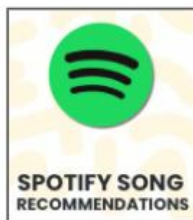
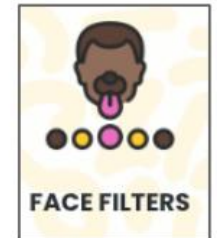
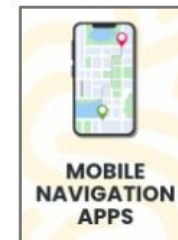
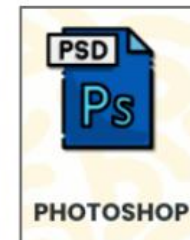
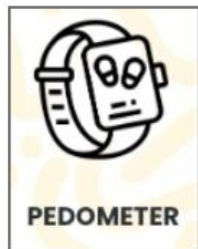


# AI Literacy & DS Ethics

Danni Presgraves, Ph.D.  
Director, Data Science Training

DEFINITELY  
**NOT** AI

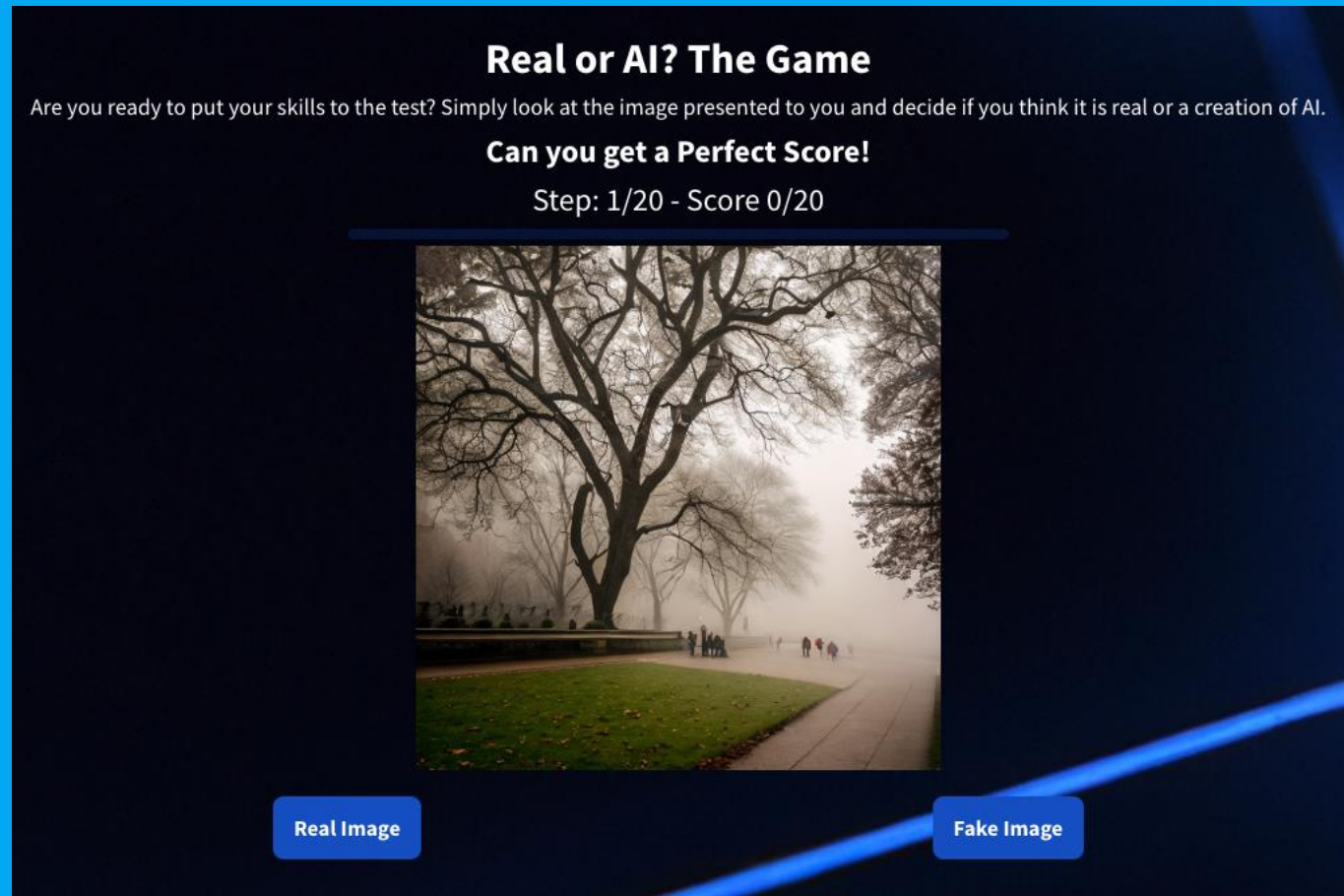
DEFINITELY  
**AI**





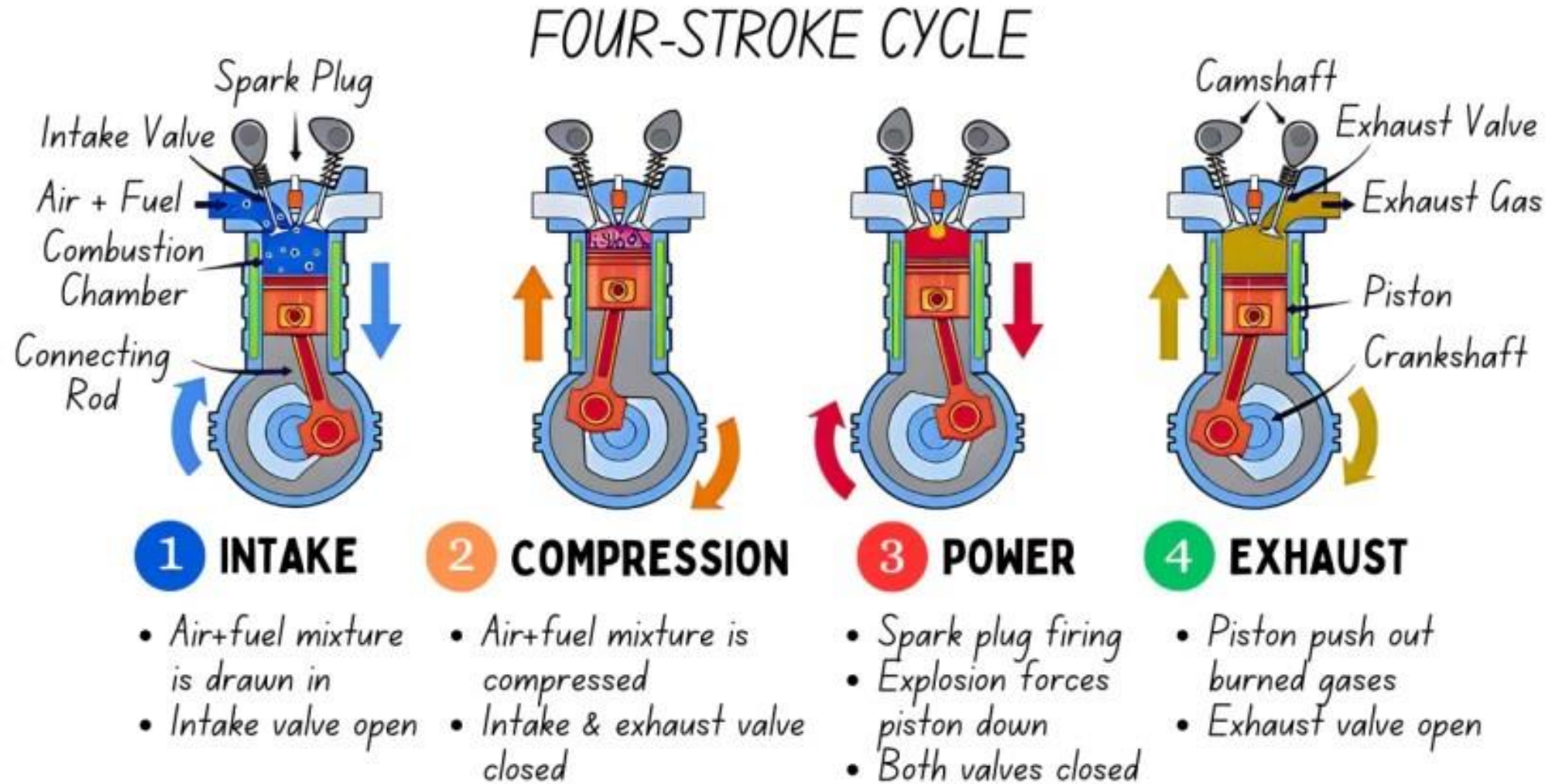
[The Canadian House Hippo](#)

What is Ai?

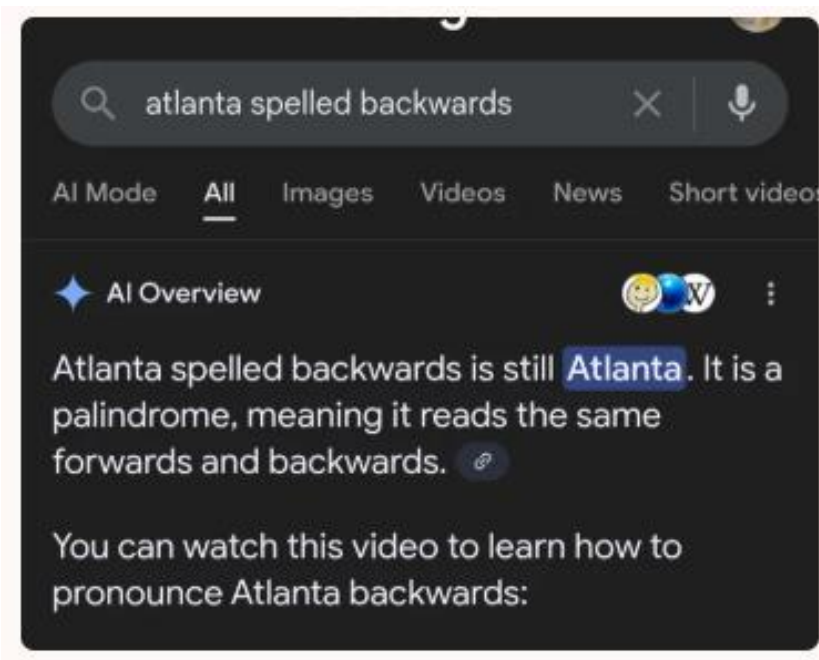
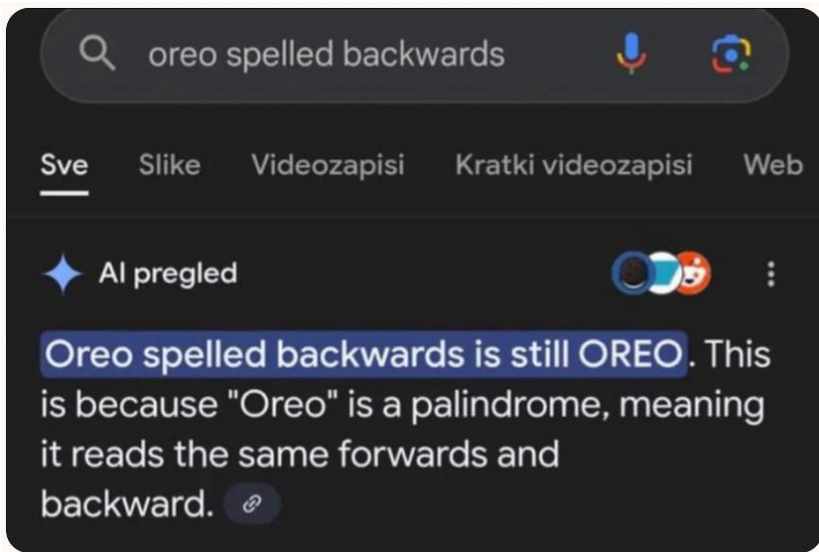


# Play the “Is it real or is it AI?” Game

# Purpose of this course:



How differential steering works (1937): <https://www.youtube.com/watch?v=yYAw79386WI>



= ChatGPT 5.1 >



Is 73 the 21st prime?

