CARTE-Enbridge Bootcamp

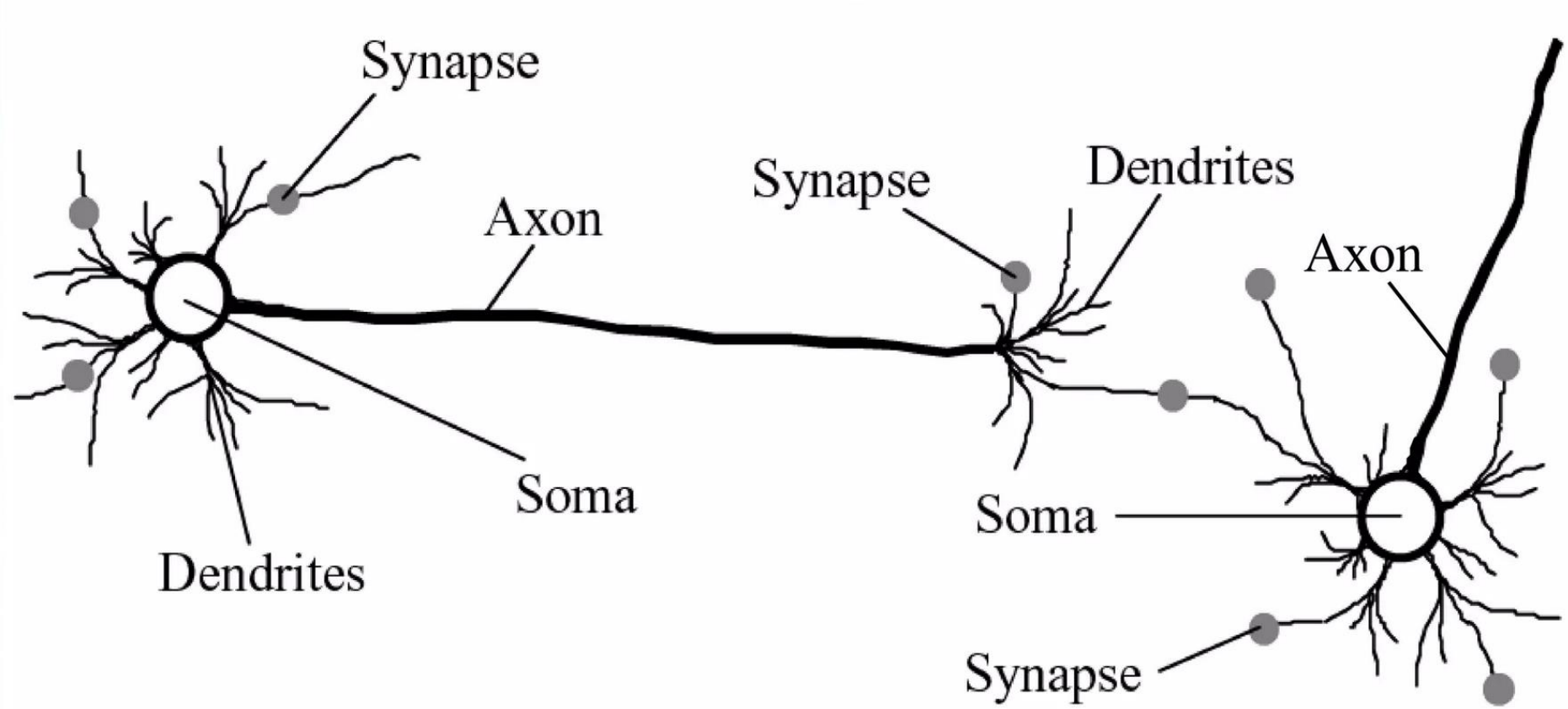
Basics of Neural Networks

UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING
Centre for Analytics and Artificial Intelligence Engineering

