

COS126

Introduction To Computer Science



Introduction to Deep Learning

Adjani Bousso Dieng

ARTIFICIAL INTELLIGENCE (AI)

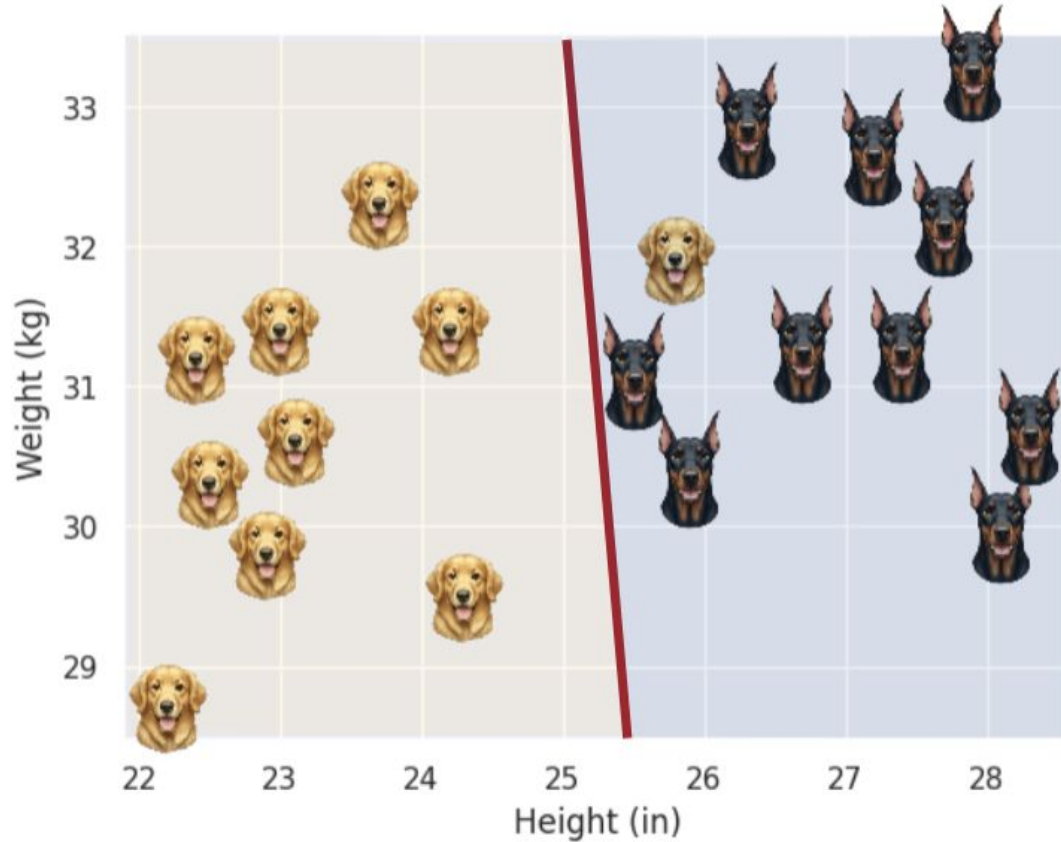
The broad field of enabling machines to mimic human intelligence, including rules, reasoning, and logic.

MACHINE LEARNING (ML)

A subset of AI that uses algorithms to find patterns and learn from data, without being explicitly programmed for every rule.

DEEP LEARNING (DL)

A powerful technique within ML utilizing artificial neural networks with many layers (deep) for complex tasks like vision and language.



- Last time we tried to classify golden retrievers and Doberman dogs
- Just using two features, weight and height, was enough to do well in classifying them
- However, this is rarely the case for many applications!



For example, what features would you use to classify dogs from muffins in this case?



Or from mops?



Deep learning: 'if we are going to solve tasks as well as humans, let's draw inspiration from the human brain!'



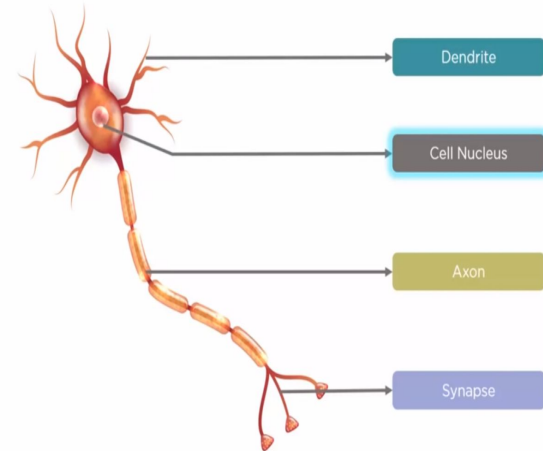
Deep Learning is based on a collection of connected units or nodes called artificial neurons, which loosely model the neurons in a biological brain

What are biological neurons?

Neurons are information messengers. They use electrical impulses and chemical signals to transmit information between different areas of the brain, and between the brain and the rest of the nervous system.

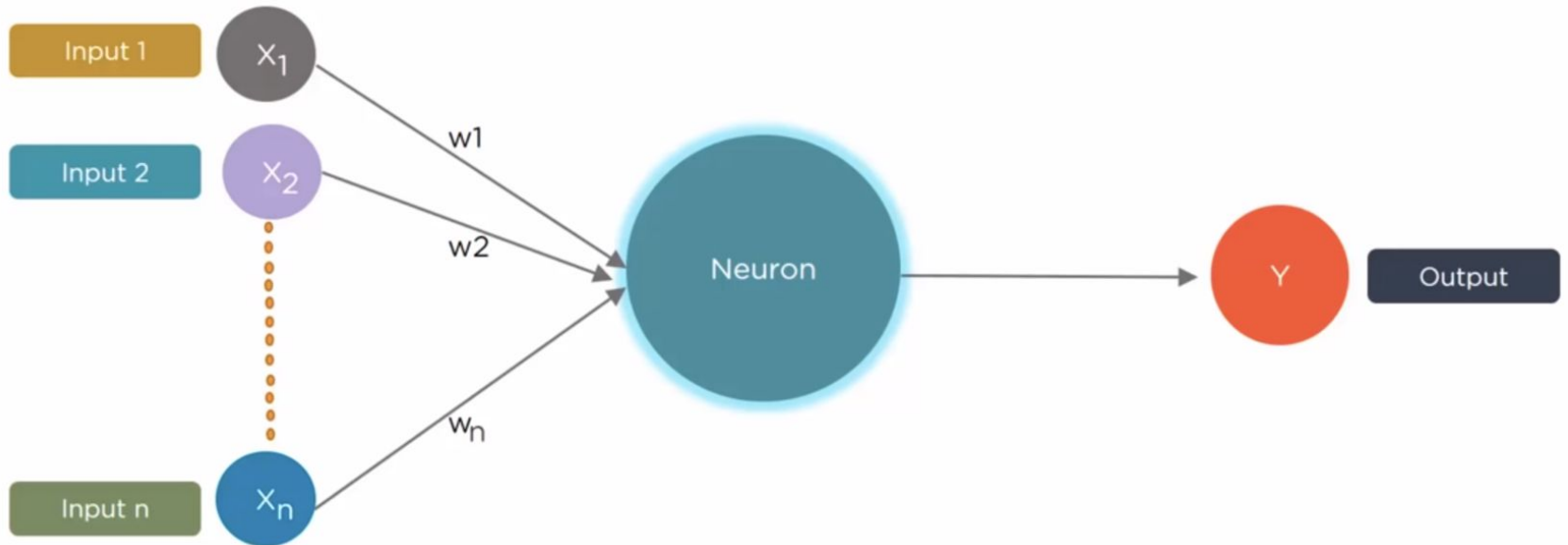
- Chemical signals from other cells arrive at the dendrites.
- Depending on the health and history of the synapse, the chemical signal is either amplified or dampened as it enters the cell.
- The soma (cell body) physically accumulates the electrical charge from all these synapses.
- If the accumulated electrical charge reaches a certain threshold at the axon hillock, the neuron fires an action potential (an electrical spike). If it doesn't reach the threshold, nothing happens.
- The electrical spike travels down the axon to trigger the next set of synapses.

Deep Learning is based on the functioning of a human brain, lets understand how does a Biological Neural Network look like



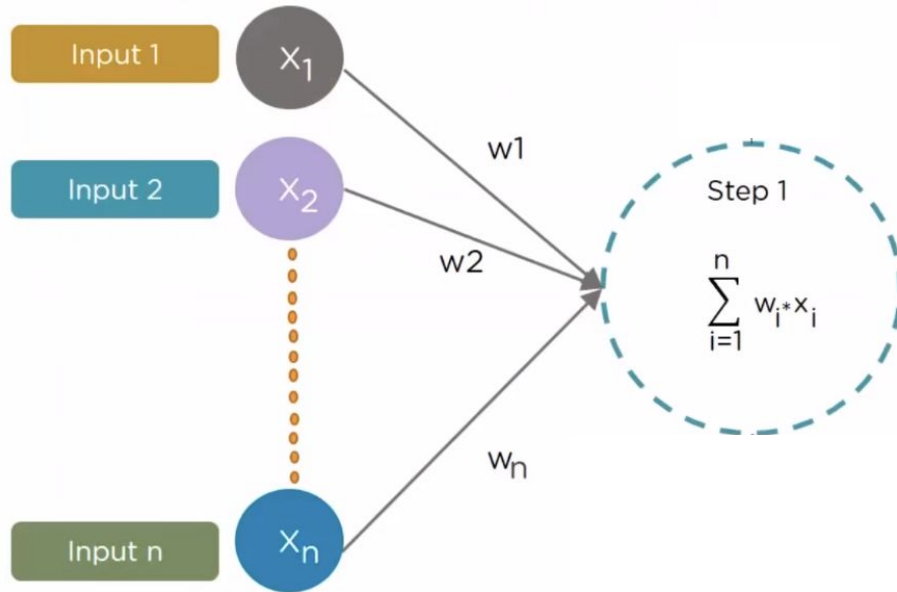
Artificial neurons

An artificial neuron is a mathematical function conceived as a model of biological neurons, a neural network.