Seventy-three is **not** the twenty-first prime.

Here are the primes around that point:

- 19th prime: **67**
- 20th prime: **71**
- 21st prime: **73**
- 22nd prime: **79**

So **73 is the 21st prime number.**

[https://www.reddit.com/r/ChatGPT/comments/1pbv54l/73 is not the 21st prime/#lightbox](https://www.reddit.com/r/ChatGPT/comments/1pbv54l/73_is_not_the_21st_prime/#lightbox)


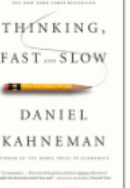
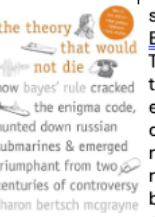
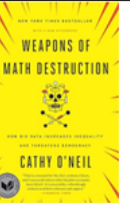
<https://www.reddit.com/r/mildlyinfuriating/comments/1mnakmj/really/>

What is AI?



## Crowd sourcing popular 'AI' (broadly defined) books via JAX

- Please add your thoughts

	Book	Summary	Additional Thoughts?	Final Rating
1		A slim book with a blunt message: either we shape technology, or it shapes us. The "commands" are about literacy - how to keep from being programmed by the very tools we've built. I treat technology as if it is neutral, when it's anything but. There is an unsettling sense that every click, scroll, or "like" is part of someone else's program, unless I resist.		
2		A classic that is about, well, thinking. <u>Premise</u> is that we have two systems of thinking. After years of teaching statistics, I already had a prior belief that our brain rebels against learning certain ideas (like statistics). I initially read this book because I realized that I did not understand what 'learning' is. I am still not sure <u>that</u> I do, but I am a more informed kind of ignorant.		
3		I will fight anyone who doesn't recognize the supremacy of Bayes' theorem in <u>Bioinformatics</u> etc. This is basically the biography of Bayes' theorem - which sounds dull, but it's weirdly engaging. The story winds from obscure clergymen to wartime codebreakers to modern data science, and along the way makes clear that probability isn't just math but a worldview. It's a reminder that even equations can be political - which is a theme for many of these broadly data science books.		
4		O'Neil takes a wrecking ball to the idea that algorithms are impartial. Her "weapons of math destruction" are the models that scale inequality while hiding behind a veneer of objectivity. I found myself nodding a lot since I tend to emphasize assumptions and limitations of models. Note: She has a fantastic blog called "mathbabe.org"		

## Popular\_AI\_books\_2025.docx

- I have written up a short summary of popular Ai books
- You can use it as a reference, or (please) add your own thoughts!

## Additional Resources for AI literacy

### Technical Skills:

Most of this list are code-free (or coding optional) resources to learn more about a particular concept. There are a handful of options on this list that do require some coding (Python or R), but they are marked.

#### 1. **Data Science & AI course:** [Data Carpentries beta version](#)

This is an ~ 2.5-hour self-directed course that gives biomedical and life-science investigators a rapid, practical overview of data-science and AI concepts. It starts with core definitions of data science, AI and machine-learning, then walks through AI for laboratory automation and for deriving data insights, before digging into common pitfalls such as bias, privacy and reproducibility, and finishing with hands-on guidance for cleaning data, running ML pipelines and responsibly reporting results.

#### 2. **Neural Networks:** [Interactive free 'textbook' by 3Blue1Brown](#)

You will find explanations of all the basics of Neural Networks, including gradient descent and backpropagation (which is sort of a fancier version of l'Hôpital's rule...). He integrates YouTube explanations into his broader explanations.

#### 3. **Interactive Neural Network:** [play with the parameters!](#)

#### 4. **Elements of AI:** [A free course](#)

Three different levels of course: introduction (non-coding), coding (pre-requisites: Python), and a level of incorporating AI into a Learning Management System (which is portable to other pipelines).

#### 5. **Machine Learning for Biologists:** [Data Carpentry](#)

Accessible with biological examples – includes logistic regression, linear regression, Decision trees, random forests, Neural Networks, Overfitting. Great place to start since there is no coding necessary.

#### 6. **Machine Learning:** [Machine Learning in Medicine](#)

**Coding is included in this course** (but you read through and watch the material to get a better understanding without coding) Developed by evolutionary biologist Pleuni Pennings (SFSU) and collaborators, this open-access course targets life-science trainees who know basic Python but are new to machine-learning. It blends brief concept videos, step-by-step editable Colab notebooks and concise Documents. It contains 7 autonomous (and self-contained) modules. Each module is ~ 2 - 3 hours long:

Term	Definition
<b>Artificial Intelligence (AI)</b>	The broad field focused on making machines perform tasks that normally require human intelligence.
<b>Machine Learning (ML)</b>	A subset of AI where systems learn from data to make predictions or decisions without being explicitly programmed.
<b>Supervised Learning</b>	Training models on labeled data where the correct answers are known.
<b>Unsupervised Learning</b>	Training models on unlabeled data to find hidden structure or patterns.
<b>Feature</b>	A measurable property or characteristic of data used as input for a model (e.g., pixel brightness, word frequency, gene expression level).
<b>Model</b>	A mathematical representation of a system that maps inputs (features) to outputs (predictions).
<b>Loss Function</b>	A formula that quantifies how far the model's predictions are from the true values.
<b>Gradient Descent</b>	An optimization process that updates model parameters to minimize loss by following the slope ("gradient") downhill.
<b>Overfitting / Underfitting</b>	When a model learns noise instead of pattern (overfitting) or fails to capture the structure (underfitting).
<b>Bias-Variance Tradeoff</b>	The balance between model complexity (variance) and generalization (bias).

# Day 1 Agenda:

## 1. AI Overview

- Big ideas, history, AI conceptual flowchart

## 2. Machine Learning Examples

- What is *learning*?
- Supervised Learning examples: regression, classification
- Unsupervised Learning examples: k-means clustering
- Simple Neural Network
- Overview of concepts: Loss Function Minimization, Gradient Descent

## Day 2 (Tuesday)

1. Review
2. Chat**GPT**
- 3.** **G**enerative
- 4.** **P**re-trained
- 5.** **T**ransformed
- 6. Dr. Derya Unutmaz**

## Day 3 (Wednesday)

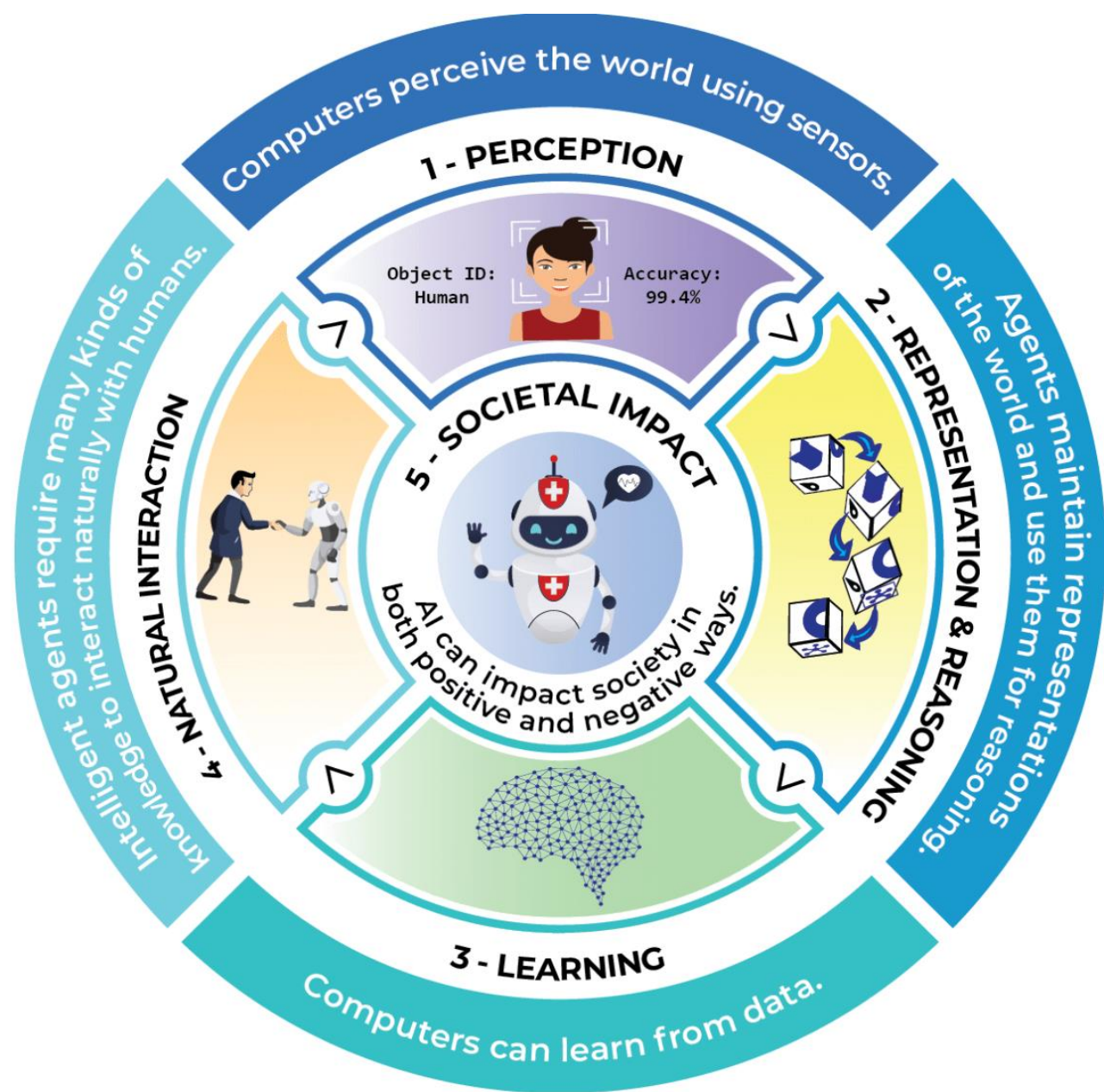
1. Spill over from yesterday
2. Alignment
3. Ethics
4. Prompt engineering
5. Gemini