Neural Networks

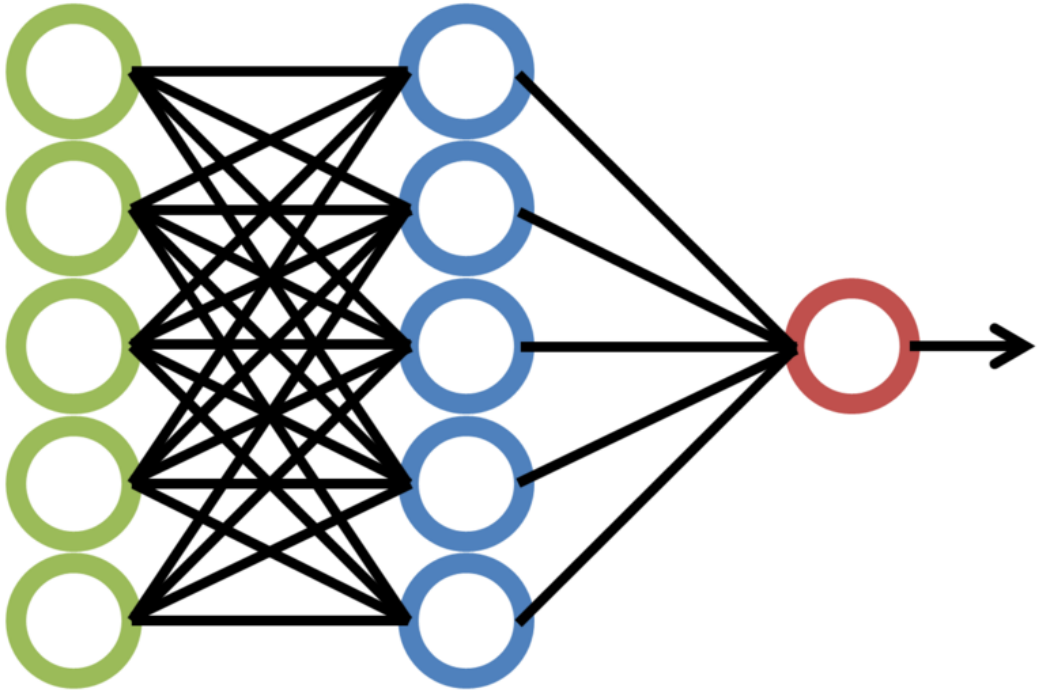
- Highly simplified model of a brain
- Brains consist of a densely connected set of nerve cells, or basic information processing unit
- Key concept: while each neuron is simple, many of them together are vastly complex