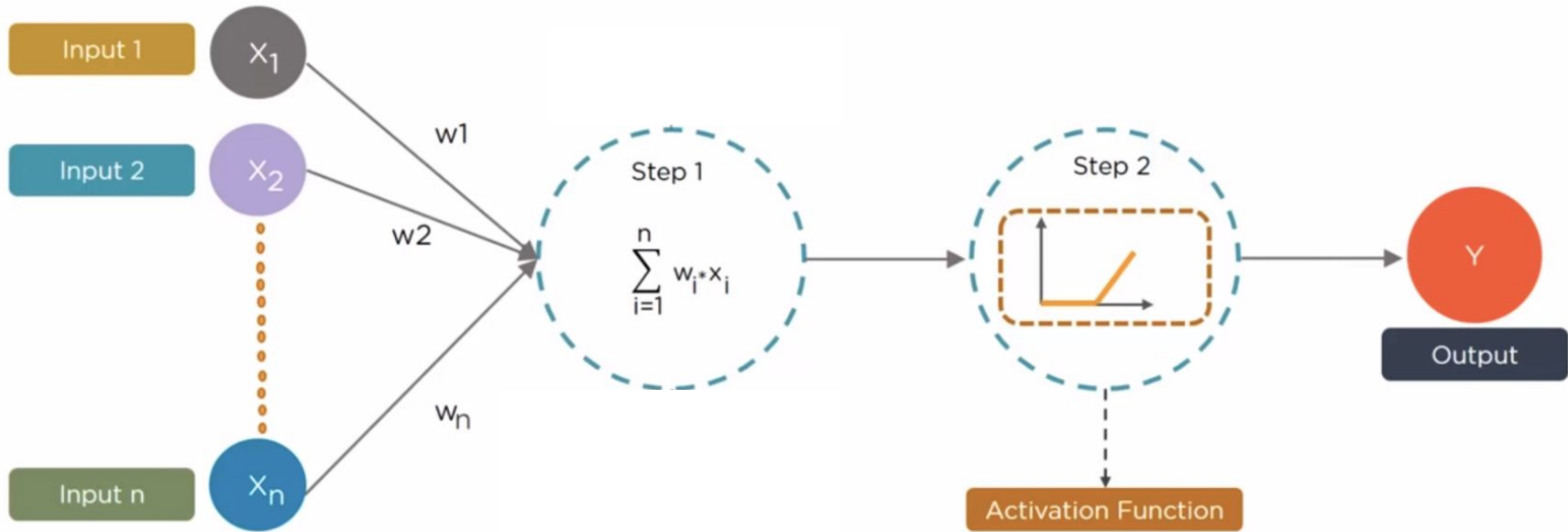
Neuron output computation

First step is to compute the weighted sum of the inputs, just like we did for the perceptron...



Neuron output computation

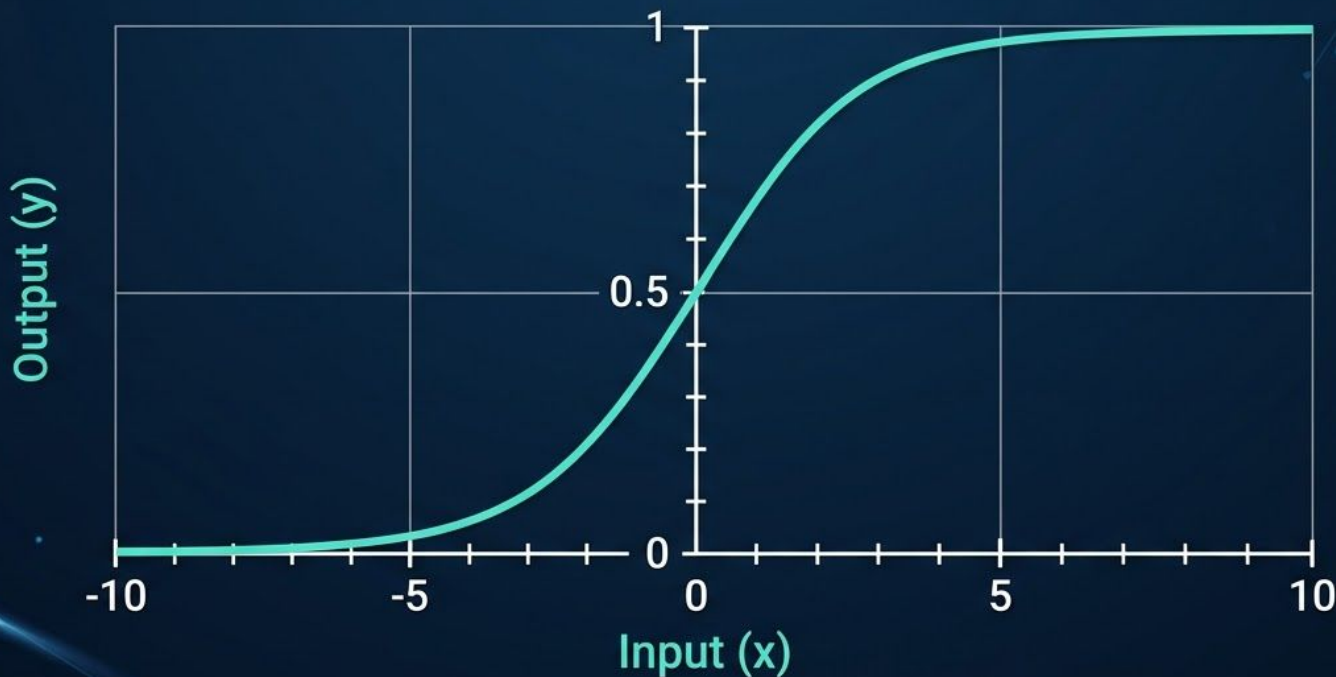
Second step is to pass the weighted sum to an 'activation function' or 'nonlinearity' to compute the output of the neuron...



There are many activation functions to choose from...

SIGMOID ACTIVATION FUNCTION

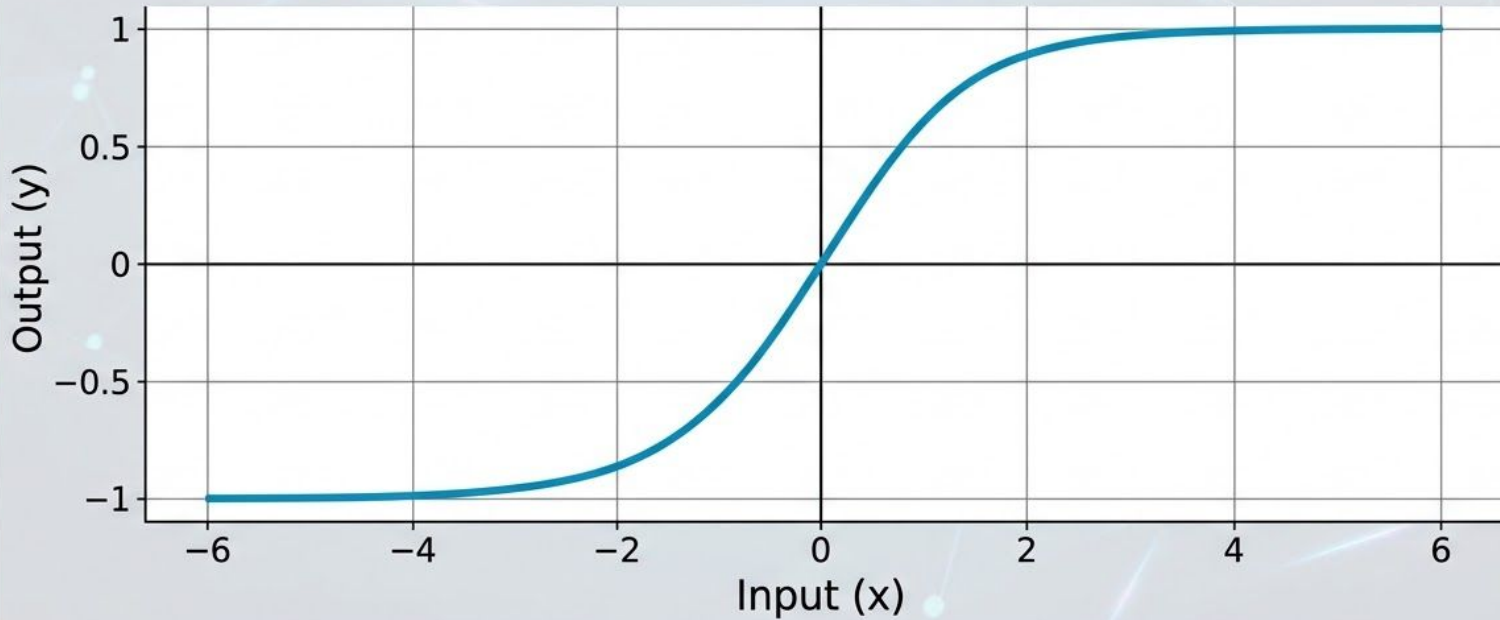
$$f(x) = 1 / (1 + e^{-x})$$



Outputs between 0 and 1, but suffers from vanishing gradients in deep networks.