## Day 4 (Thursday)

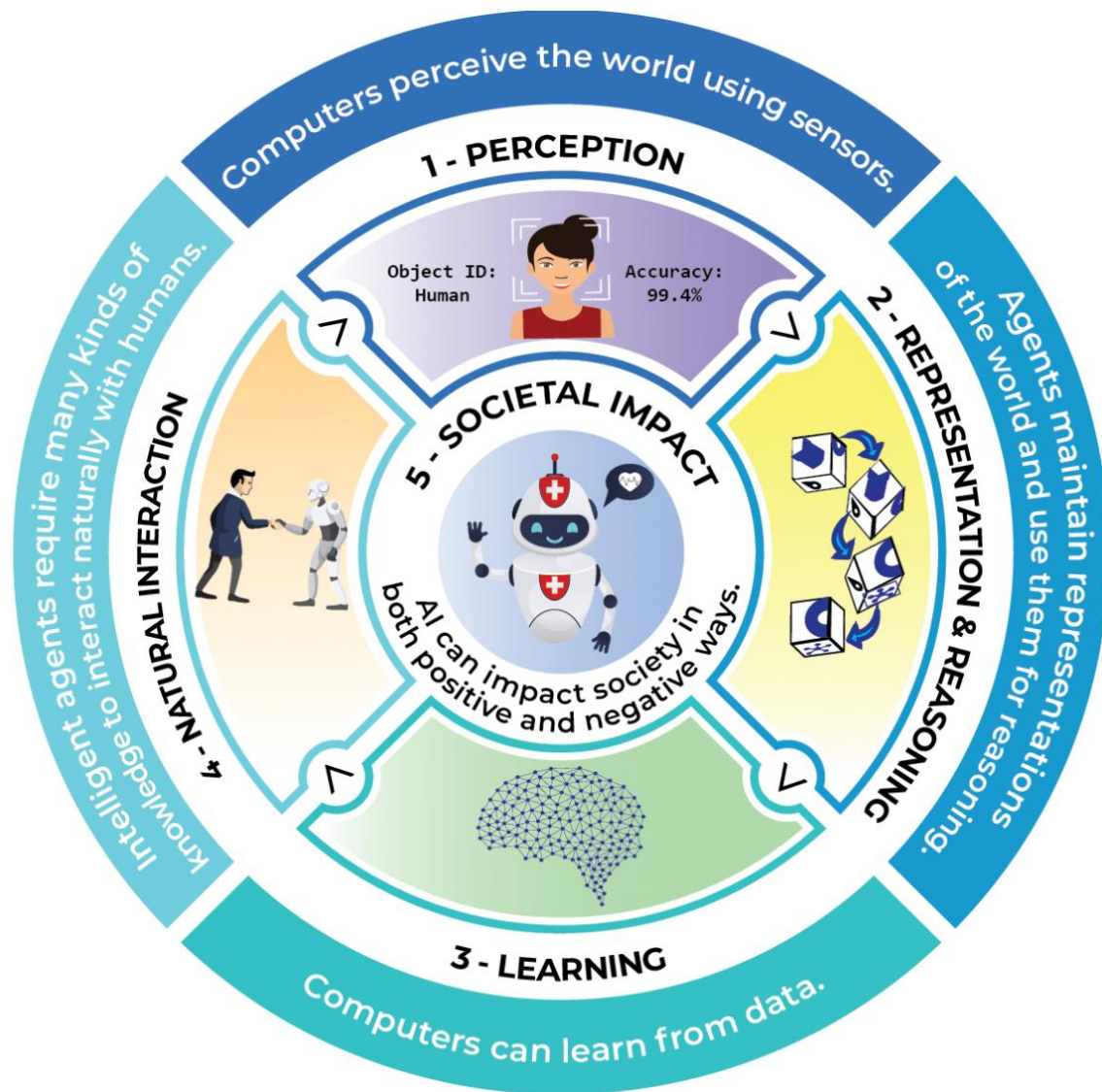
1. Review
2. Artificial Neurons
3. CNNs
4. RNNs

## Day 5 (Friday)

1. Generative modules & Transformers
2. GANs
3. DS Balderdash
4. Transformers
5. **Jeff Chuang (2:30-3:30)**



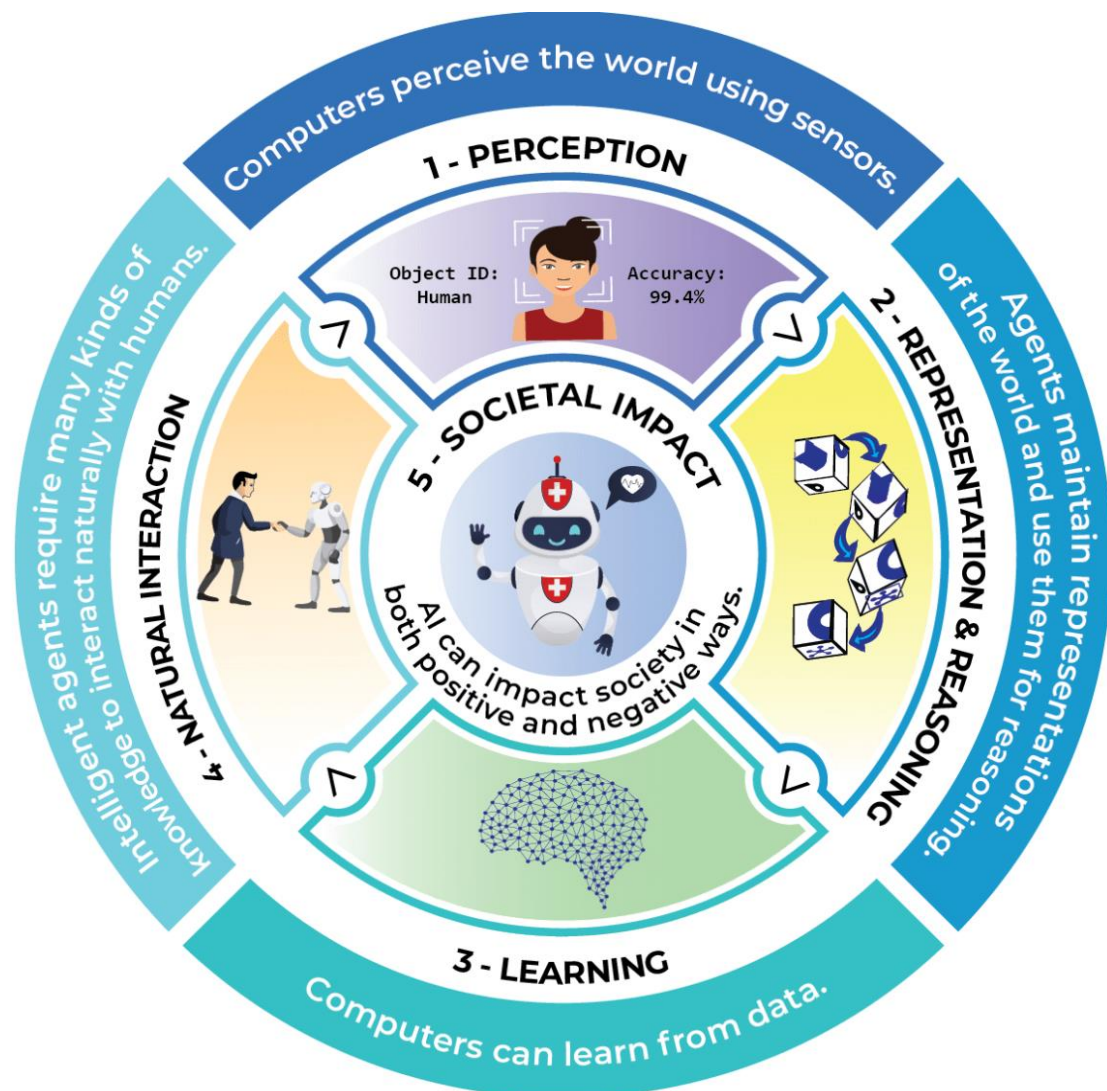
Five Big Ideas Wheel



How do real-world data, like text, images, audio become numerical input?

AI doesn't *eliminate* the need for thinking clearly, it amplifies the consequences of thinking *poorly*.

What *is* learning?



Five Big Ideas Wheel

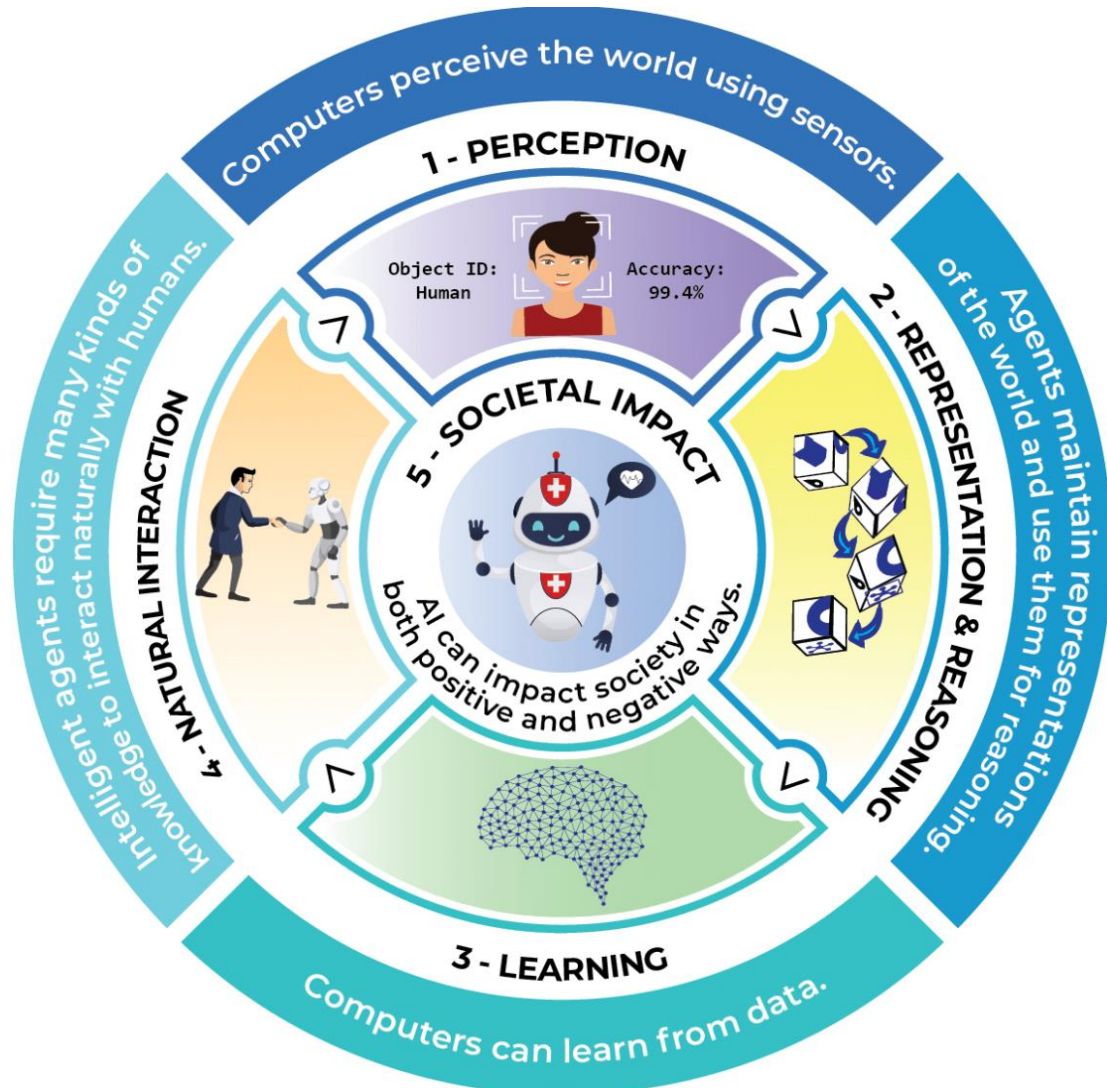
How do real-world data, like text, images, audio become numerical input?



