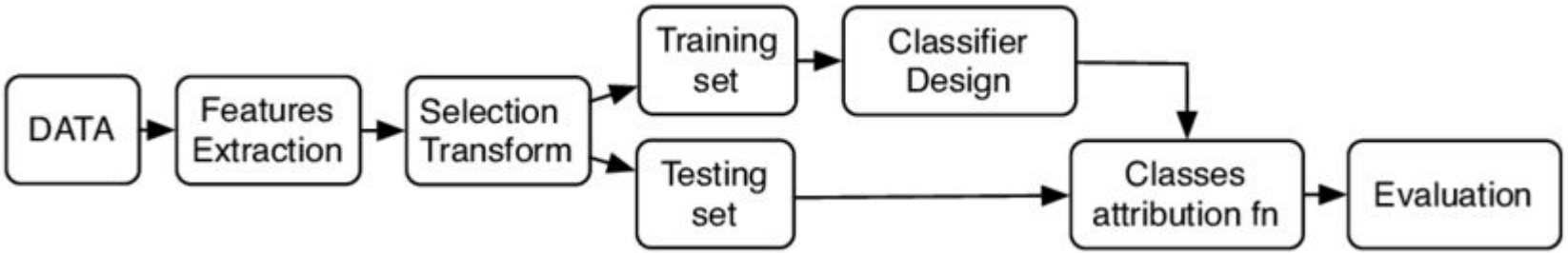
How do we represent the (data, label) pairs with numbers??

Output is easy: 0 (false) or 1 (true)

Input is harder: "My foodstock is not the only one which is . . . "

- we represent a text as a **feature vector** : a list of values for different features

Supervised Classifiaction: Test data



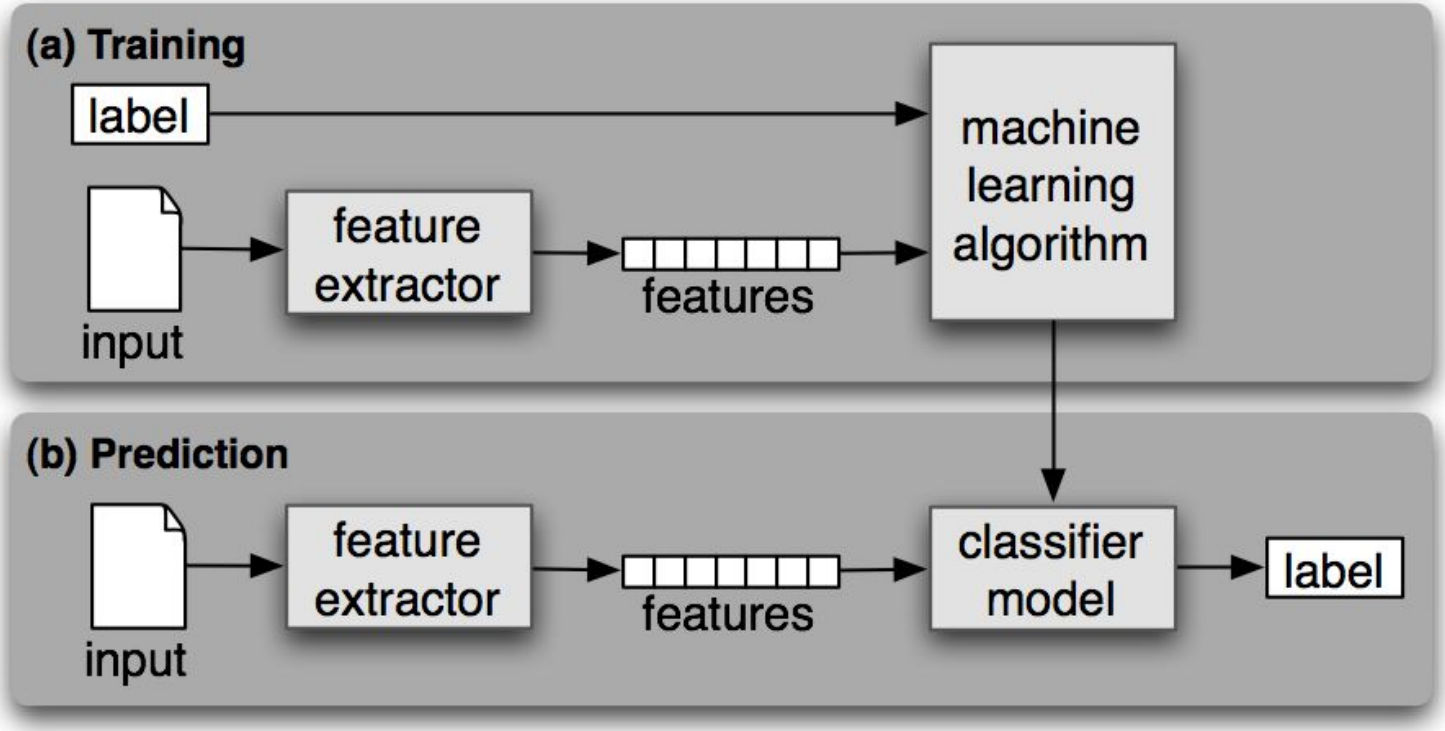
We want to be able to generalize to **unseen future examples**