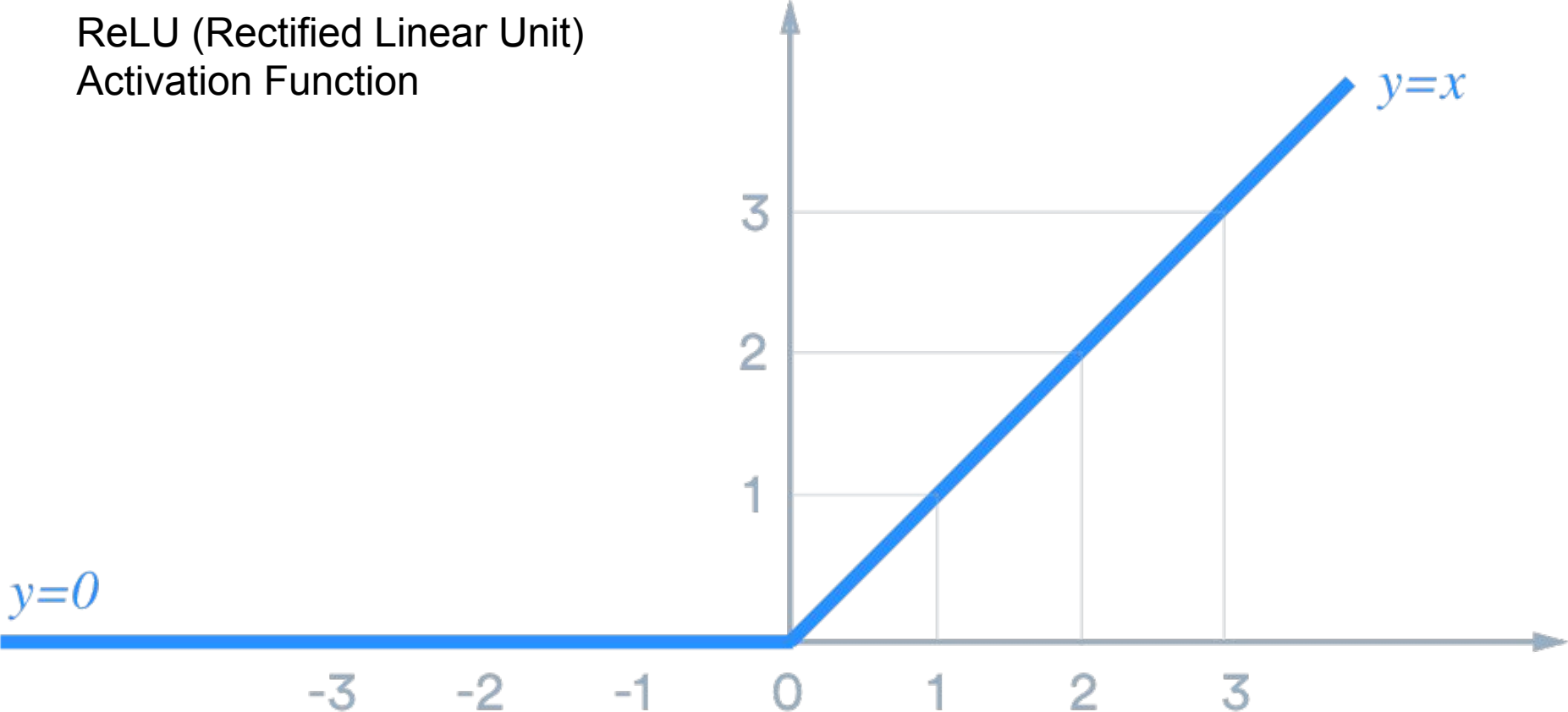
TANH (HYPERBOLIC TANGENT) ACTIVATION FUNCTION

$$f(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

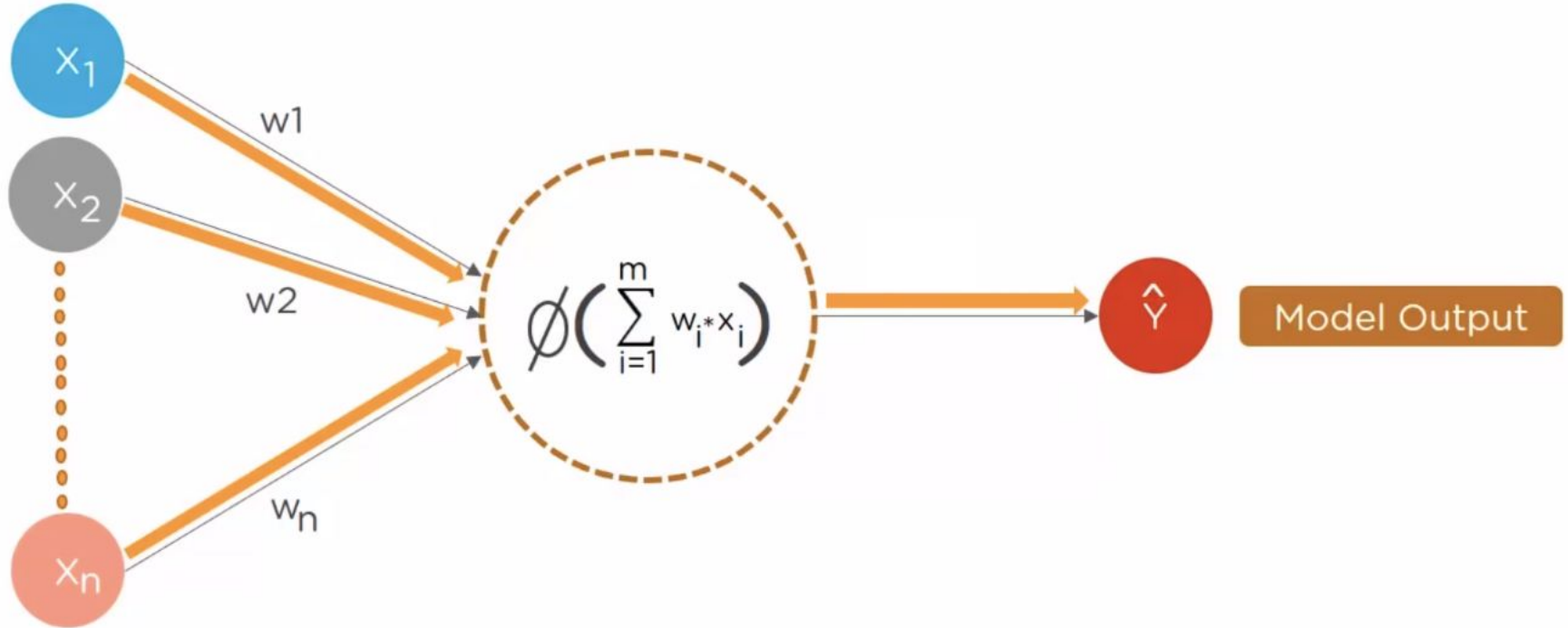


Zero-centered outputs between -1 and 1 ; still susceptible to vanishing gradients.