AI doesn't *eliminate* the need for thinking clearly, it amplifies the consequences of thinking *poorly*.



What *is* learning?



Five Big Ideas Wheel

How do real-world data, like text, images, audio become numerical input?

- Images → pixel matrices
- Text → tokens → numbers
- Categories → one hot encoding

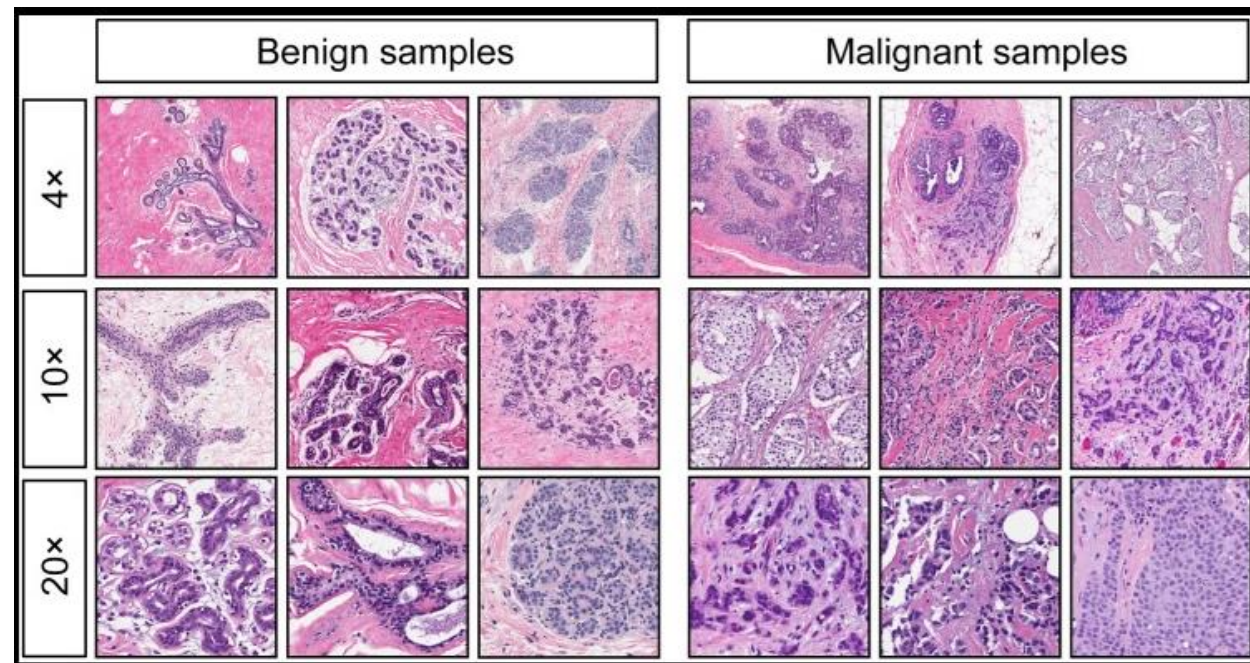
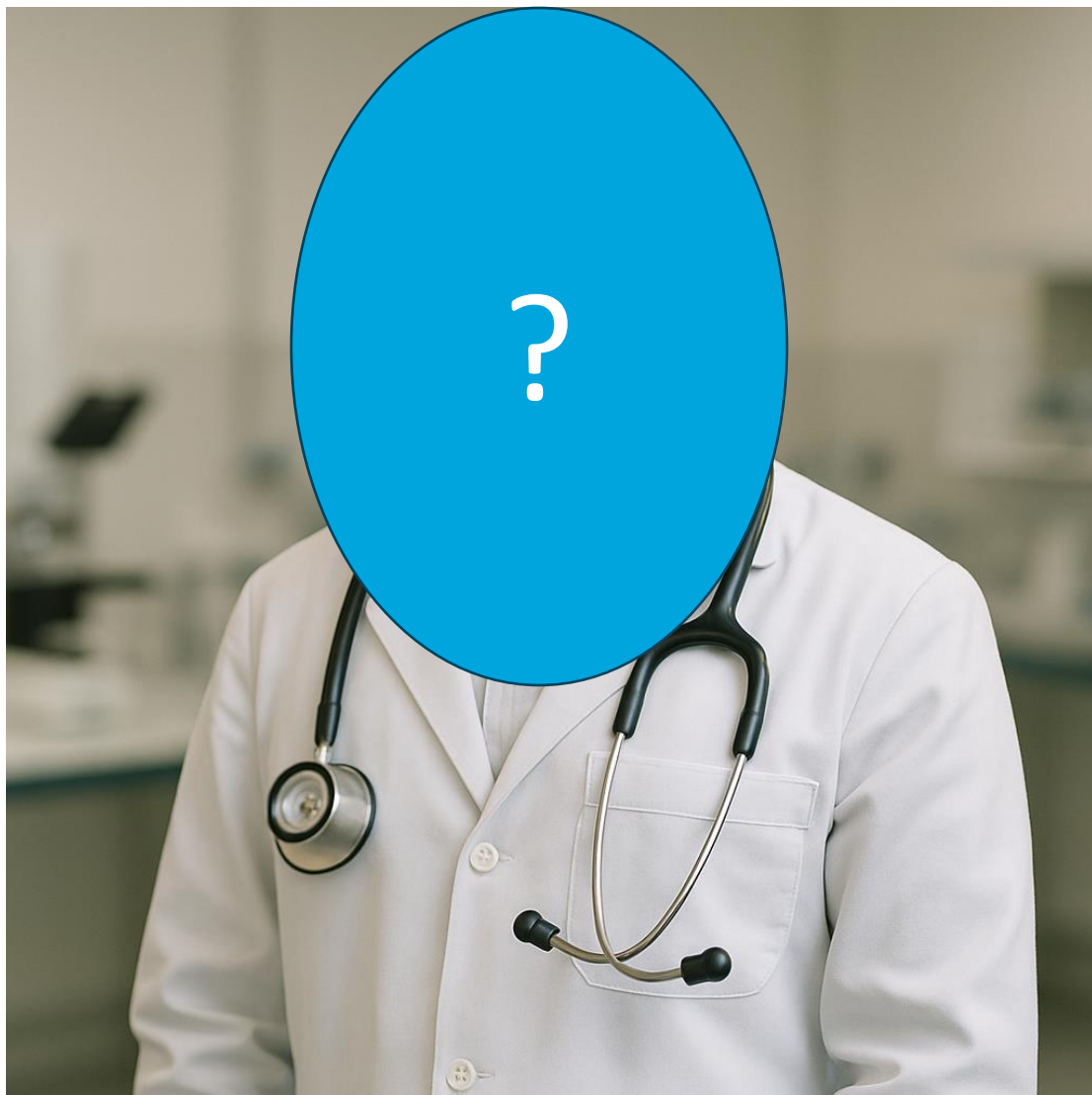


AI doesn't *eliminate* the need for thinking clearly, it amplifies the consequences of thinking *poorly*.



What *is* learning?

## Side bar: What *is* learning?



# What is learning?



► PLoS One. 2015 Nov 18;10(11):e0141357. doi: [10.1371/journal.pone.0141357](https://doi.org/10.1371/journal.pone.0141357)

## Pigeons (*Columba livia*) as Trainable Observers of Pathology and Radiology Breast Cancer Images

[Richard M Levenson](#) <sup>1,\*</sup>, [Elizabeth A Krupinski](#) <sup>3</sup>, [Victor M Navarro](#) <sup>2</sup>, [Edward A Wasserman](#) <sup>2,\*</sup>

Editor: Jonathan A Coles<sup>4</sup>

► [Author information](#) ► [Article notes](#) ► [Copyright and License information](#)

PMCID: PMC4651348 PMID: [26581091](https://pubmed.ncbi.nlm.nih.gov/26581091/)

### Flock wisdom:

- By day 14, individual birds had 85% accuracy
- A 4 pigeon ‘flock’ had 99% accuracy → consensus

[Home](#) > [Learning & Behavior](#) > [Article](#)

## Taking pigeons to heart: Birds proficiently diagnose human cardiac disease

Published: 21 January 2020

Volume 48, pages 9–21, (2020) [Cite this article](#)

# Is Artificial Intelligence new?

1943	<b>McCullouch and Pitts</b> <ul style="list-style-type: none"> <li>- A Logical Calculus of the Ideas Immanent in Nervous Activity</li> <li>- laid foundations for “artificial neurons”</li> </ul>
1949	<b>Hebb</b> <ul style="list-style-type: none"> <li>- “fire together, wire together”</li> </ul>
1950	<b>Turing</b> <ul style="list-style-type: none"> <li>- “Computing Machinery and Intelligence”</li> </ul>
1956	<b>McCarthy</b> <ul style="list-style-type: none"> <li>- Coined the term: Artificial Intelligence</li> <li>- thinking involves both <i>knowledge</i> and <i>reasoning and</i> can be encoded mathematically</li> </ul>
1957	<b>Rosenblatt</b> <p>perceptrons → Neural Networks</p>
1967	<b>Minsky</b>
2009	<b>Fei-Fei Li</b> <ul style="list-style-type: none"> <li>- ImageNet (14 million images labeled)</li> </ul>
2012	<b>AlexNet</b>
2024	<b>Nobel Prizes</b> <ul style="list-style-type: none"> <li>- <b>Physics:</b> Geoffrey Hinton, John J. Hopfield</li> <li>- <b>Chemistry:</b> John J Jumper, David Baker, Demis Hassabis</li> </ul>

1308

Catalan poet and theologian [Ramon Llull](#) publishes *Ars generalis ultima* (The Ultimate General Art), further perfecting his method of using paper-based mechanical means to create new knowledge from combinations of concepts.

1666

Mathematician and philosopher Gottfried Leibniz publishes *Dissertatio de arte combinatoria* (On the Combinatorial Art), following Ramon Llull in proposing an alphabet of human thought and arguing that all ideas are nothing but combinations of a relatively small number of simple concepts.

1726

Jonathan Swift publishes *Gulliver's Travels*, which includes a description of [the Engine](#), a machine on the island of Laputa (and a parody of Llull's ideas): "a Project for improving speculative Knowledge by practical and mechanical Operations." By using this "Contrivance," "the most ignorant Person at a reasonable Charge, and with a little bodily Labour, may write Books in Philosophy, Poetry, Politicks, Law, Mathematicks, and Theology, with the least Assistance from Genius or study."

1763

Thomas Bayes develops a framework for reasoning about the probability of events. [Bayesian inference](#) will become a leading approach in machine learning.

1854

[George Boole](#) argues that logical reasoning could be performed systematically in the same manner as solving a system of equations.