So we **hold out** some of our data during training

That is, we don't let the model see it during the training process

After training, we use this data to evaluate the performance of the classifier

Putting it Together: Supervised Classification

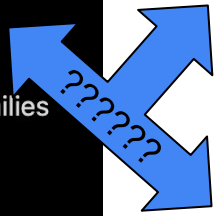


Supervised or Unsupervised?

Supervised or Unsupervised?

NEG

POS



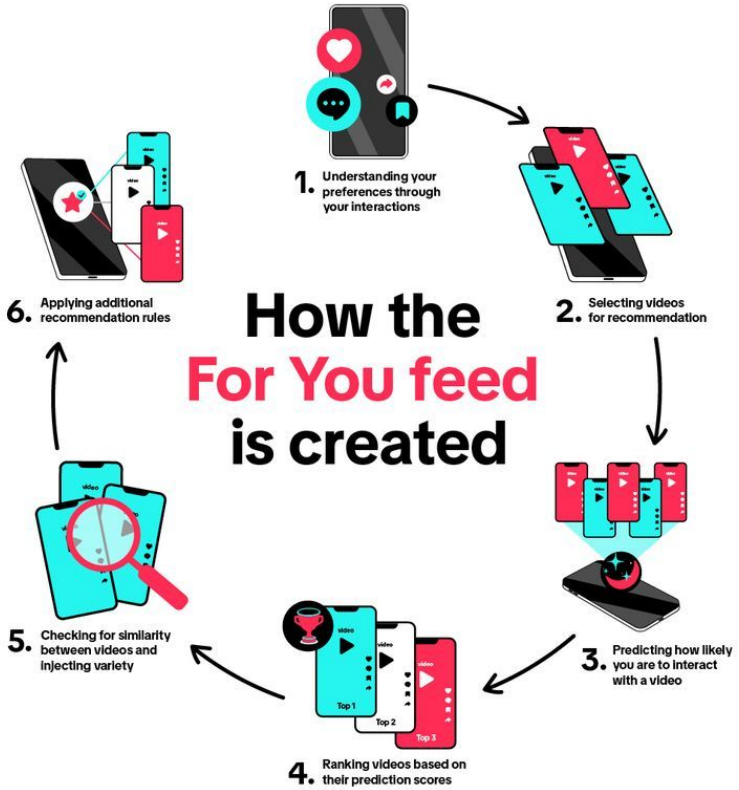
Our Naive Bayes vegemite classifier

Supervised or Unsupervised?

- what is being categorized?
- what are the categories?
- are the categories known ahead of time?

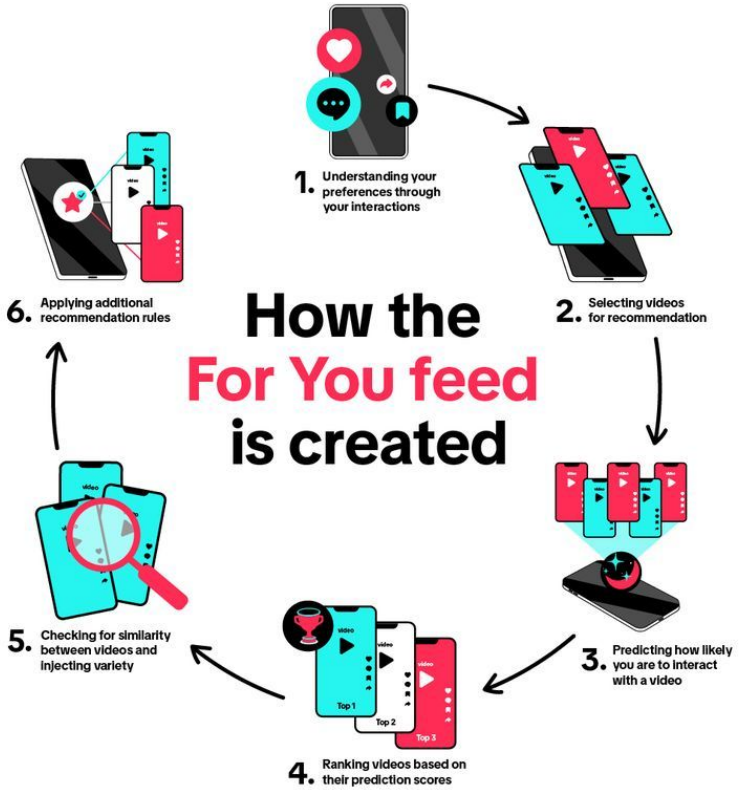
Unsupervised

- "the algorithm" categorizes YOU!
- Categories of people (users) are not known ahead of time.
- You are clustered with other people with similar patterns of behavior