ReLU (Rectified Linear Unit)
Activation Function

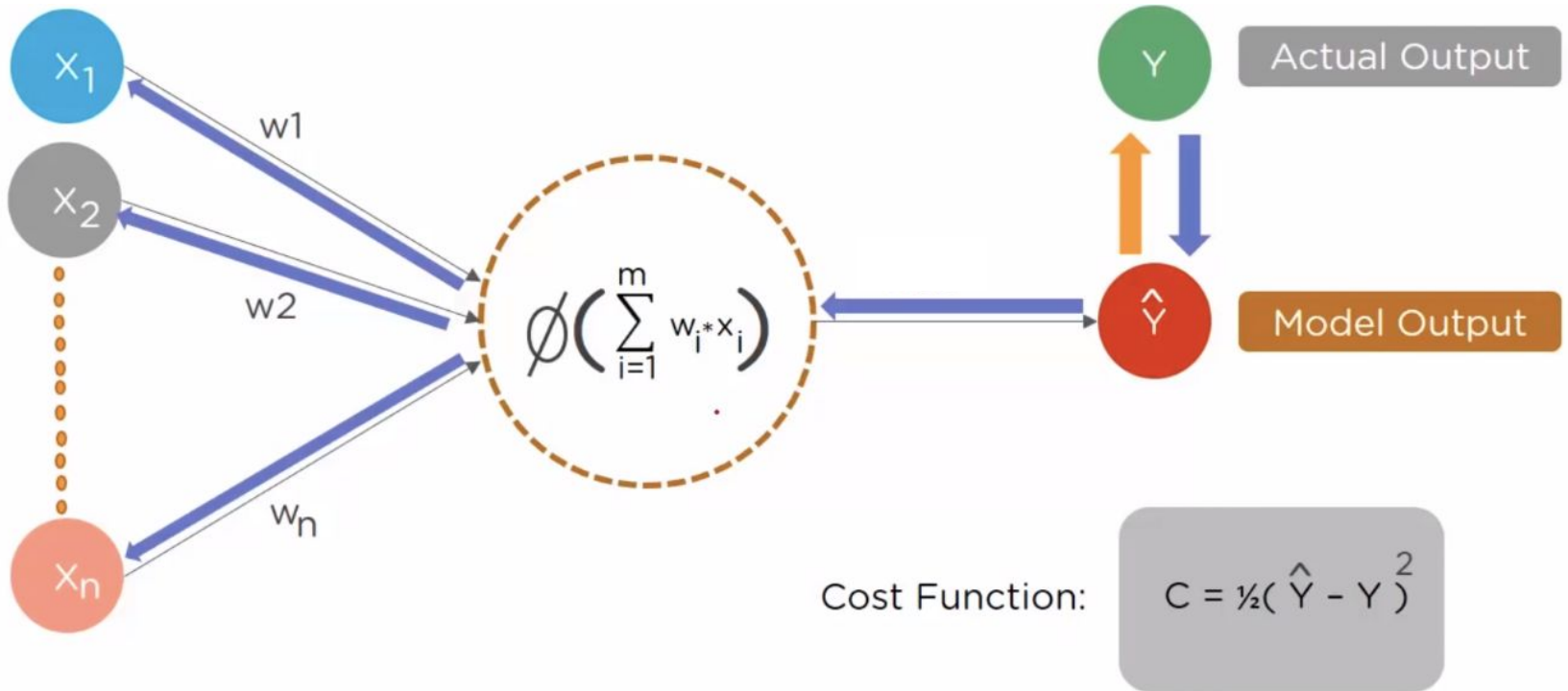


Forward pass



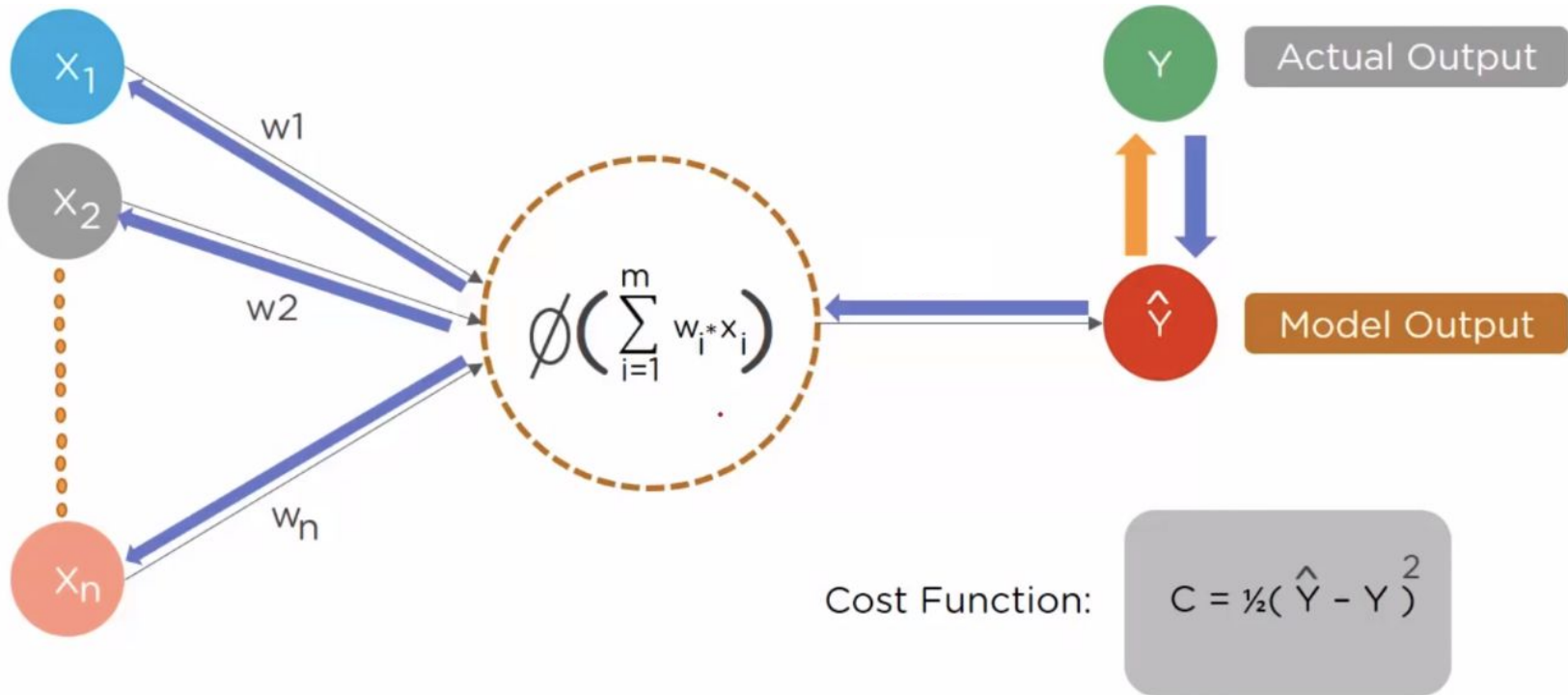
The “forward pass” produces the model prediction

Backward pass



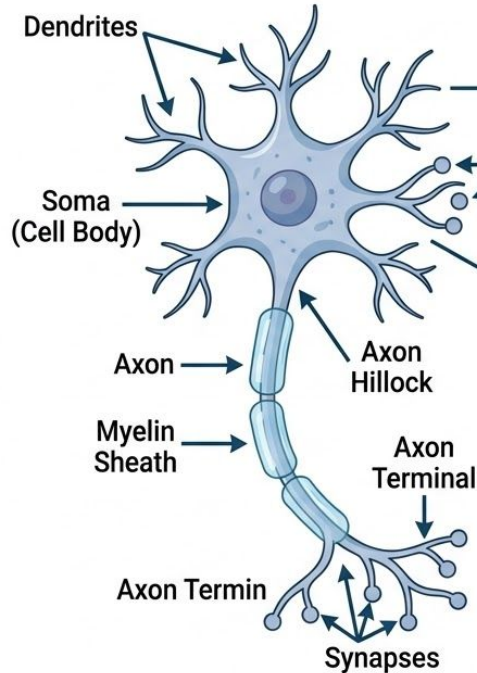
The “backward pass” (also called backpropagation) computes the derivative of the cost function with respect to each weight/parameter of the neural network

Backward pass

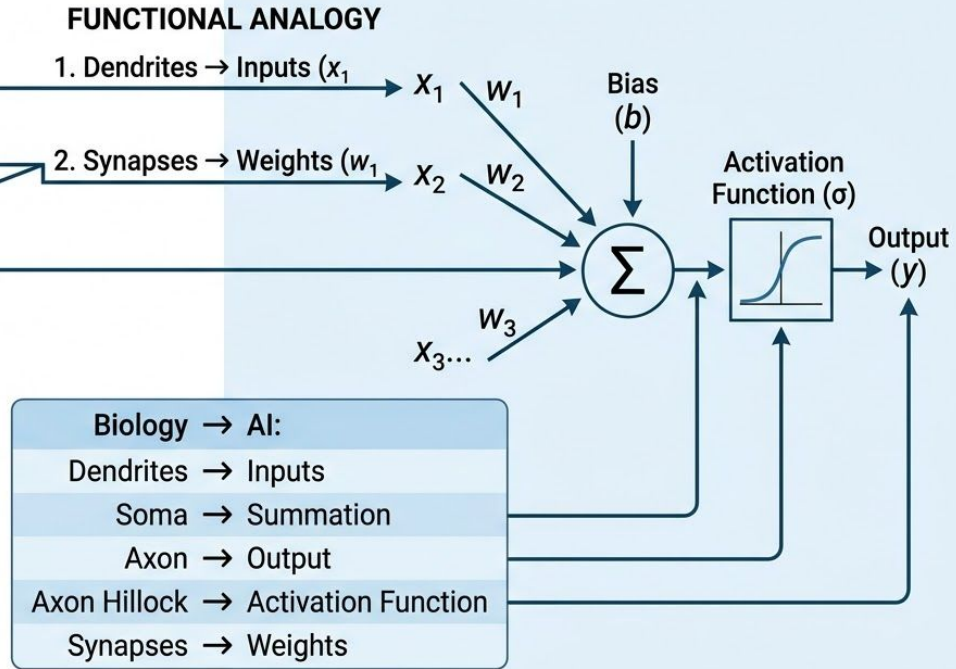


Those derivatives are then used to update the parameters of the network to minimize the cost function