Biological Neural Network

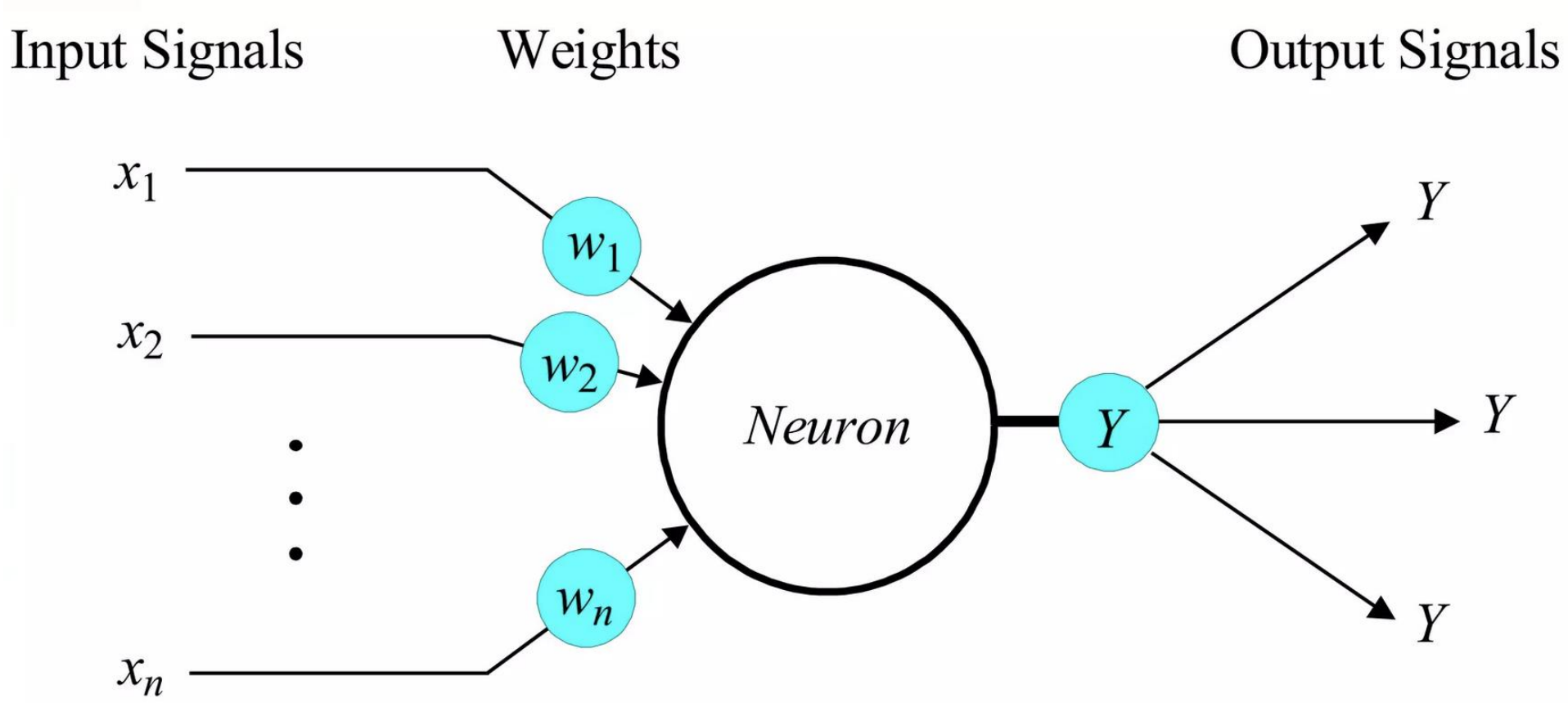


Artificial Neural Network (ANN)

- ANN also consists of numerous simple processors that are analogous to neurons in the brain
- Neurons are connected by weighted links passing signals in one direction

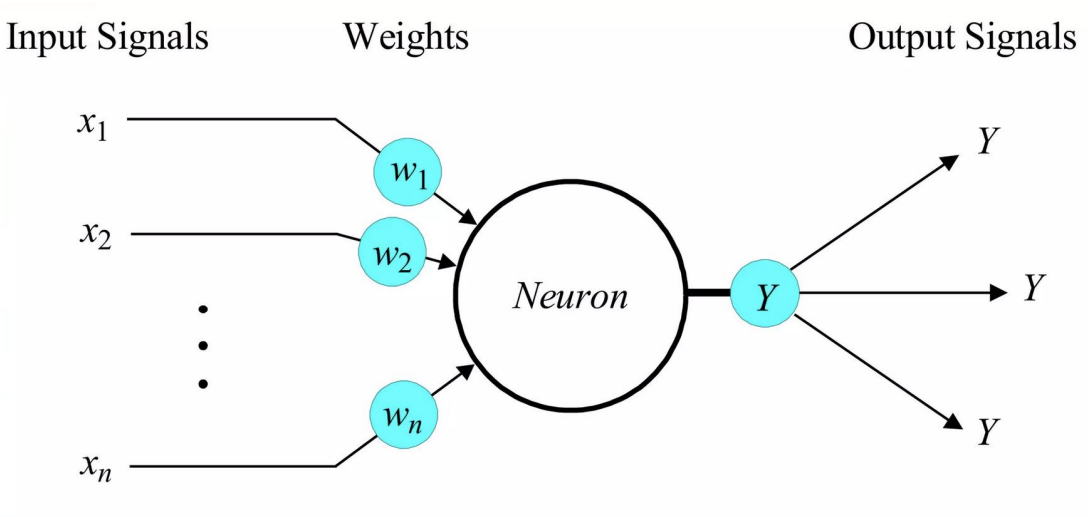


View inside an artificial neuron



View inside an artificial neuron

- Behaves like a linear regression model:
- $w_1 x_1 + w_2 x_2 + \ \ldots + w_n x_n$
- Weights correspond to how much the neuron “cares” about each input