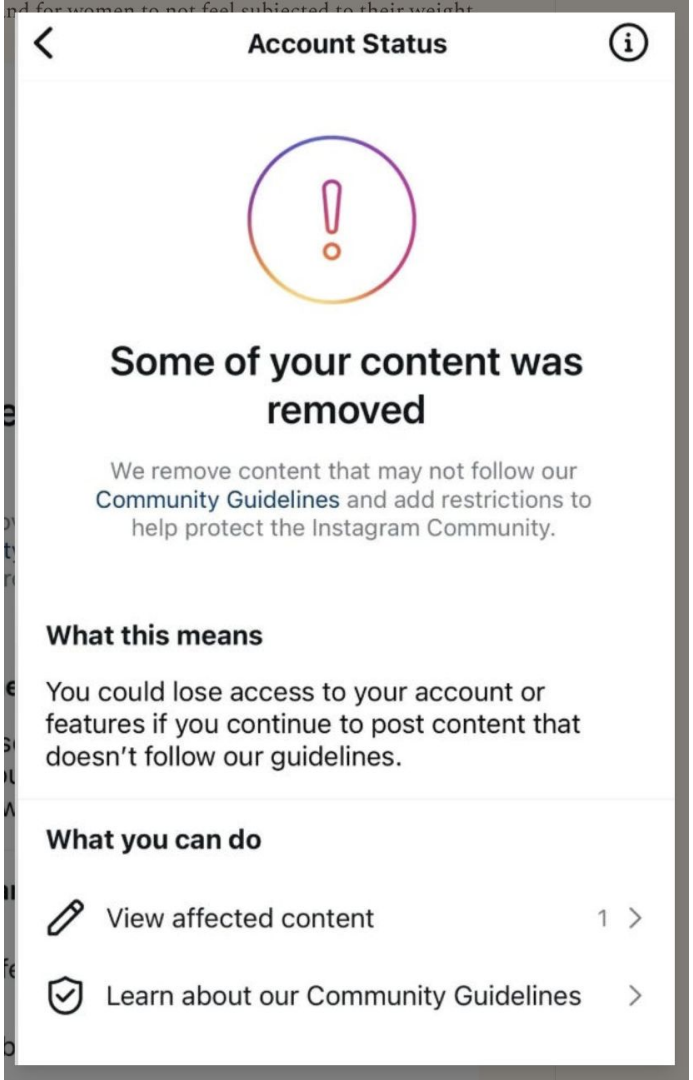


Unsupervised

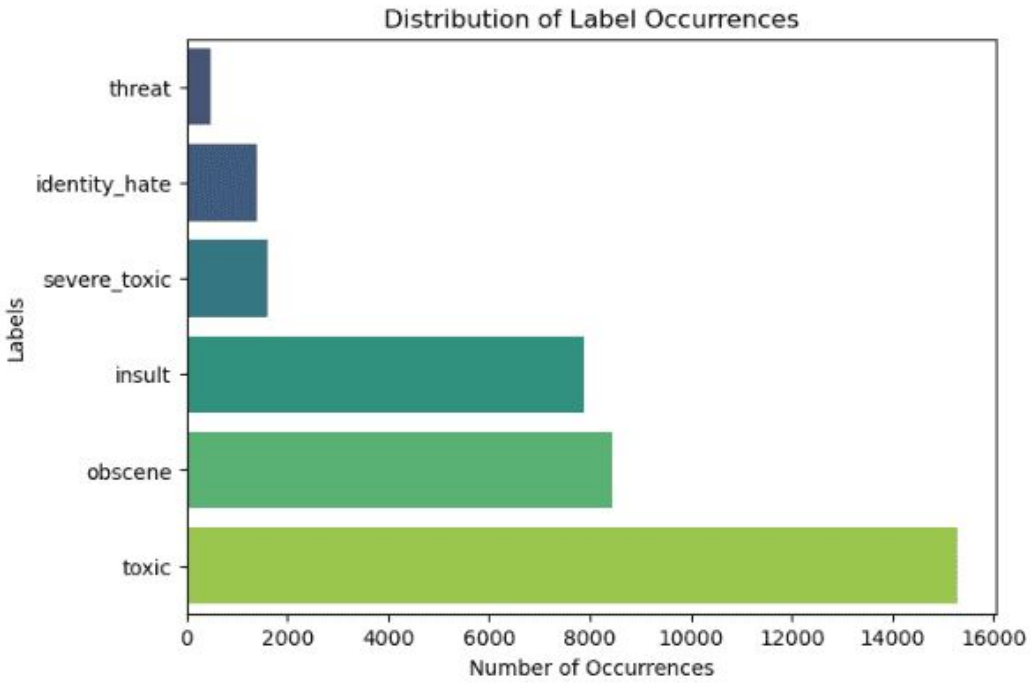
- "the algorithm" categorizes YOU!
- Categories of people (users) are not known ahead of time.
- You are clustered with other people with similar patterns of behavior