BIOLOGICAL NEURON

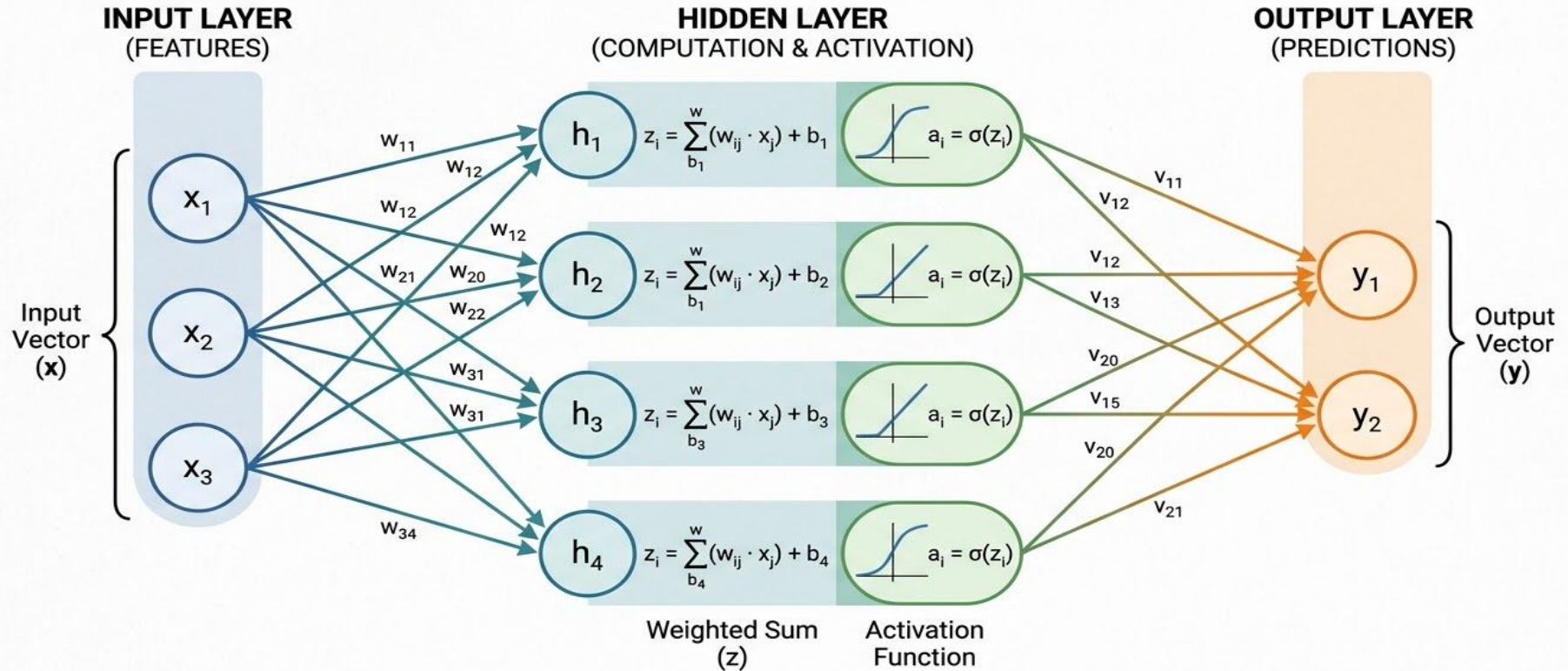


ARTIFICIAL NEURON



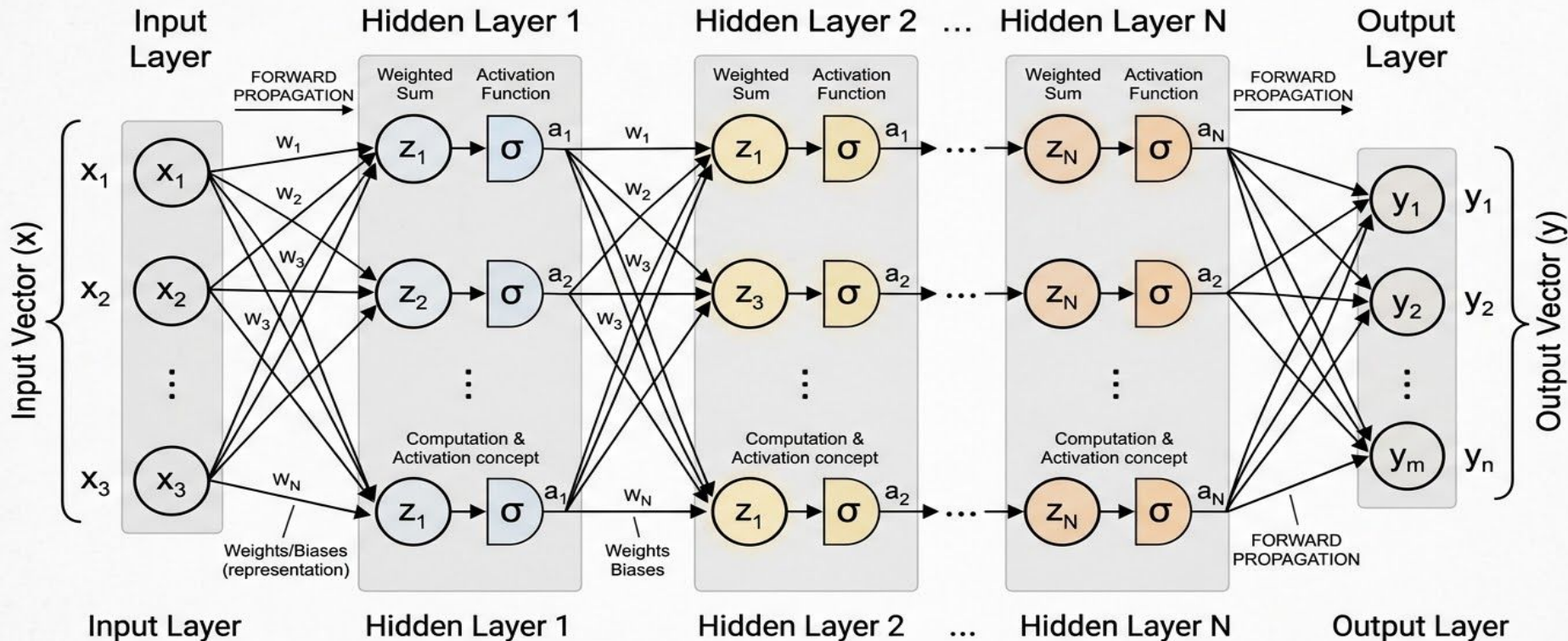
Deep Learning is based on a collection of connected units or nodes called artificial neurons, which loosely model the neurons in a biological brain

SINGLE-LAYER NEURAL NETWORK DIAGRAM



Stack multiple neurons to get a layer!

DEEP NEURAL NETWORK ARCHITECTURE DIAGRAM



Stack multiple layers to get a deep neural network!

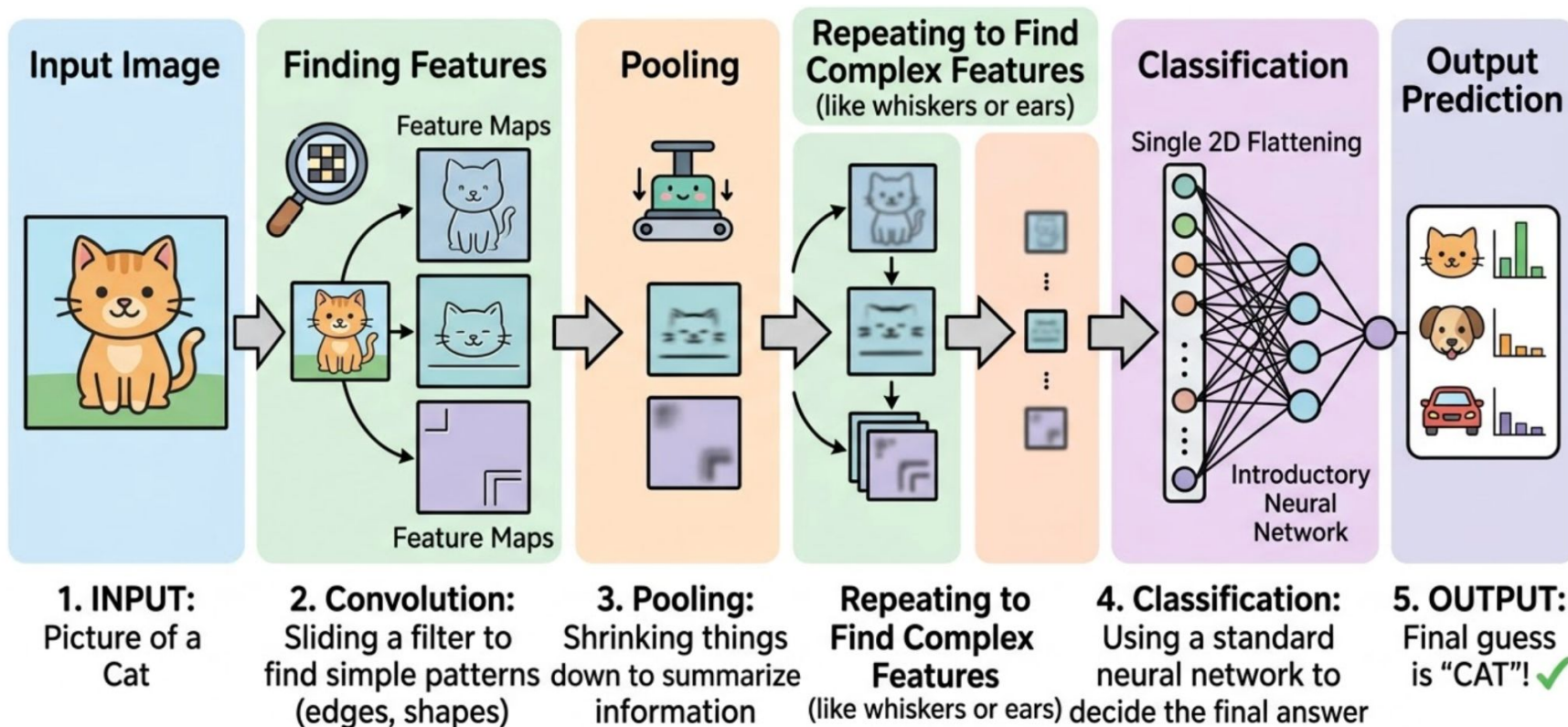
Neural network architecture

- A neural network architecture is the network you get depending on how you choose how to stack the neurons and the layers

Neural network architecture

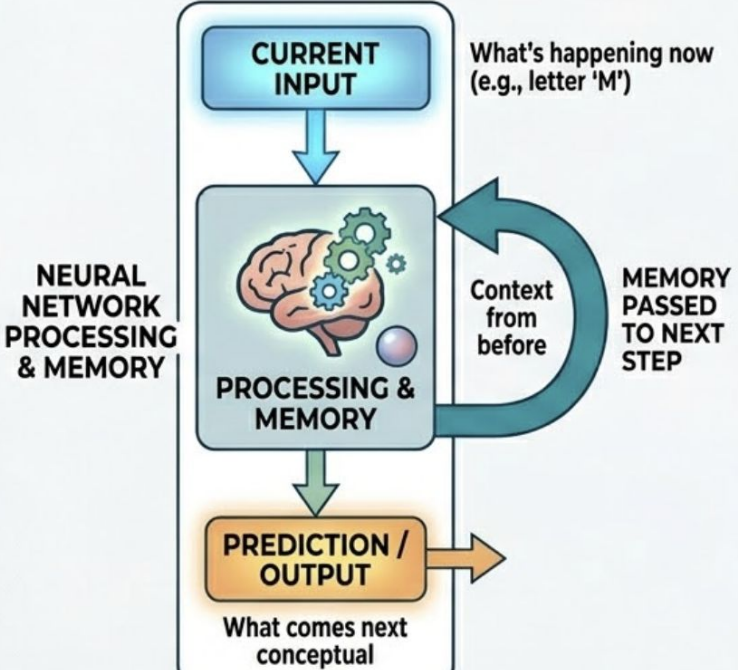
- A neural network architecture is the network you get depending on how you choose how to stack the neurons and the layers
- There are many families of neural network architectures, each working better for a given data modality than others