1898

At an electrical exhibition in the recently completed Madison Square Garden, Nikola Tesla makes a demonstration of [the world's first radio-controlled vessel](#). The boat was equipped with, as Tesla described, "a borrowed mind."

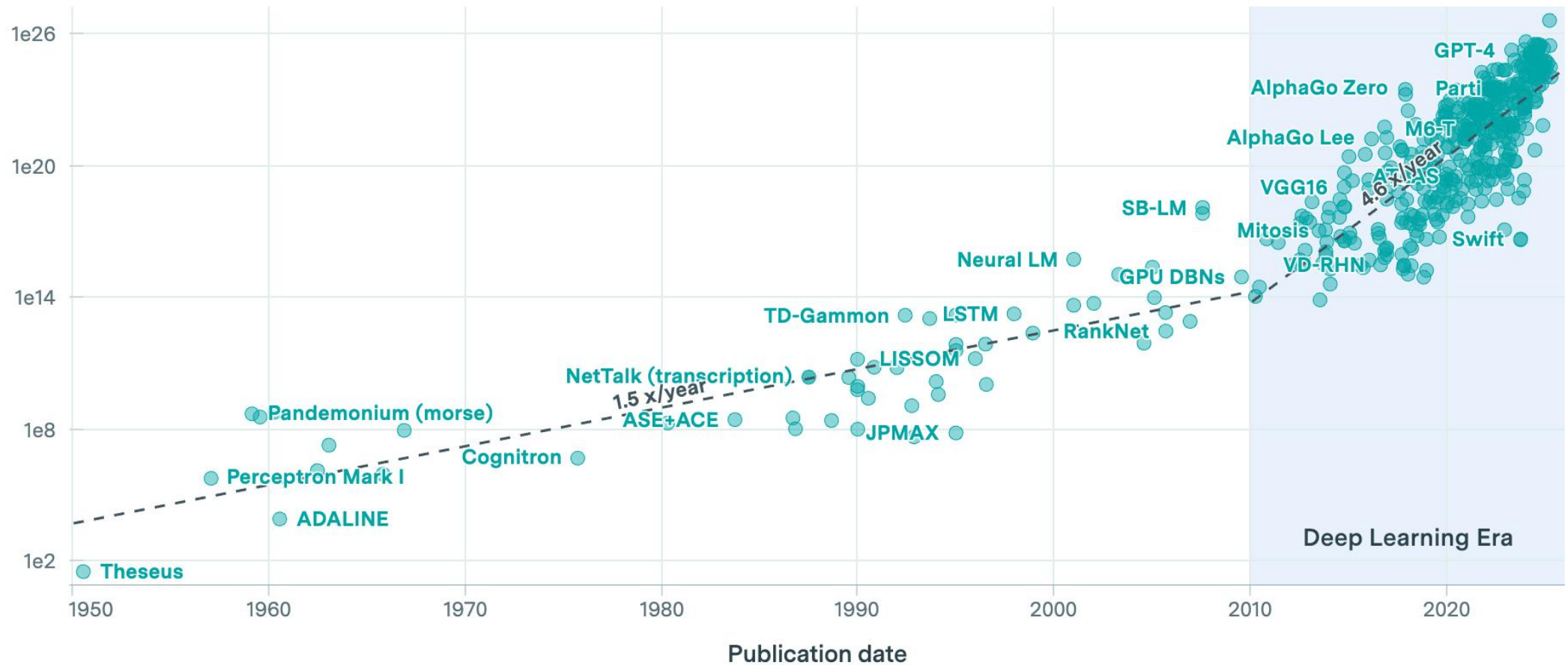
1914

The Spanish engineer [Leonardo Torres y Quevedo](#) demonstrates the first chess-playing machine, capable of king and rook against king endgames without any human intervention.

# Notable AI Models

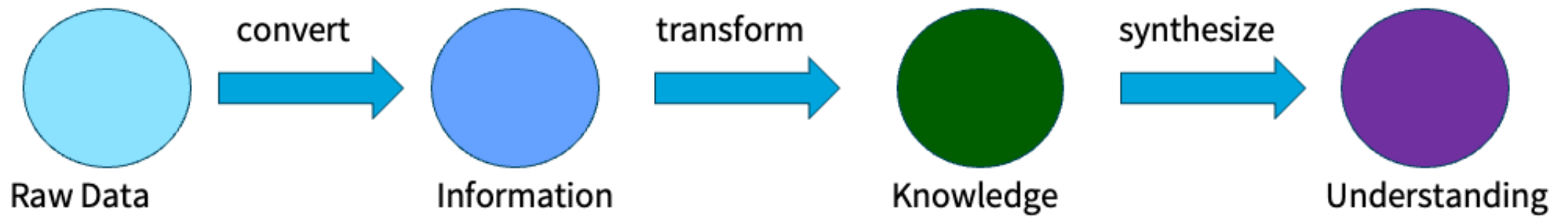
Training compute (FLOP)

469 Results

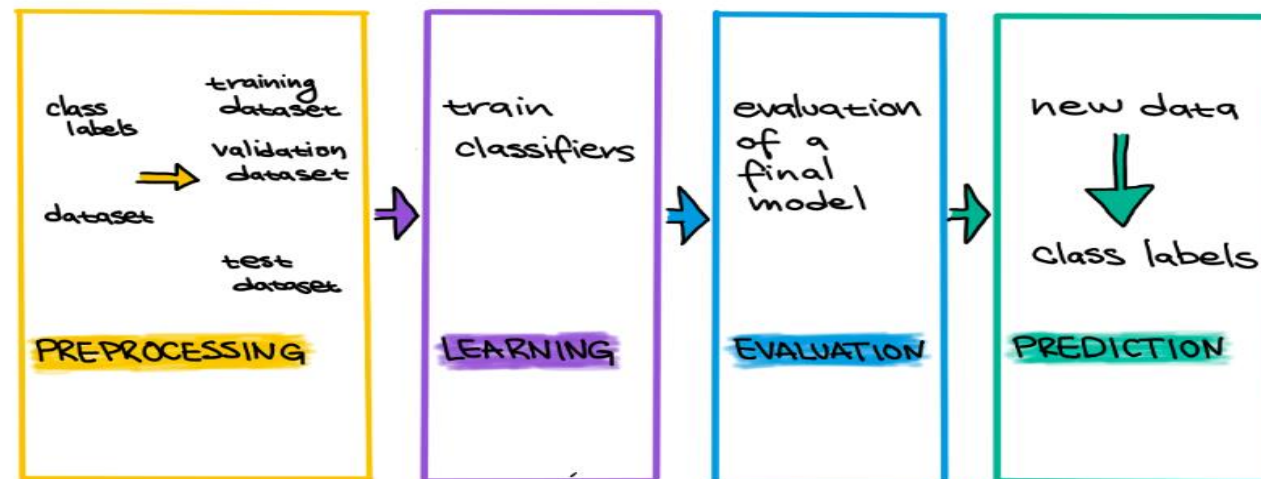


CC-BY

epoch.ai



Machine learning workflow



# Raw Data Becomes Numbers:

- All inputs – pixels, words, sounds – must be converted into numbers so that models can process them