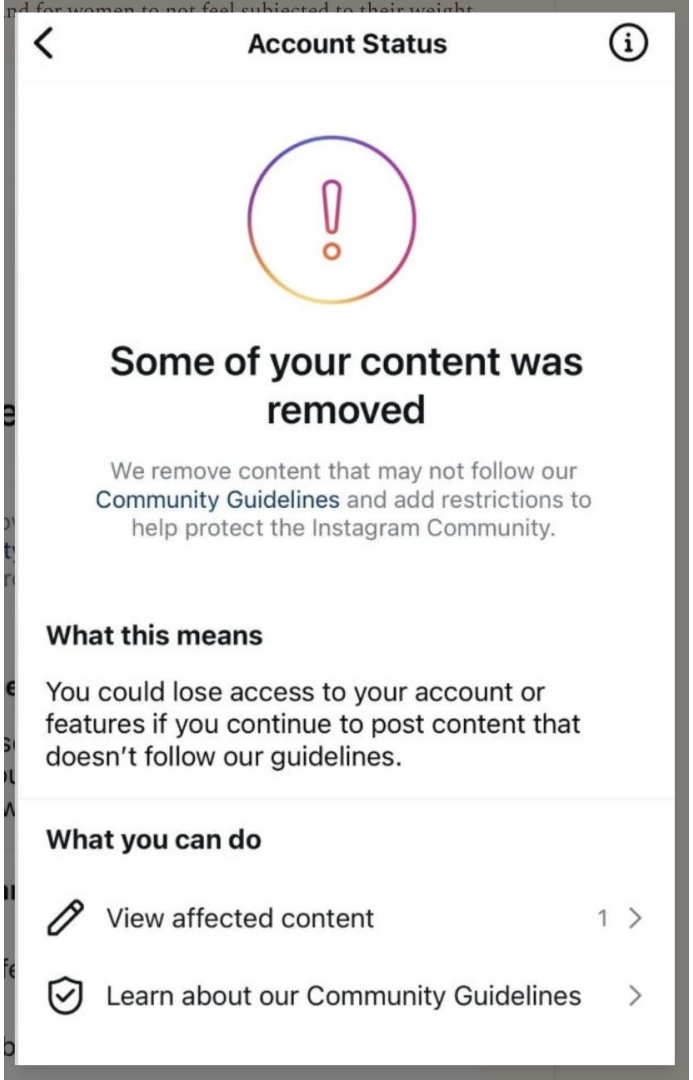


Supervised or Unsupervised?