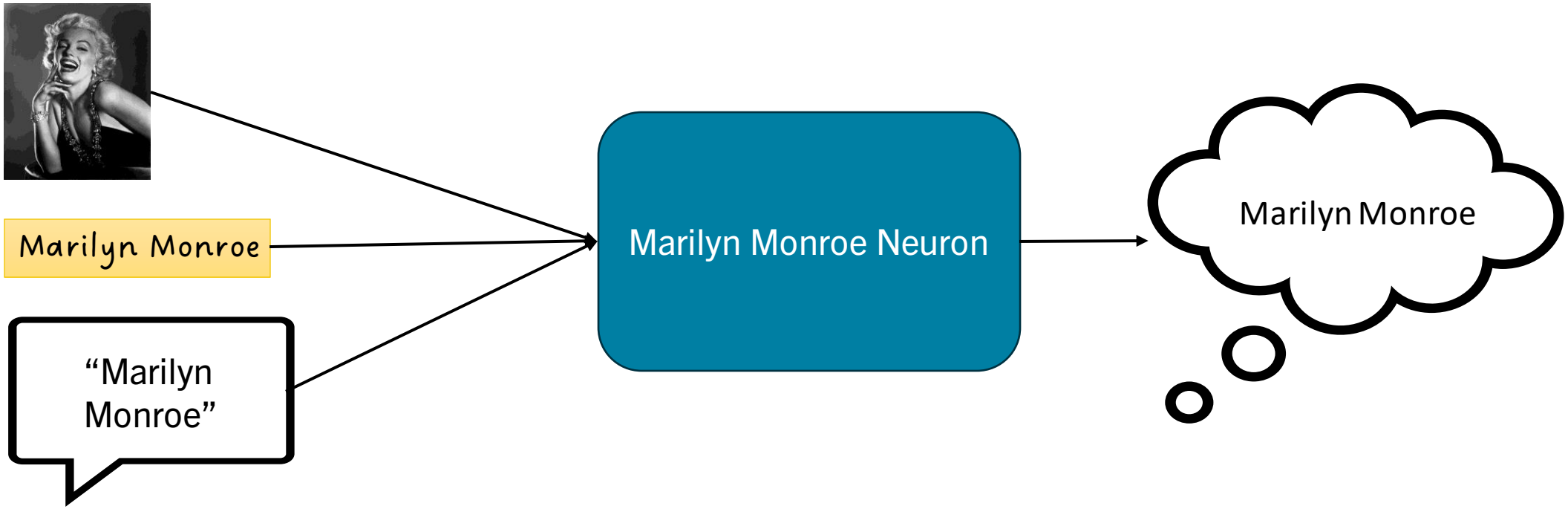


Back to the brain: the Marilyn Monroe Neuron

- Study conducted on patients with epilepsy
- Researchers use specialized equipment to measure the “excitement” of individual neurons in a patient’s brain
- Measuring a neuron, the researchers showed patients a series of images
- In each patient, they found around five neurons that fired when the patient looked at a specific person

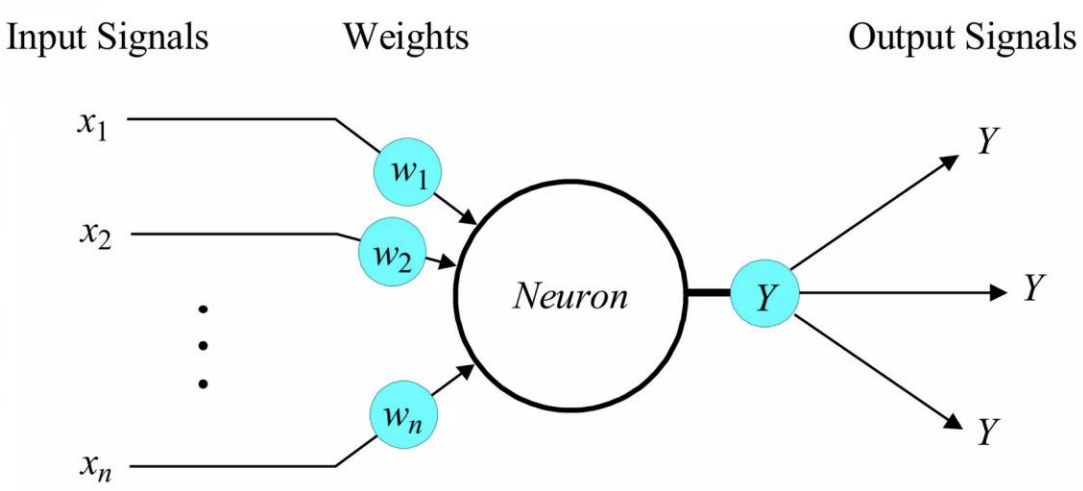
Back to the brain: the Marilyn Monroe Neuron