## Images

Pixels → RGB values → Matrices

## Categories

Labels → One hot Encoding

## Text

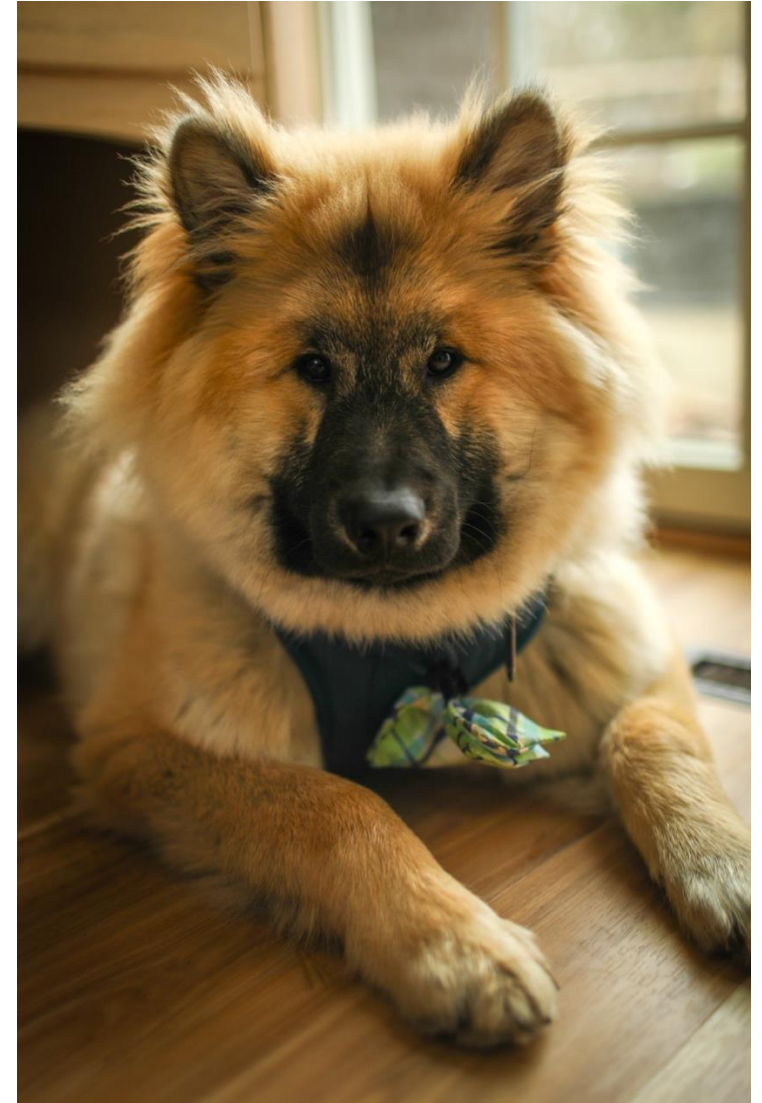
Words → Tokens → Embeddings

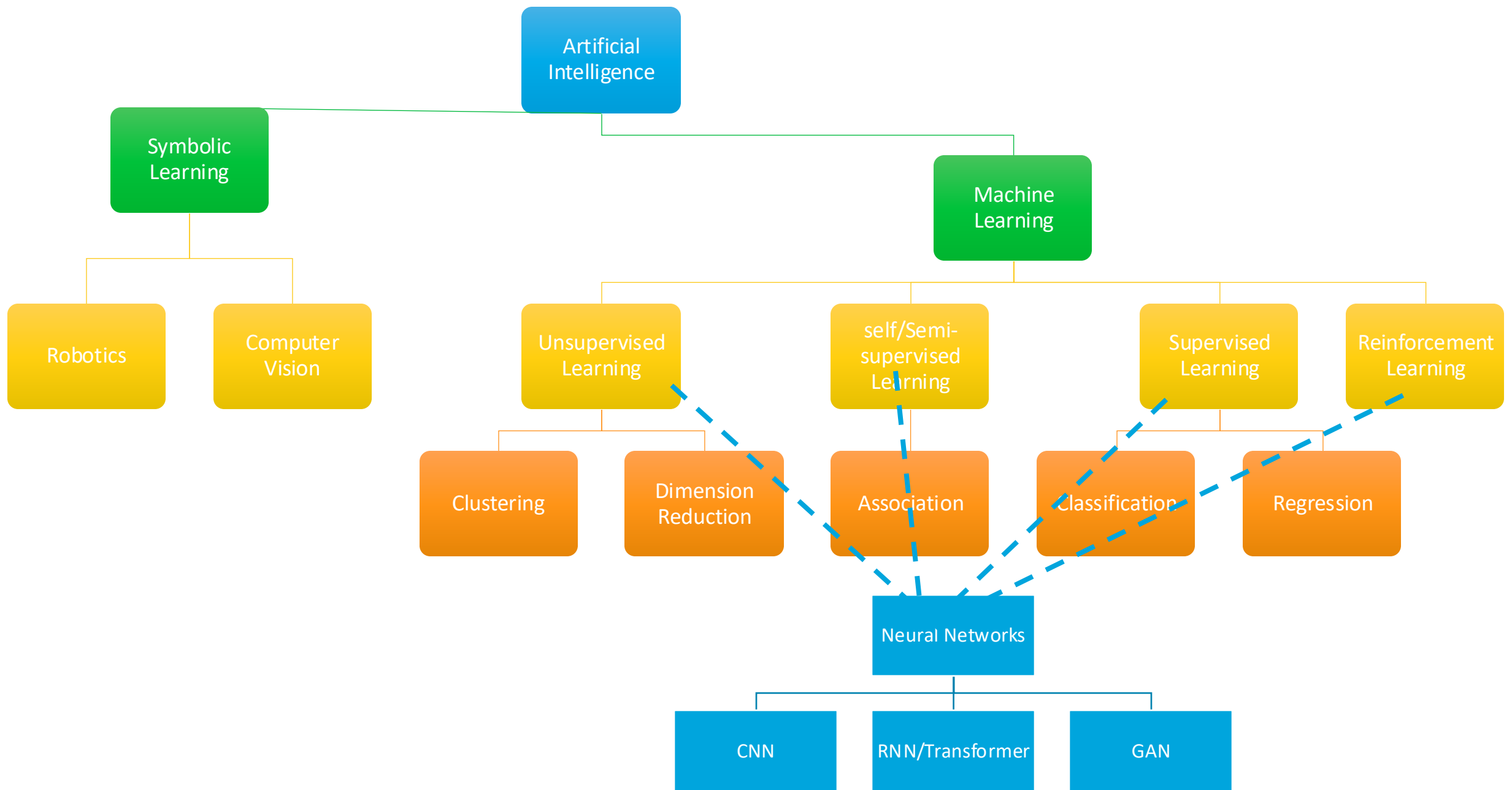
## Audio

Waveforms → Frequencies → Vectors

# How does raw data become numbers?

Huxley is an 18-month-old Eurasier (male, not neutered) who came in on 110/03/2025 for a sick visit because his owner said he hasn't been eating normally and vomited twice. The tech wrote his weight as **40 pounds**, though later in the notes someone jotted “~18 kg?” with a question mark next to it. On exam, Huxley's temperature was 102.1 F, heart rate 96 bpm, and his respiratory rate was listed as “**about 30/min**,” although the structured intake sheet earlier said 28/min. His body condition score was 6/9 and hydration was estimated around 5% dehydrated. Pain was mild (1 on the 0–4 scale). CBC showed WBC 14.2 (slightly high), RBC 6.3, hemoglobin 15.1 g/dL, and platelets 268. Chemistry panel had BUN 21, creatinine 1.1, ALT 68 (a bit high), and glucose 87. Abdominal radiographs noted some gastric fluid but no foreign body. The assessment was possible gastritis with mild dehydration, no obstruction suspected. Plan: maropitant at 1 mg/kg once daily for 3 days, 200 mL SQ fluids, bland diet for 2–3 days with small frequent meals and return if vomiting continues for more than 48 hours.





# Machine Learning

- Not necessarily **predictive**
- Group and interpret data based on **input**

Develop **predictive model** based on **input** and **output**

## Unsupervised Learning

## Supervised Learning

Data is labelled

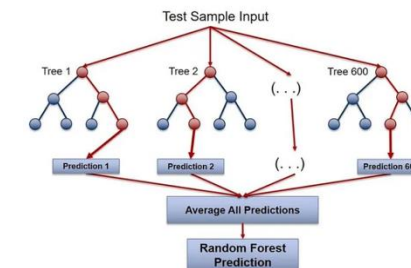
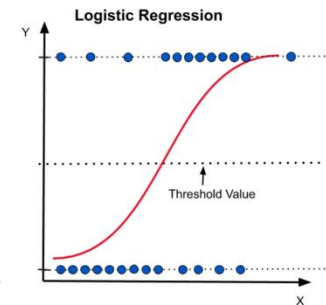
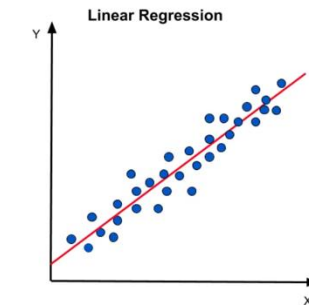
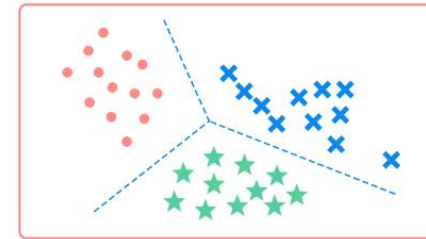
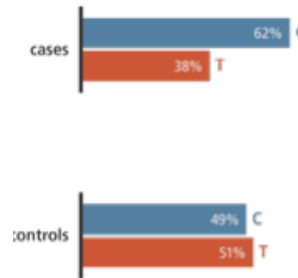
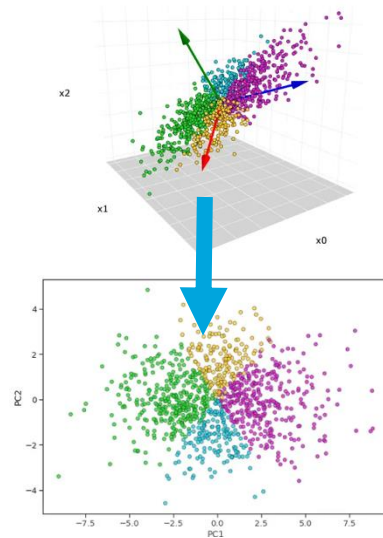
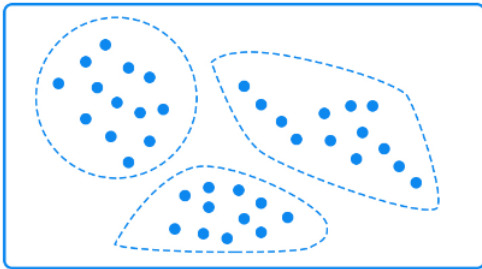
Clustering

Dimension Reduction

Association

Classification

Regression



# Machine Learning

## Unsupervised Learning

## Supervised Learning

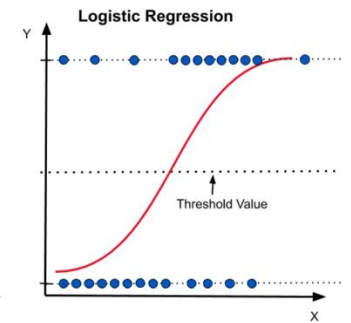
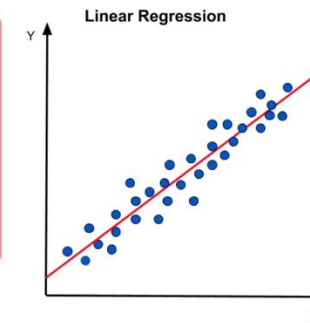
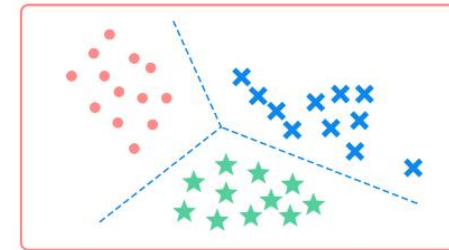
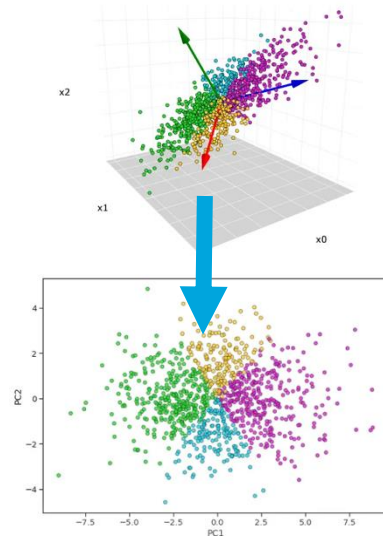
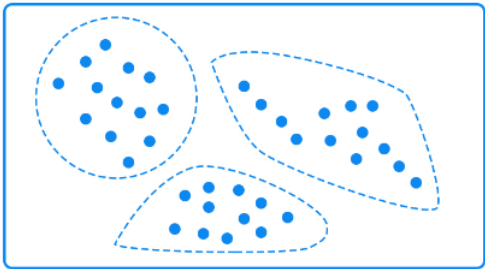
Clustering

Dimension Reduction

Association

Classification

Regression

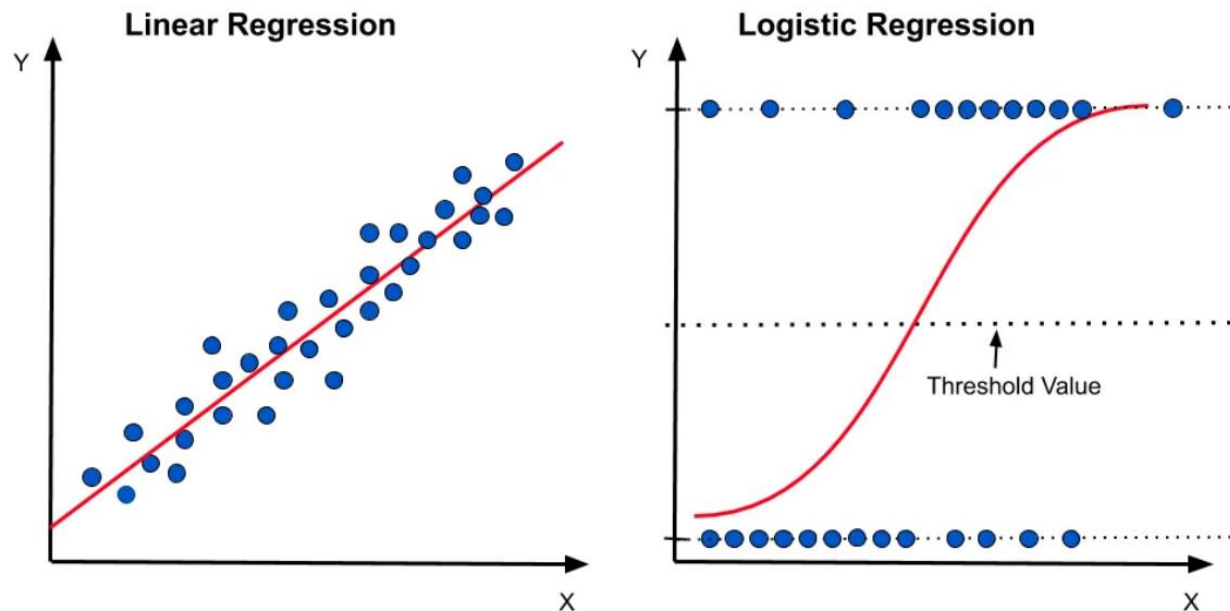


# 10-minute Break

Return at:

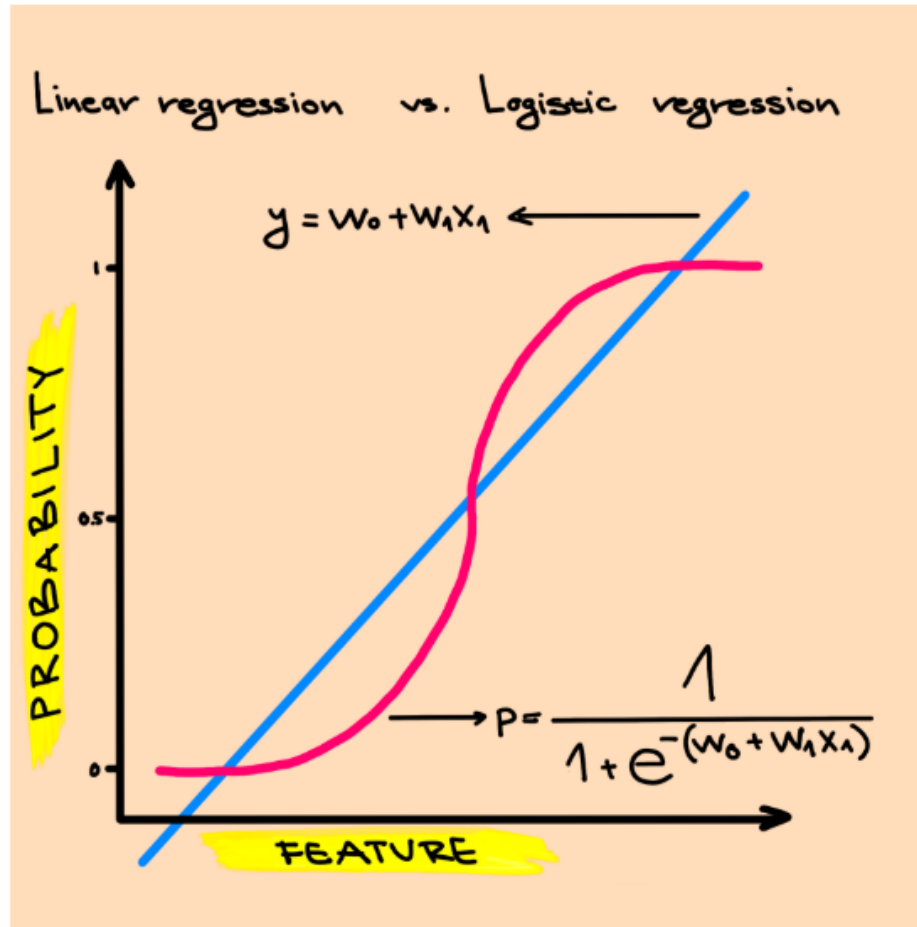
# In machine learning, computers learn patterns that help them to make decisions without being explicitly programmed.

- A familiar way to make predictions is with regression:



For decision trees example: <http://www.r2d3.us/visual-intro-to-machine-learning-part-1/>

# Supervised ML: **Logistic & Linear Regression:**



<https://carpentries-incubator.github.io/ml4bio-workshop/05-logit-ann/index.html>

## Questions to consider for Logistic Regression:

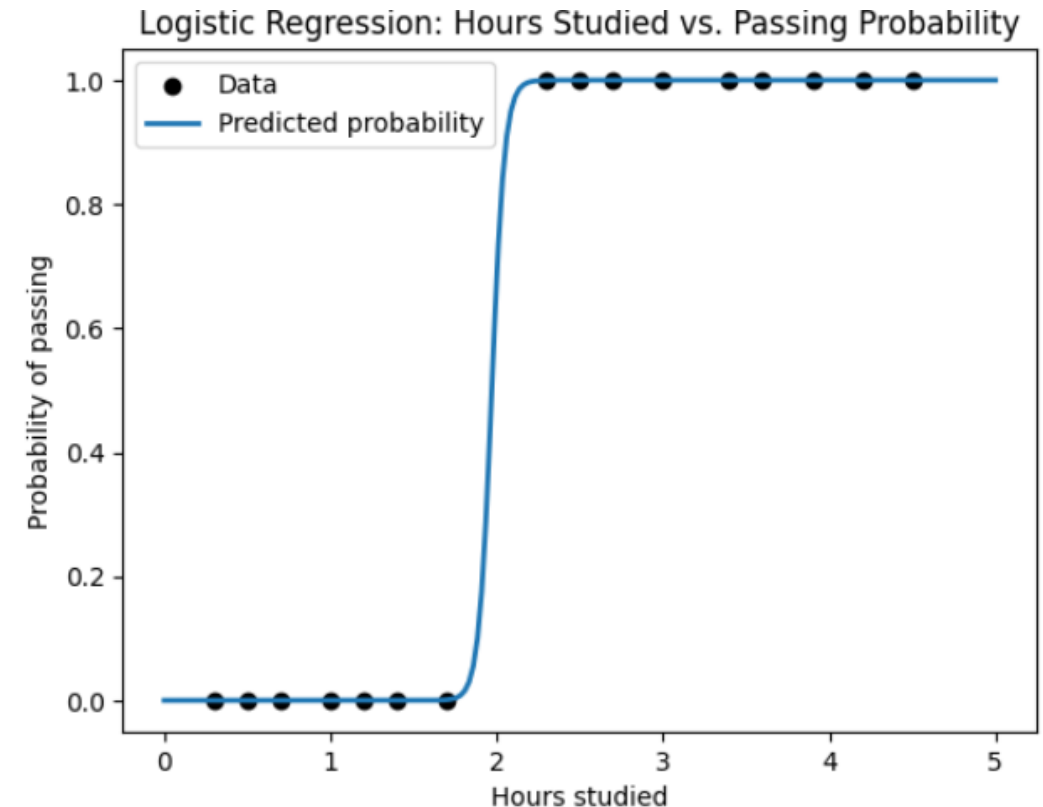
- How do AI systems use features (input variables) to generate predictions?
- How do supervised models make decisions?
- Can we impact the decision that it makes?
- Consider where logistic regression models are used (healthcare, criminal justice, finance). What are the risks of defining the cut-off value? Of including certain features and not others?

# Supervised ML: **Logistic Regression:**

Hours studied	Result
0.3, 0.5, 0.7, 1.0, 1.2, 1.4, 1.7	Fail
2.3, 2.5, 2.7, 3.0, 3.4, 3.6, 3.9, 4.2, 4.5	Pass

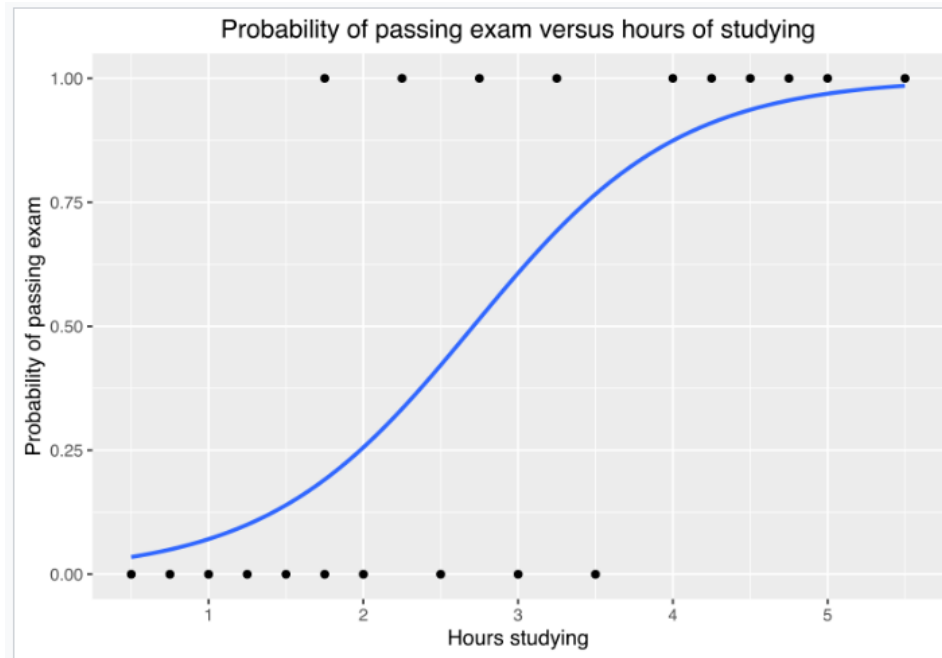
# Supervised ML: **Logistic Regression:**

Hours studied	Result
0.3, 0.5, 0.7, 1.0, 1.2, 1.4, 1.7	Fail
2.3, 2.5, 2.7, 3.0, 3.4, 3.6, 3.9, 4.2, 4.5	Pass



# Supervised ML: **Logistic Regression:**

Hours studied	Result
0.5, 0.75, 1.0, 1.25, 1.5, 1.75, 2.0, 2.5, 3.0, 3.5	Fail
1.75, 2.25, 2.75, 3.25, 4.0, 4.25, 4.5, 4.75, 5.0, 5.5	Pass



Graph of a logistic regression curve fitted to the  $(x_m, y_m)$  data. The curve shows the probability of passing an exam versus hours studying.

# Questions to consider for the following simulation:

- Why are k-means clustering *unsupervised*?
- Can unsupervised algorithms uncover real patterns?
- How does k-means demonstrate how learning is done via algorithms?
- Are these clusters ‘real’?
- Parameter ‘knobs’:
  - Does it matter how many k are chosen (k=2,k=3,etc)
  - Do different initial placements change the results?