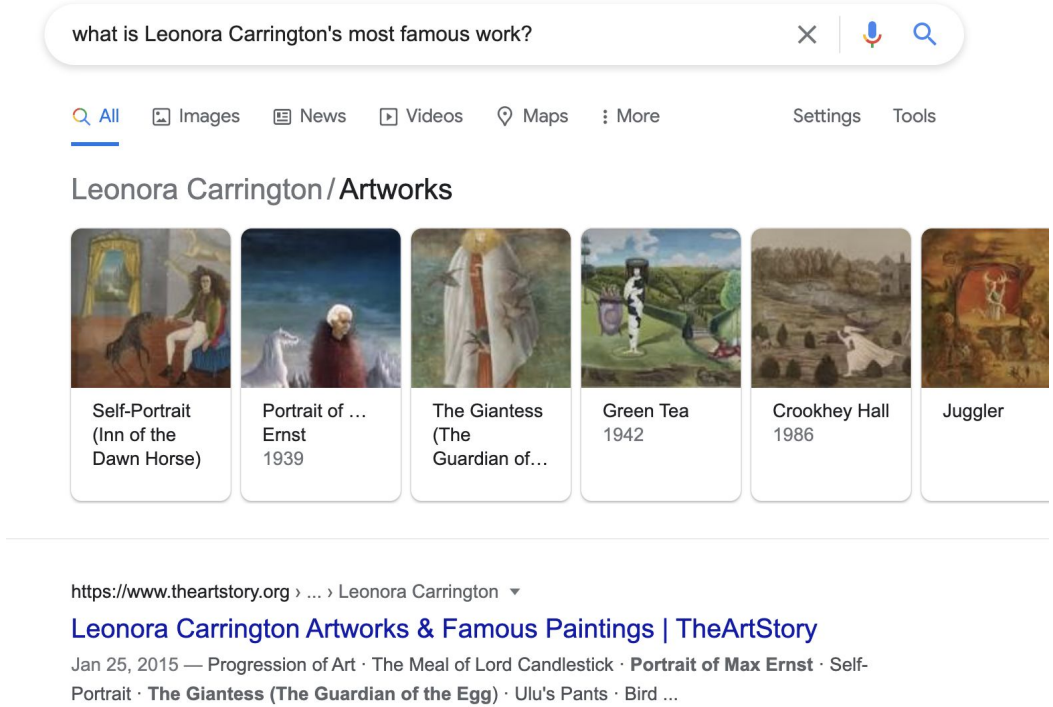


Supervised!


Distribution of Label Occurrences
threat
identity_hate
severe_toxic
insult
obscene
toxic
Labels
0
2000
4000
6000
8000
10000
12000
14000
16000
Number of Occurrences

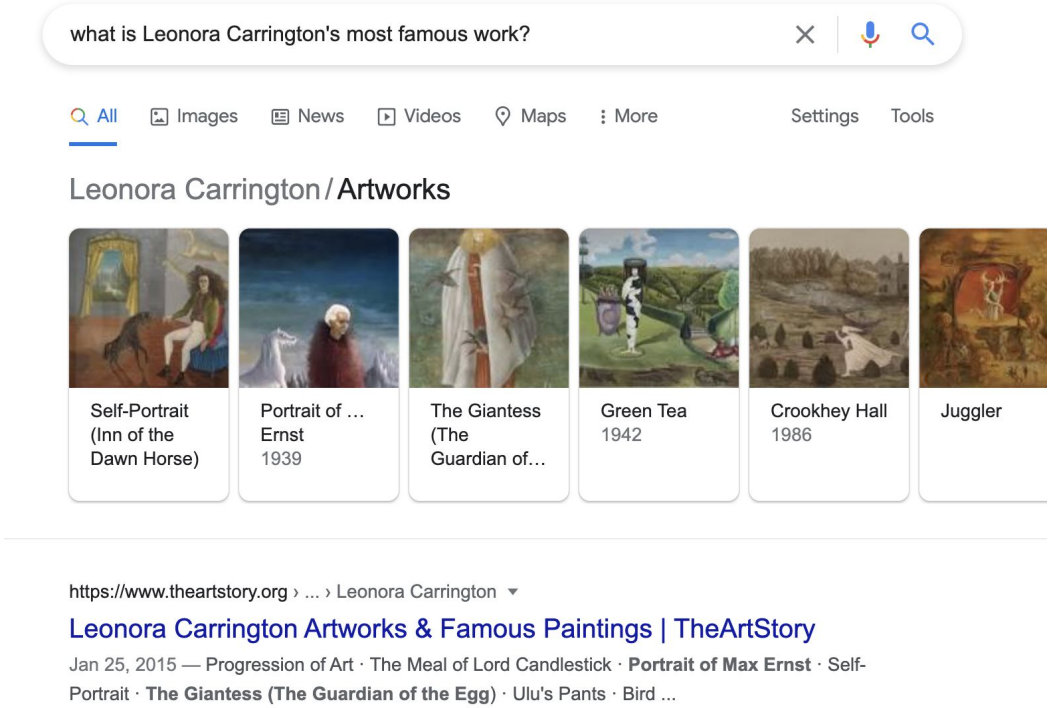

Account Status
Some of your content was removed
We remove content that may not follow our Community Guidelines and add restrictions to help protect the Instagram Community.
What this means
You could lose access to your account or features if you continue to post content that doesn't follow our guidelines.
What you can do
View affected content
1
Learn about our Community Guidelines

Supervised or Unsupervised?