- Once a “celebrity” neuron was identified, the researchers wanted to know if it would still fire for representations other than images



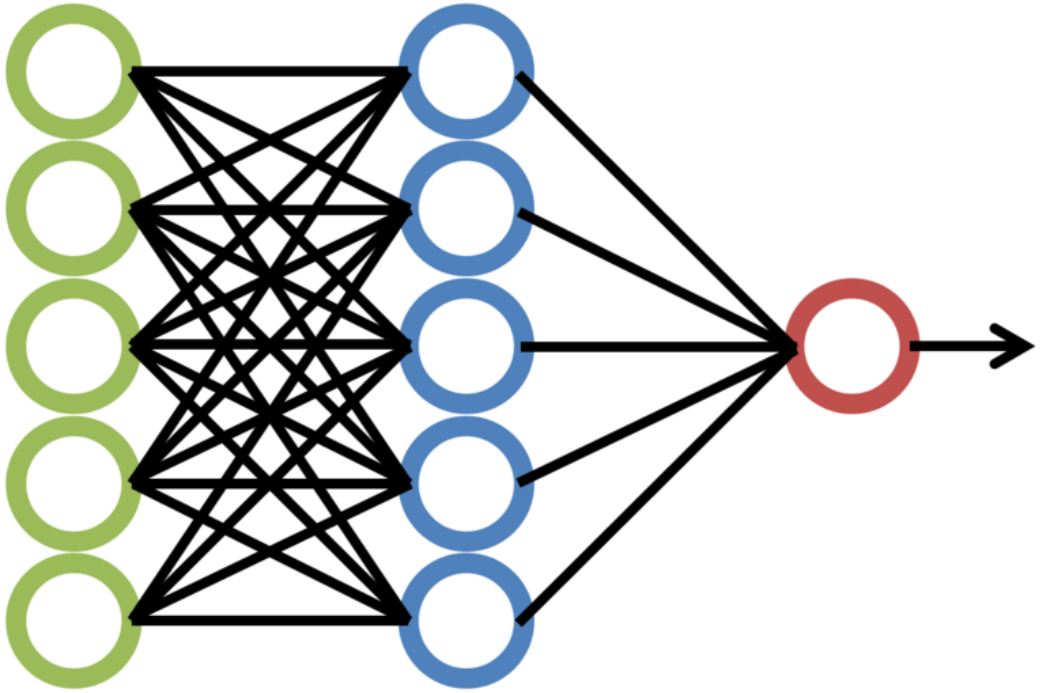
Marilyn Monroe ANN

- Weights would be high from neurons that react to different representations of Marilyn Monroe
- Weights would be low for neurons that react to other people, or concepts



ANNs

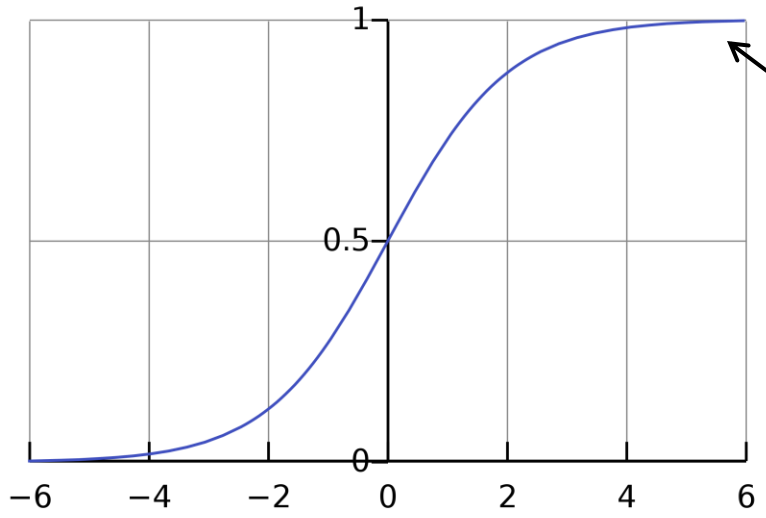
- Each neuron considers the responses of the neurons in the previous layer
- It learns to pay attention to the neurons that are excited about what it's excited about
- Ignores the neurons that are excited about other things