# Unsupervised ML: **Classification**

## Visualizing K-Means Clustering

January 19, 2014

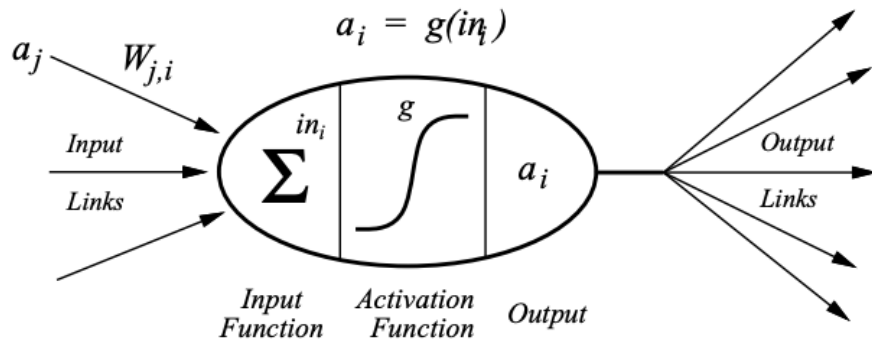
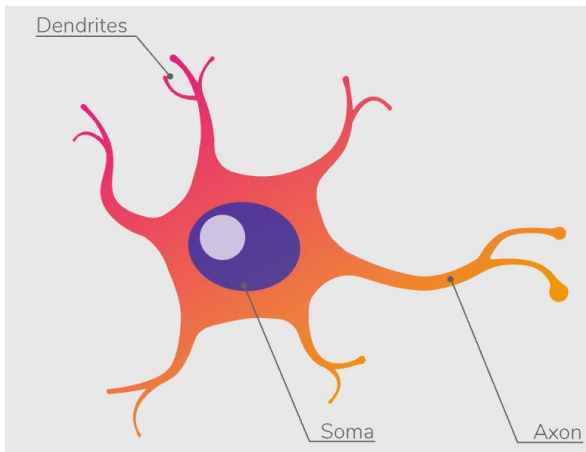
Suppose you plotted the screen width and height of all the devices accessing this website. You'd probably find that the points form three clumps: one clump with small dimensions, (smartphones), one with moderate dimensions, (tablets), and one with large dimensions, (laptops and desktops). Getting an algorithm to recognize these clumps of points without help is called *clustering*. To gain insight into how common clustering techniques work (and don't work), I've been making some visualizations that illustrate three fundamentally different approaches. This post, the first in this series of three, covers the k-means algorithm. To begin, click an initialization strategy below:

## How to pick the initial centroids?

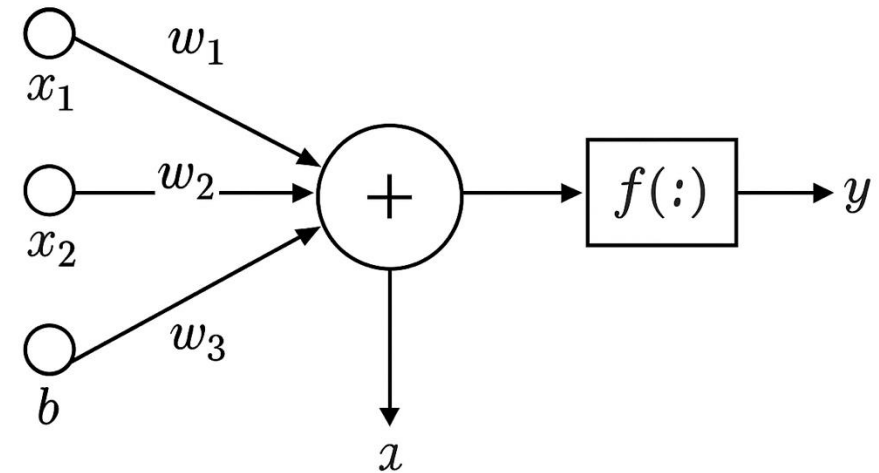
I'll Choose

Randomly

Farthest Point



$$a_i = g\left(\sum_j W_{j,i} a_j\right)$$



Inputs+ Bias + weights  $\rightarrow$  hidden layers  $\rightarrow$  predictions

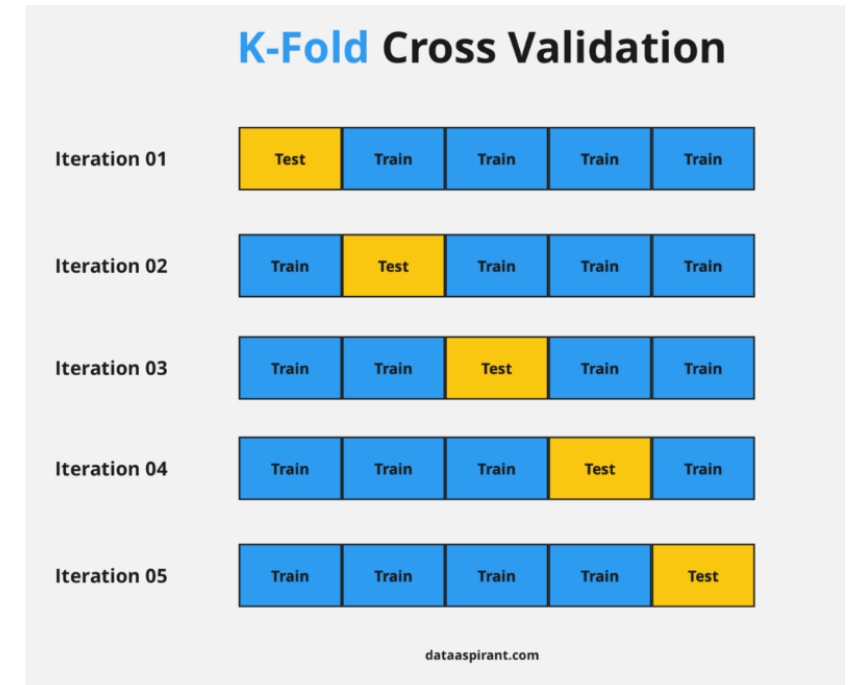
# Training versus Testing

- Strategy:

split the data and use some of it for **training** the NN to **make predictions** and **reserve** some of it **to test** the NN



- K-fold cross-validation



Hours Studied	Result
0.3	Fail
0.5	Fail
0.7	Fail
1.0	Fail
1.2	Fail
1.4	Fail
1.7	Fail
2.3	Pass
2.5	Pass
2.7	Pass
3.0	Pass
3.4	Pass
3.6	Pass
3.9	Pass
4.2	Pass
4.5	Pass

K=5  
→

**Fold 1:** 0.3, 1.2, 2.5

**Fold 2:** 0.5, 2.7, 3.9

**Fold 3:** 0.7, 1.7, 3.0

**Fold 4:** 1.0, 2.3, 3.4

**Fold 5:** 3.6, 1.4, 4.2, 4.5

Training

1,2,3,4

1,2,3,5

1,2,4,5

1,3,4,5

2,3,4,5

Testing

5

4

3

2

1

Hours Studied	Result
0.5	Fail
0.75	Fail
1.0	Fail
1.25	Fail
1.5	Fail
1.75	Fail
2.0	Fail
2.5	Fail
3.0	Fail
3.5	Fail
1.75	Pass
2.25	Pass
2.75	Pass
3.25	Pass
4.0	Pass
4.25	Pass
4.5	Pass
4.75	Pass
5.0	Pass
5.5	Pass

K=5  
→

**Fold 1:** 0.5,1.5,3.0,3.25

**Fold 2:** 0.75, 1.75, 3.5, 4.0

**Fold 3:** 1.0,2.75, 4.0, 5.5

**Fold 4:** 1.25, 2.5, 3.25, 5.0

**Fold 5:** 2.0, 4.25, 4.5, 4.75

Training

Testing

1,2,3,4

5

1,2,3,5

4

1,2,4,5

3

1,3,4,5

2

2,3,4,5

1

## **Question to consider for The Weight Adjustment Game:**

- How do we get the optimal weights?

## **Question to consider for Fair Treasure Splitting Game:**

- What trade-offs were you willing to accept?

# The Weight Adjustment Game

Age	Glucose	Has Diabetes (Yes = 1, No = 0)
40	130	1
25	100	0

# 10-minute Break

Return at: XX:XX

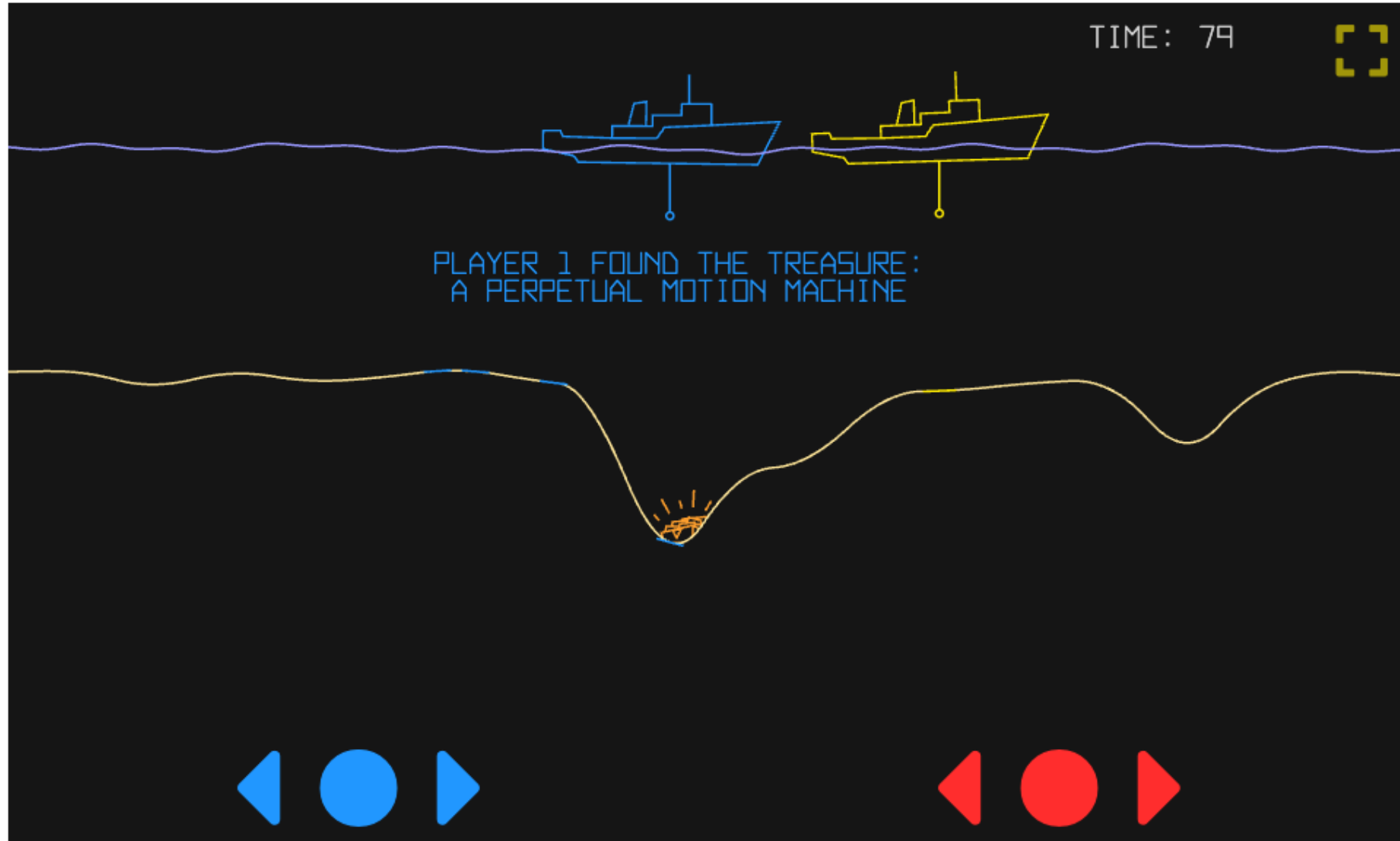
# Fair Treasure Splitting Game

1. Your group will receive one set of diamonds
2. You need to split the diamonds into 2 groups
3. Calculate the total value of each pile
4. Determine the **loss** (difference in values)
5. The team with the smallest loss wins

## Thing to think about:

- What was your strategy for dividing the diamonds?
- Was your split fair?
- If given more time, is there another approach that you might take?

# Gradient Descent



**Think about questions you have about  
today's material.**

We will see you tomorrow at 1:30pm!