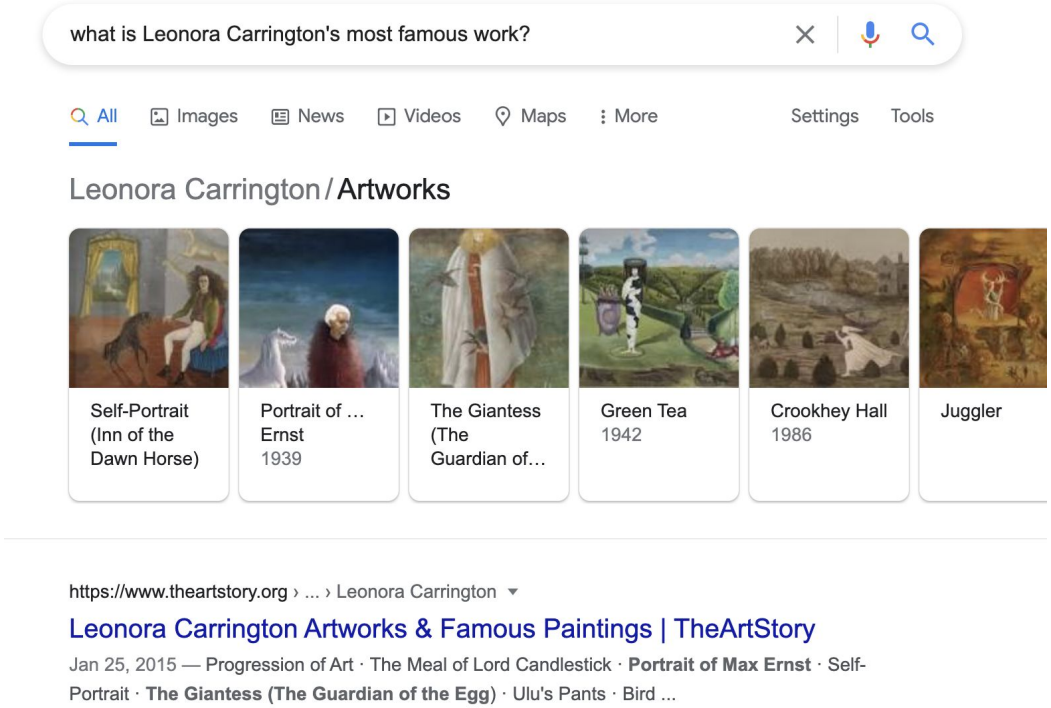
Supervised or Unsupervised?

- What is being categorized?
- What kinds of categories are there?
- Are they known ahead of time?

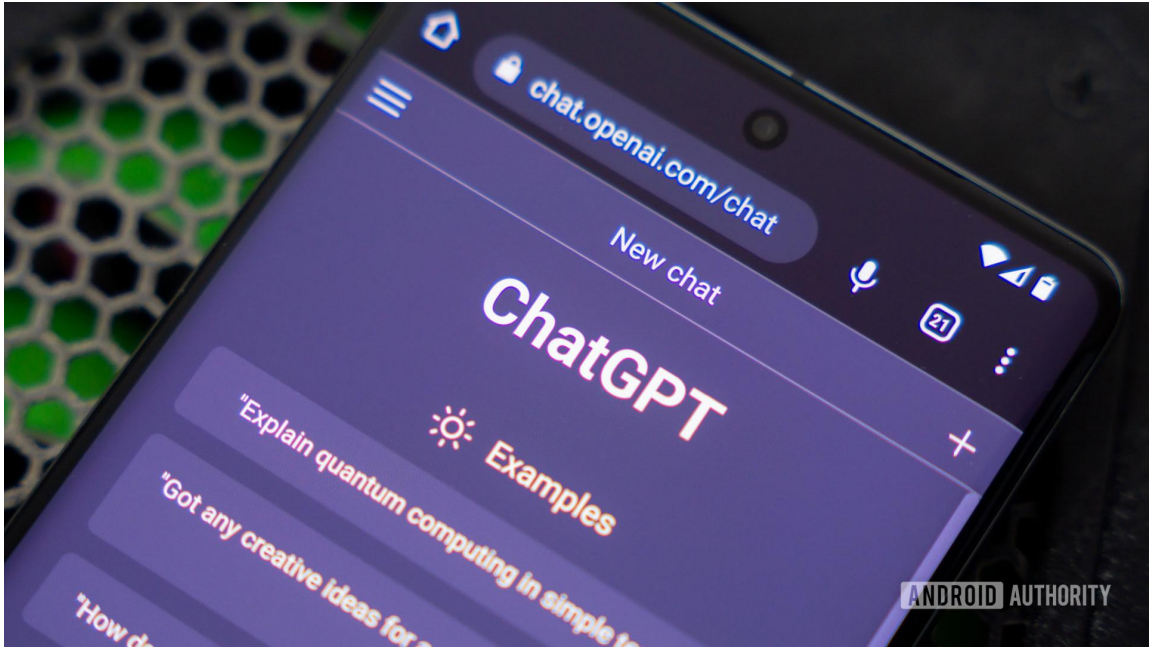


Unsupervised!

- What is being categorized?
- What kinds of categories are there?
- Are they known ahead of time?



Supervised or Unsupervised?