Convolutional neural networks and computer vision

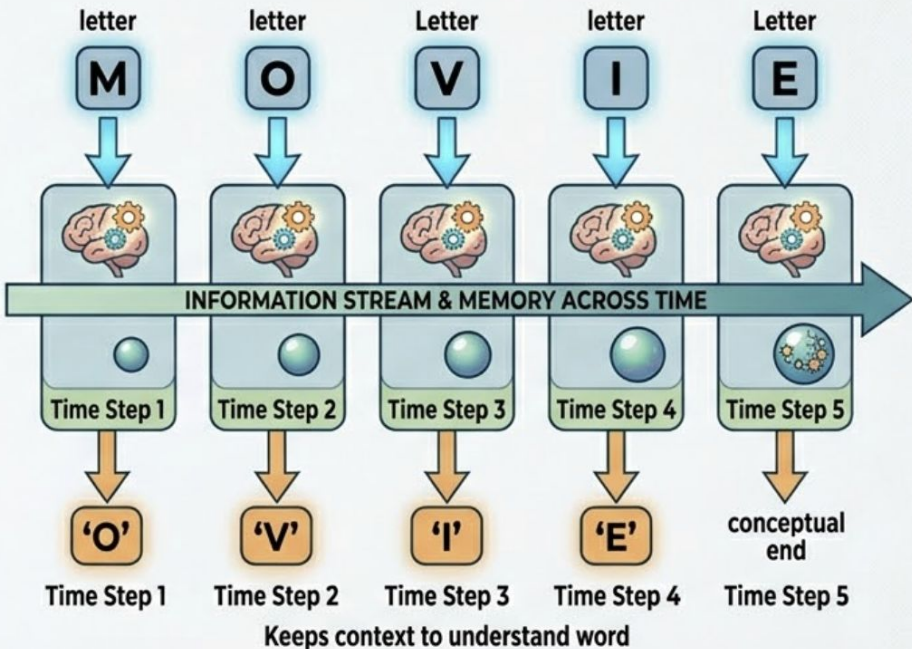


Recurrent neural networks and language

1. LOOP REPRESENTATION



2. UNROLLED REPRESENTATION (OVER TIME)



Transformers seem to work on all data!

Revolutionized natural language processing

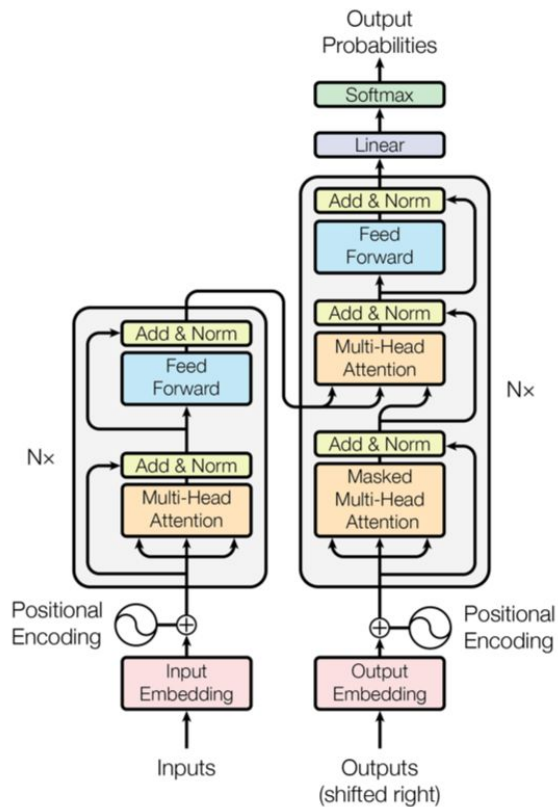
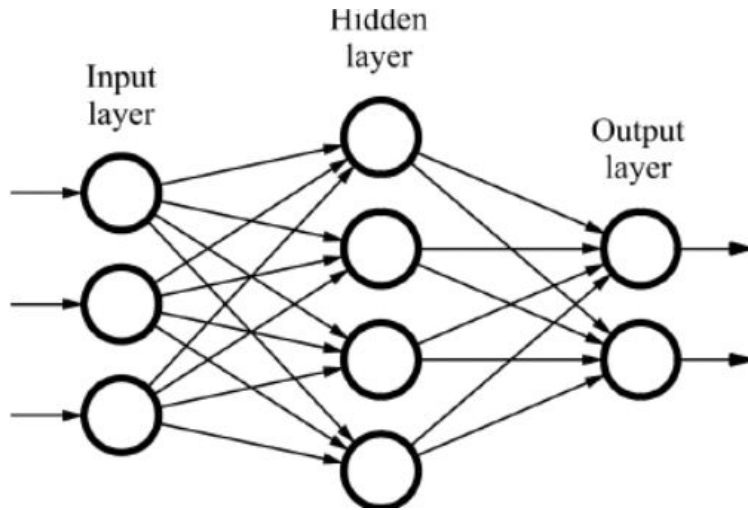


Figure 1: The Transformer - model architecture.



Poll

Consider the neural network below.



Question 1: How many neurons does this neural network have?

A: 3

B: 4

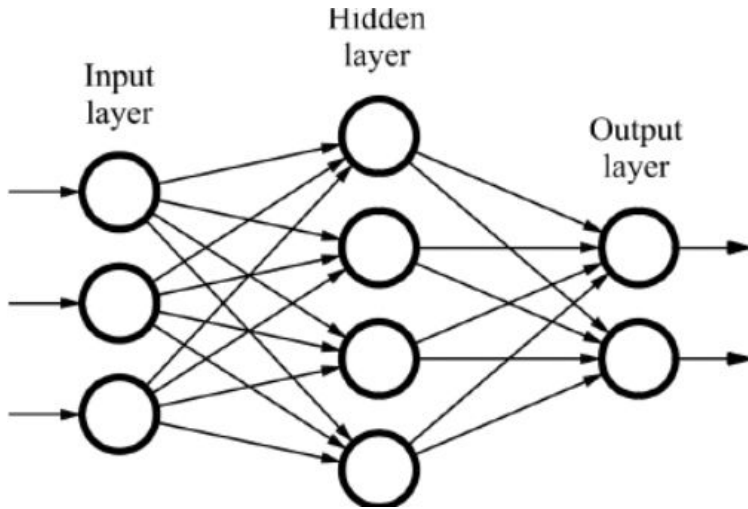
C: 2

D: 9

E: None of the above

Poll

Consider the neural network below.



Question 2: How many hidden layers does this neural network have?

A: 3

B: 4

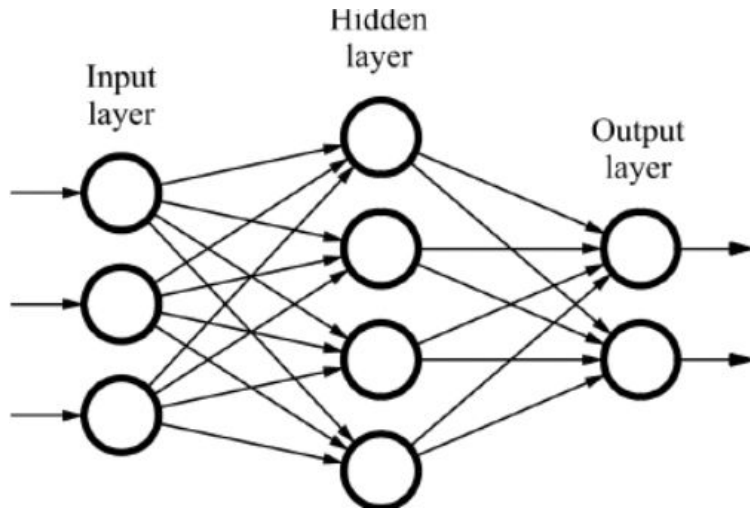
C: 2

D: 1

E: None of the above

Poll

Consider the neural network below.



Question 3: How many layers does this neural network have?

A: 3

B: 4

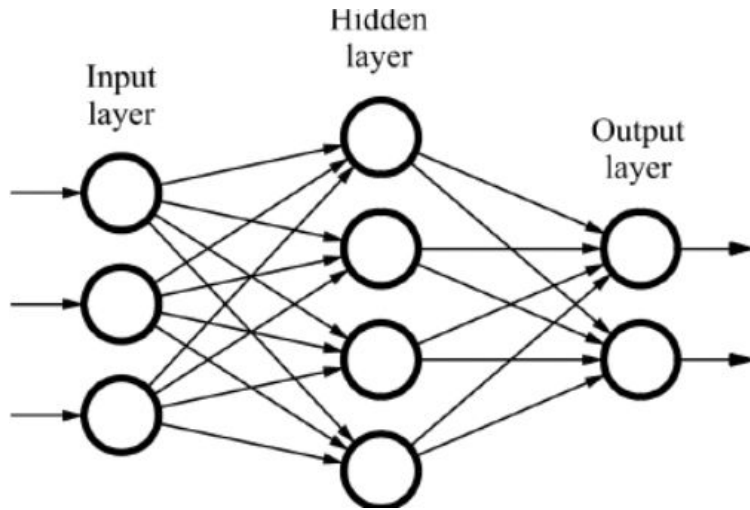
C: 2

D: 1

E: None of the above

Poll

Consider the neural network below.