Activation Functions

- Basic approach: when I see enough activity, I get excited
- More useful: gradually increase excitement as we see more activity
- In practice: many different activation functions!

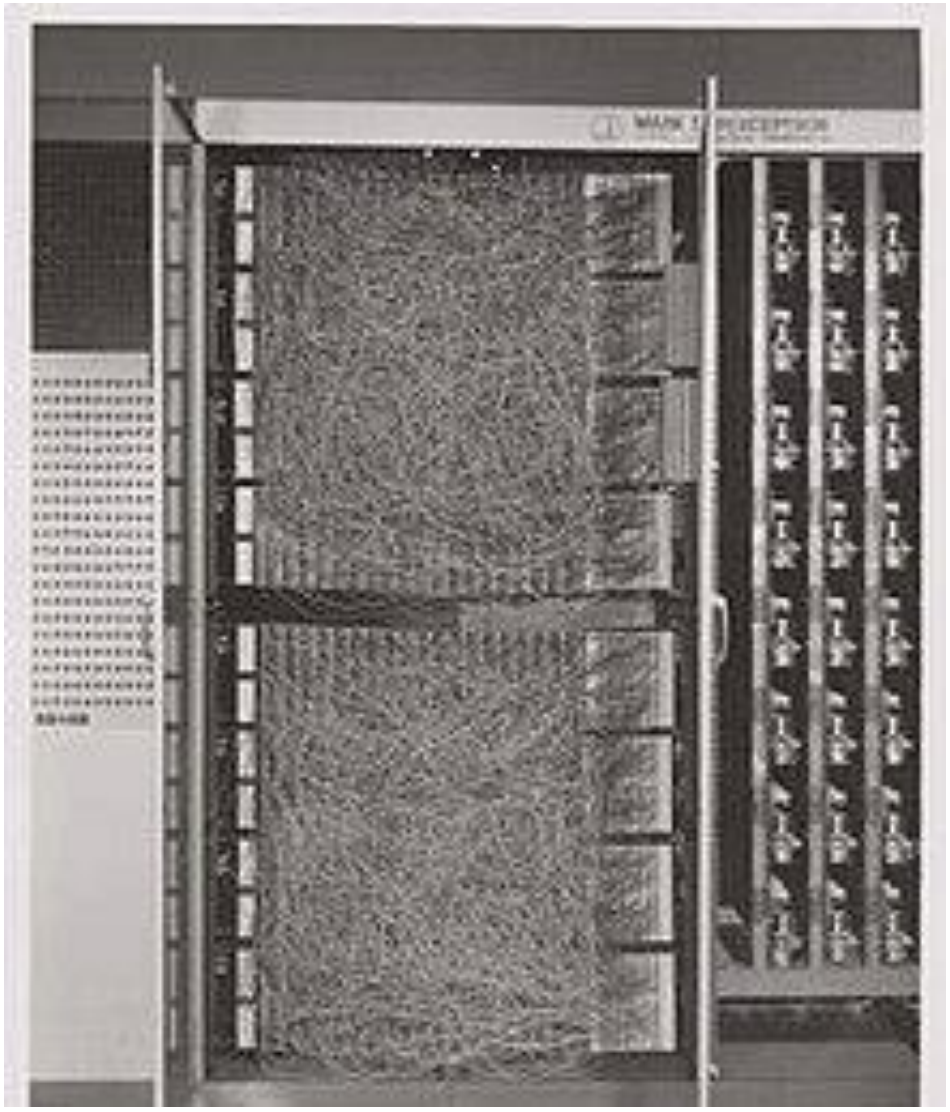
- Below threshold: 0
- Above threshold: 1