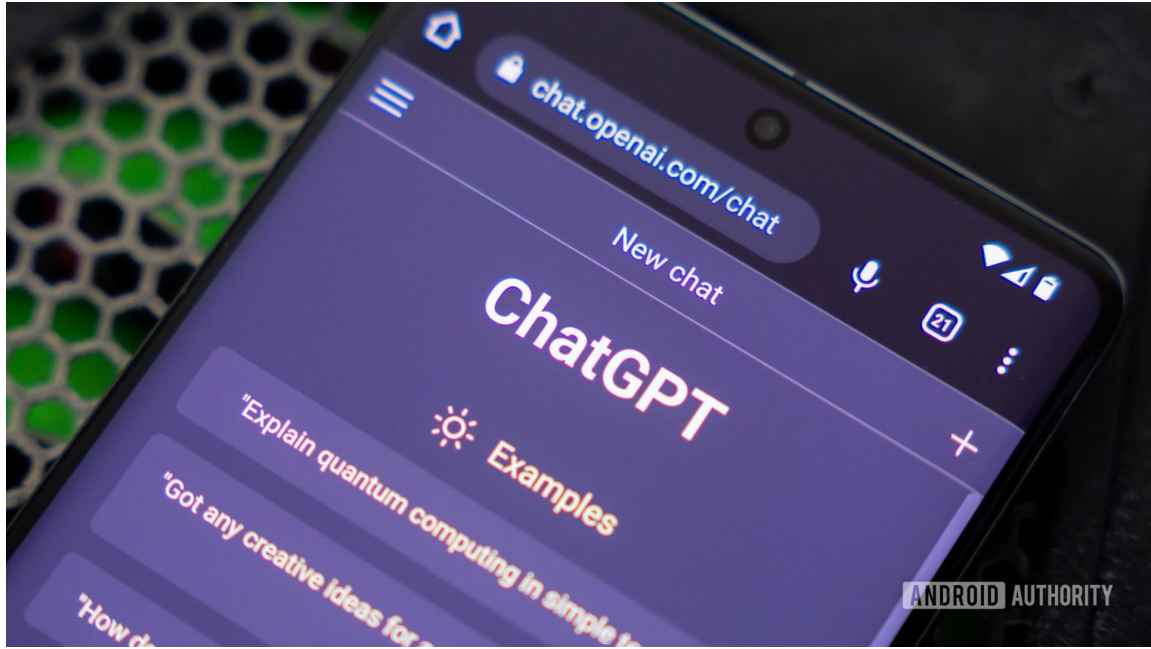


Supervised or Unsupervised?

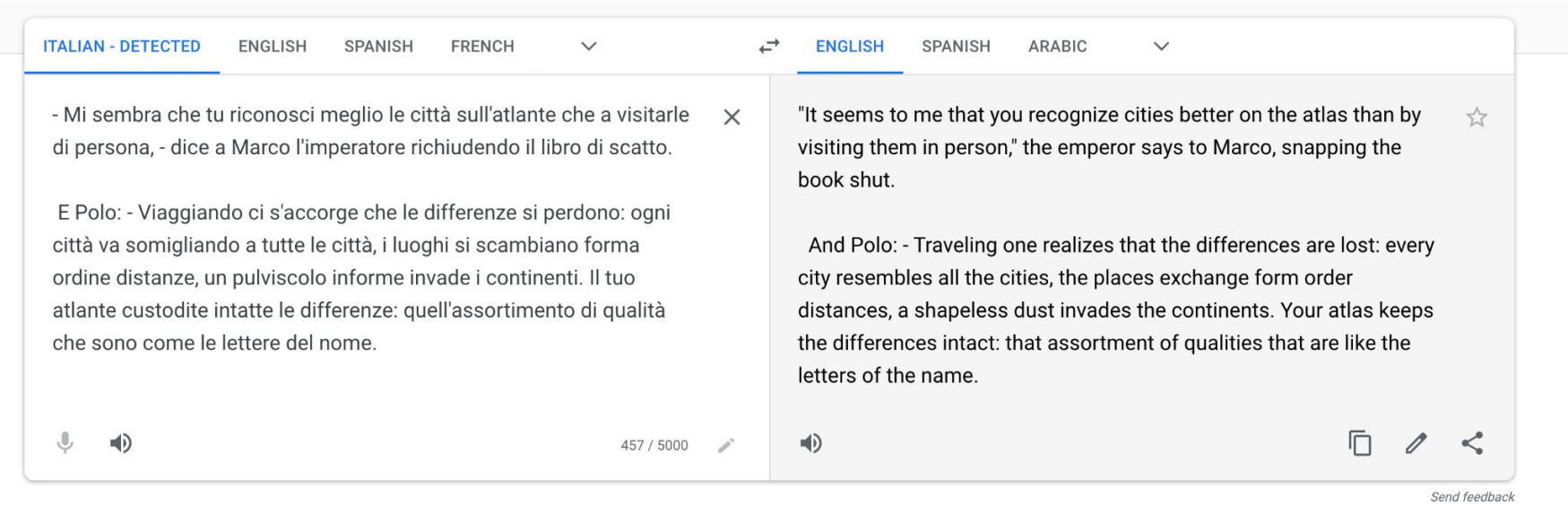
We haven't talked details yet, but you already know about the task of language modeling.