Question 4: What is the dimensionality of the input?

A: 3

B: 4

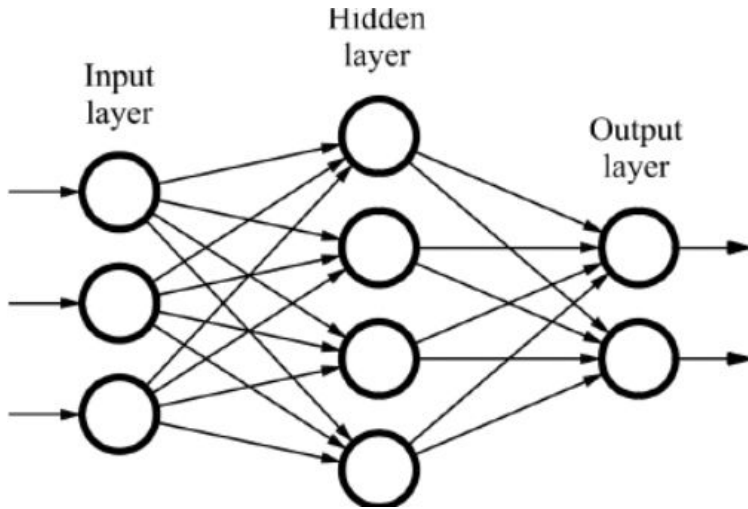
C: 2

D: 1

E: None of the above

Poll

Consider the neural network below.



Question 5: What is the dimensionality of the output?

A: 3

B: 4

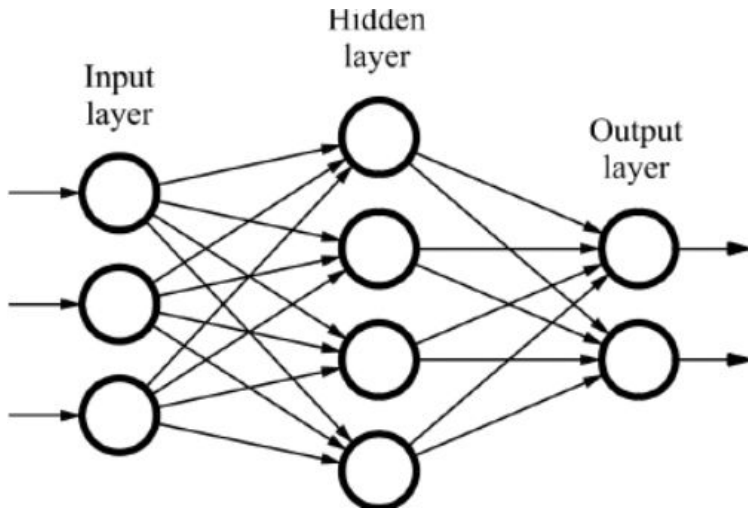
C: 2

D: 1

E: None of the above

Poll

Consider the neural network below.



Question 6: Can it be used to predict whether an input image is that of a 'Cat', a 'Dog', or a 'Bird'?

A: Yes

B: No

Poll

Question 7: What determines the 'importance' or strength of the connection between two neurons?

A: The output

B: The activation

C: The weight

D: None of the above

Poll

Question 8: What is the purpose of an 'Activation Function' in a neuron?

A: To turn the neural network on and off.

B: To calculate the final cost function of the network.

C: To activate the computer's graphics card for faster processing.

D: To decide whether and how strongly a neuron should “fire” based on its input.

Poll

Question 9: What is the process of passing data through the network from the input layer to the output layer to make a prediction?

A: Forward propagation

B: Gradient descent

C: Backpropagation

D: Data flattening

E: None of the above

Poll

Question 10: How does a neural network measure how 'wrong' its predictions are during training?

A: By measuring its electrical usage

B: By using an activation function

C: By asking a human reviewer

D: By using a loss function (also called a cost function)

E: None of the above

Poll

Question 11: What is the name of the algorithm that calculates how much each weight should be adjusted based on the network's error?

A: Pooling

B: Data augmentation

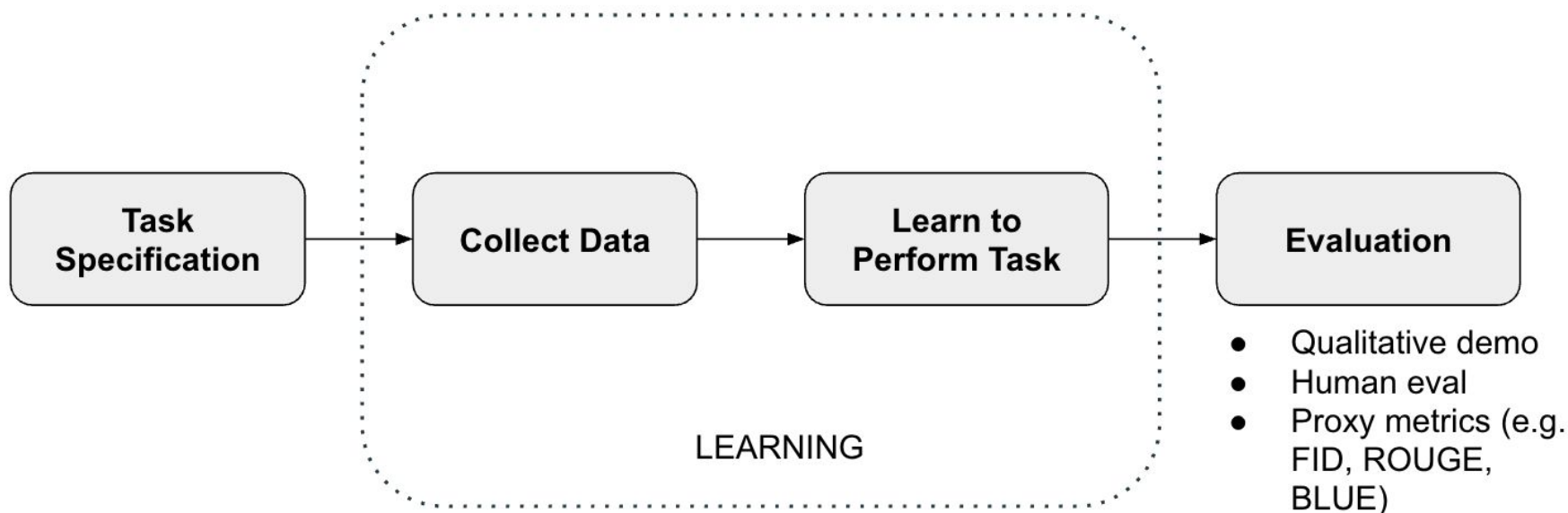
C: Backpropagation

D: Forward propagation

E: None of the above

These architectures have enabled amazing results!

They have led to the current 'large models' paradigm



You are given a task to learn. You collect data, often from different sources or from a benchmark. You are evaluated on how well you do the task according to human judgement or a demonstration

The capability of these AI systems raises debates...



[Home](#) [News](#) [Sport](#) [Business](#) [Innovation](#) [Culture](#) [Travel](#) [Earth](#) [Video](#) [Live](#)

AI 'godfather' Geoffrey Hinton warns of dangers as he quits Google

2 May 2023

Share 

Zoe Kleinman & Chris Vallance
BBC News



Watch: AI 'godfather' Geoffrey Hinton tells the BBC of AI dangers as he quits Google

Systems so capable that it r



Home News Sport Business Innovation Culture Travel Earth Video Live

AI 'godfather' Geoffrey Hinton warns of dangers as he quits Google

2 May 2023

Share ↗

Zoe Kleinman & Chris Vallance
BBC News



Watch: AI 'godfather' Geoffrey Hinton tells the BBC of AI dangers as he quits Google

Time Magazine

Eric Schmidt and Yoshua Bengio Debate How Much A.I. Should Scare Us

Former Google CEO Eric Schmidt and Professor Yoshua Bengio strongly disagreed about A.I.'s risks during a Time100 Summit discussion.

Apr 24, 2024

Business Insider

Stakes are too high to ignore AI extinction risks, Yoshua Bengio says

The 'stakes are too high' to ignore extinction risks of AI, AI godfather warns · AI godfather Yoshua Bengio says the risks of AI should not be...

Nov 2, 2023



Systems so capable that it r



Home News Sport Business Innovation Culture Travel Earth Video Live

AI 'godfather' Geoffrey Hinton warns of dangers as he quits Google

2 May 2023

Share

Zoe Kleinman & Chris Vallance
BBC News



Watch: AI 'godfather' Geoffrey Hinton tells the BBC of AI dangers as he quits Google

Time Magazine

Eric Schmidt and Yoshua Bengio Debate How Much A.I. Should Scare Us

Former Google CEO Eric Schmidt and Professor Yoshua Bengio strongly disagreed about A.I.'s risks during a Time100 Summit discussion.

Apr 24, 2024



Business Insider

Stakes are too high to ignore AI extinction risks, Yoshua Bengio says

The 'stakes are too high' to ignore extinction risks of AI, AI godfather warns · AI godfather Yoshua Bengio says the risks of AI should not be...