N-gram LMs are one type of algorithm for this task.

So really the question is: N-grams supervised or unsupervised?

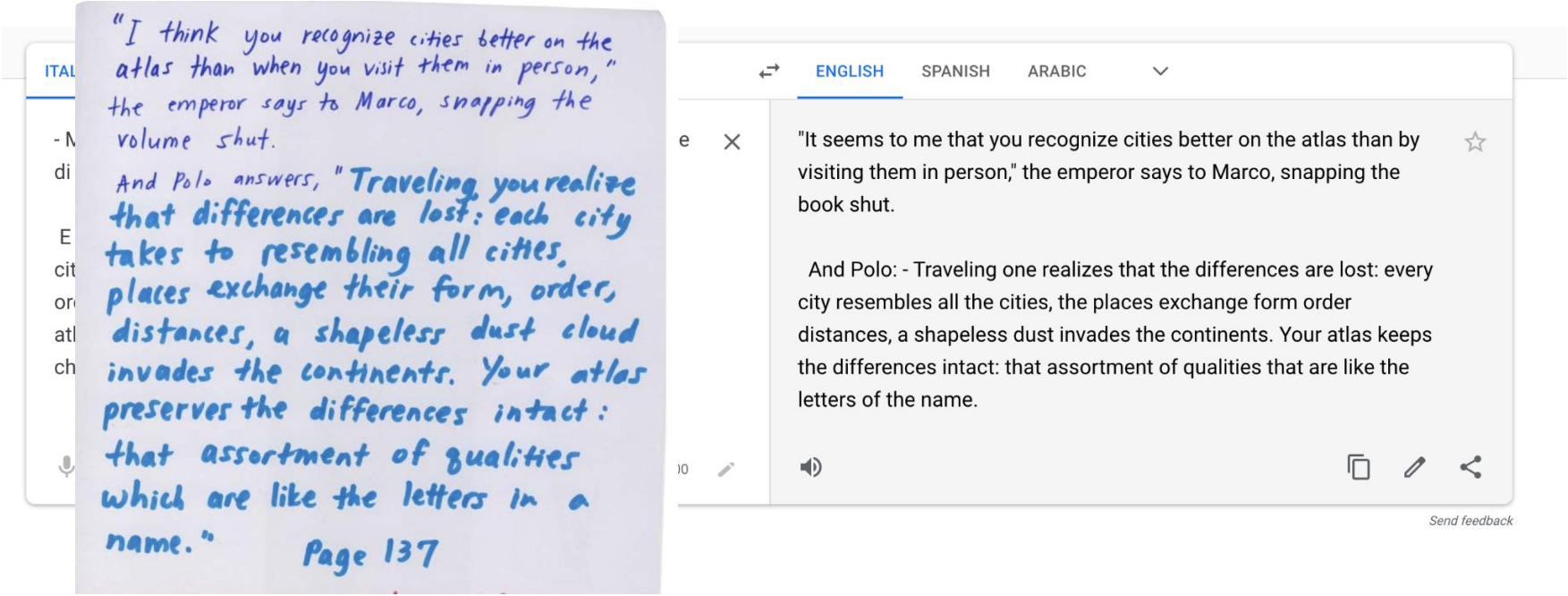


Supervised or Unsupervised?



Italo Calvino, *Invisible Cities*
https://monoskop.org/images/c/c4/Calvino_Italo_Le_citta_invisibili.pdf

Supervised!



Italo Calvino, *Invisible Cities*
https://monoskop.org/images/c/c4/Calvino_Italo_Le_citta_invisibili.pdf
Authoritative translation on the left by William Weaver

Supervised or Unsupervised?

- Speech-segmentation
- Spam filtering
- Auto-generated tags / topic labels
- Smart search results
- Fraud detection
- What are the different senses of a word?

The Curly Fry Conundrum (Jennifer Golbeck)

https://www.youtube.com/watch?v=hgWie9dnssU

Demographic study shows that “Liking” the page for curly fries is one of the strongest indicators of intelligence.

How can this be?

Latent Bias

Even if we remove demographic information from consideration (exclude features that correspond to culturally sensitive categories / identity labels), these distinctions are often still learned through correlation with other, less problematic categories.

Lipstick on a Pig:
Debiasing Methods Cover up Systematic Gender Biases in Word Embeddings But do not Remove Them

Hila Gonen[1] and **Yoav Goldberg**[1,2]
[1]Department of Computer Science, Bar-Ilan University
[2]Allen Institute for Artificial Intelligence
{hilagnn,yoav.goldberg}@gmail.com