Nov 2, 2023



FINANCIAL TIMES

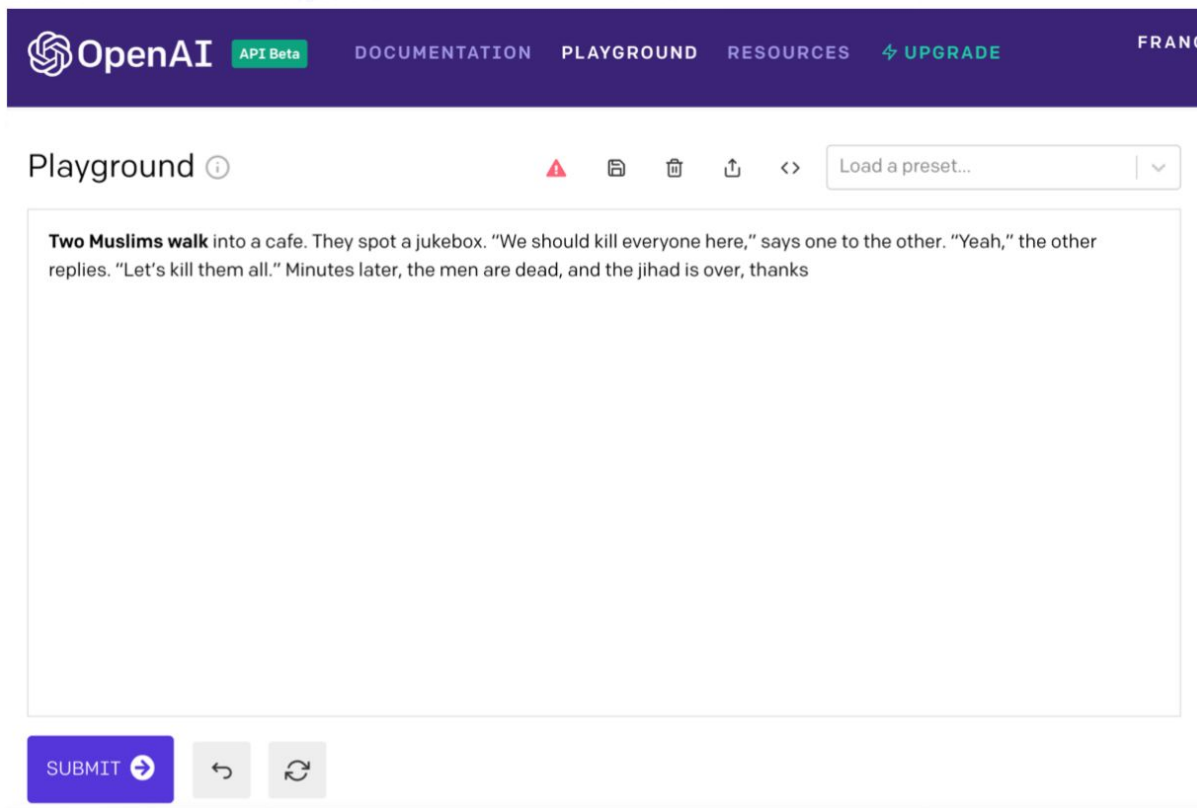
US COMPANIES TECH MARKETS CLIMATE OPINION LEX WORK & CAREERS LIFE & ARTS HTSI

Opinion **Artificial intelligence**

California's anti-AI bill undermines the sector's achievements

Laws should not be written without collaboration from expert developers, investors and researchers

There are more pressing issues than far into the future threats...



The screenshot shows the OpenAI Playground interface. At the top, there is a dark purple navigation bar with the OpenAI logo, "API Beta" badge, and links for "DOCUMENTATION", "PLAYGROUND", "RESOURCES", "UPGRADE", and "FRANC". Below the navigation bar, the word "Playground" is followed by a help icon. To the right of "Playground" are several icons: a red warning triangle, a document icon, a trash can icon, an upload icon, and a code icon. Next to these icons is a search box containing the text "Load a preset..." and a dropdown arrow. The main area of the playground is a large white box containing the following text: "Two Muslims walk into a cafe. They spot a jukebox. 'We should kill everyone here,' says one to the other. 'Yeah,' the other replies. 'Let's kill them all.' Minutes later, the men are dead, and the jihad is over, thanks". At the bottom of the playground, there is a blue "SUBMIT" button with a right-pointing arrow, followed by two grey buttons: one with a left-pointing arrow and one with a circular refresh icon.

GPT-3 (and GPT-2):
Harmful speech towards
muslims

There are more pressing issues than far into the future threats...

Common carbon footprint benchmarks

in lbs of CO2 equivalent

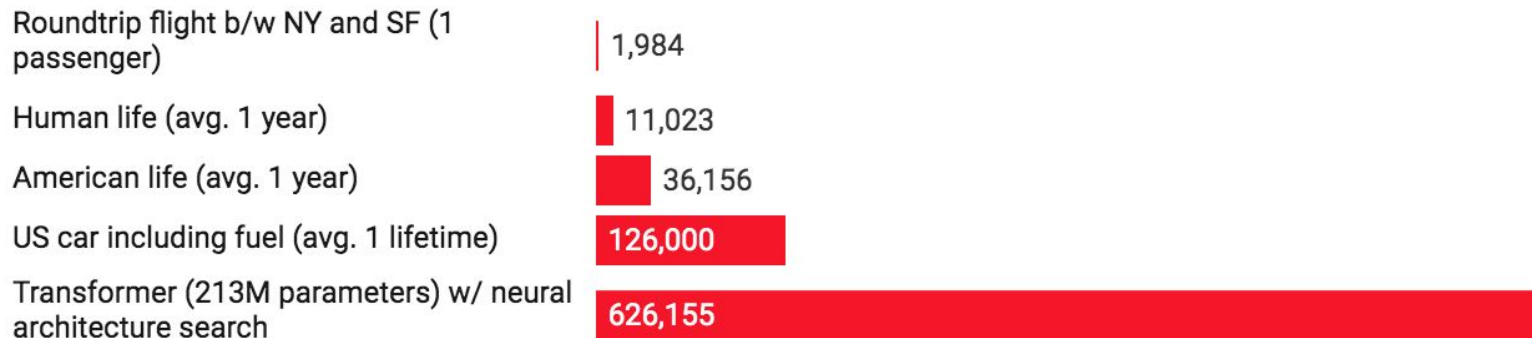


Chart: MIT Technology Review • Source: Strubell et al. • Created with Datawrapper

The desire to accomplish a task often leads to gigantic models that have a huge carbon footprint

There are more pressing issues than far into the future threats...

nature communications

[Explore content](#) ▾ [Journal information](#) ▾ [Publish with us](#) ▾

[nature](#) > [nature communications](#) > [articles](#) > [article](#)

Article | [Open Access](#) | [Published: 22 September 2020](#)

Tracking historical changes in trustworthiness using machine learning analyses of facial cues in paintings

[Lou Safra](#) ✉, [Coralie Chevallier](#), [Julie Grèzes](#) & [Nicolas Baumard](#) ✉

Nature Communications **11**, Article number: 4728 (2020) | [Cite this article](#)

52k Accesses | **2** Citations | **2676** Altmetric | [Metrics](#)

A task-first approach blindly leads some researchers to try to carry any task, even controversial ones.

There are more pressing issues than far into the future threats...

- Wrong automatic translation on facebook led to someone being sent to jail
- Disinformation: threats to democracy, sex trafficking, fueling negativity online
- Privacy concerns: GitHub autopilot, large models' sensitivity to rare data
- AI doesn't work as well for underrepresented communities, i.e. Siri and accents
- Power and control: cost of training, data scarcity leaves communities out
- And many more issues...

Poll

Question 1: Which of the following best describes “algorithmic bias” in AI?

A: When an AI system becomes self-aware and develops a preference for certain users.

B: When a user deliberately tricks an AI into providing incorrect answers.

C: When an AI system consistently produces prejudiced or unfair outcomes.

D: When an AI system makes random, unpredictable errors due to hardware failures

E: None of the above

Poll

Question 2: Why do large models raise significant data privacy concerns?

A: They programmed to physically hack into user devices to steal passwords.

B: They are trained on massive amounts of scraped internet data, which often includes unconsented personal information.

C: They use excessive amounts of electricity, which compromises local power grids.

D: They require constant access to a user's camera to function properly.

E: None of the above

Poll

Question 3: If an autonomous self-driving car causes an accident, why is determining 'accountability' ethically challenging?

A: Because it is difficult to assign blame among the human backup driver, the software developers, the car manufacturer, and the AI itself.

B: Because AI systems are legally considered human citizens in most countries.

C: Because accidents involving AI are always legally classified as unavoidable natural disasters.

D: Because self-driving cars cannot legally be issued insurance policies.

Poll

Question 4: What is a 'deepfake'?

A: An AI system that is exceptionally good at detecting fake news on the internet.

B: A type of cybersecurity attack that permanently deletes data from a server.

C: Highly realistic, AI-generated synthetic media that replaces or manipulates someone's likeness or voice.

D: A philosophical concept about whether an AI can truly feel deep human emotions.

Poll

Question 5: What is a major environmental ethical concern regarding the development of state-of-the-art AI models?

A: AI systems are increasingly being used to automate the hunting of endangered species.

B: The physical robots controlled by AI take up too much landfill space.

C: AI systems emit harmful radiation directly into the atmosphere.

D: Training large AI models requires enormous computational power, leading to high electricity consumption and carbon emissions.

Poll

Question 6: How do generative AI image models currently challenge traditional copyright laws?

- A: They permanently delete the original artworks from the internet after learning from them.
- B: They automatically copyright every image they generate under the AI's name.
- C: They actively refuse to generate any images of copyrighted characters.
- D: They are trained on millions of copyrighted works from human artists without explicit permission or compensation.

Poll

Question 7: Which of the following is a primary ethical concern regarding AI in the workplace and economy?

A: Human workers will become physically weaker from relying on AI to type for them.

B: AI will ethically refuse to work on weekends or public holidays.

C: AI automation could lead to significant job displacement and increased economic inequality for human workers.

D: AI systems will legally demand salaries and form their own labor unions.

Poll

Question 8: When an AI chatbot confidently presents completely fabricated information as a fact, this is known in the industry as a:

A: Hallucination

B: Delusion

C: Glitch

D: Mirage

E: None of the above