Desired value gets max output

History of NNs

- 1940s – Early Beginnings
  - Concept of a neural network is first proposed: “A Logical Calculus of Ideas Immanent in Nervous Activity”
- 1950s – The Perceptron
  - With funding from the US Navy, Cornell builds the Mark 1 Perceptron, a physical neural network
  - The New York Times reported the perceptron to be "the embryo of an electronic computer that [the Navy] expects will be able to walk, talk, see, write, reproduce itself and be conscious of its existence.”



History of NNs

- 1960s – The First AI Winter
  - Despite the excitement of the 50s, NN research stalled
  - A highly influential book – *Perceptrons (1969)* – showed that these early neural networks were severely limited
- 1980s – Backpropagation
  - The discovery of backpropagation allowed for the first time the creation of multi-layer neural networks that could efficiently learn from examples
- 1990s – Support Vector Machines and the Second AI Winter
  - NN research stalled again due to the rising popularity of SVMs, which provided a better theoretical framework and outperformed the NNs of the day

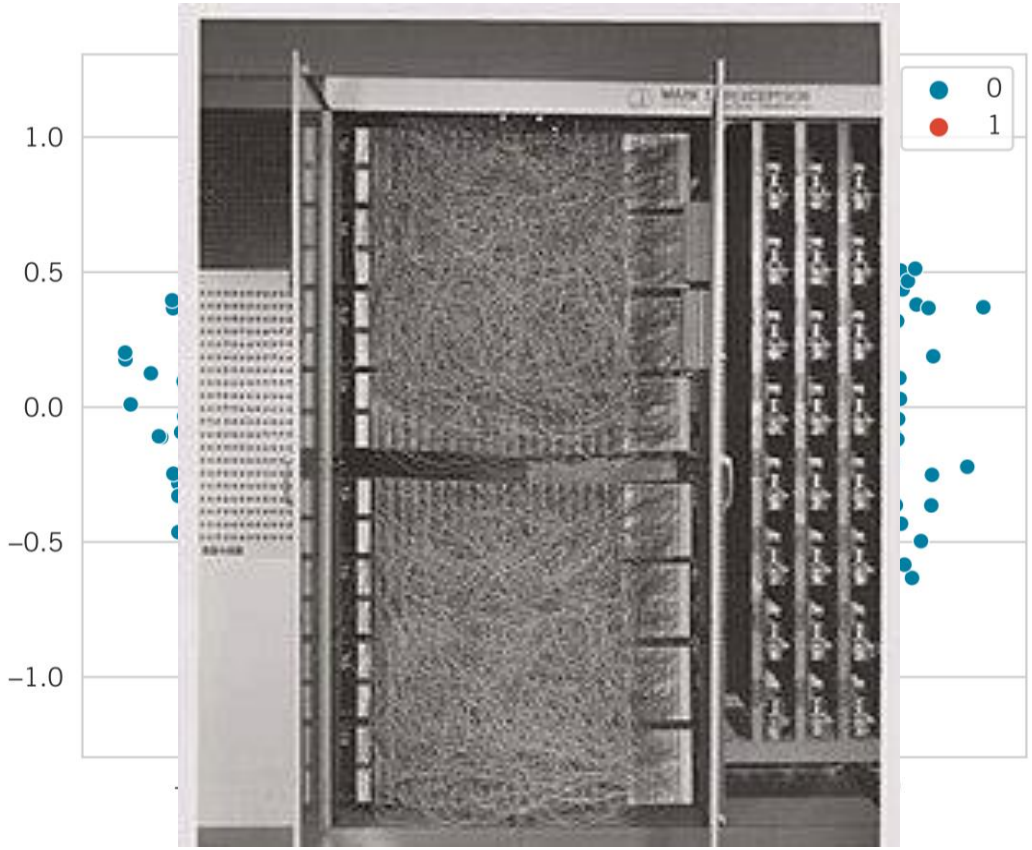
History of NNs

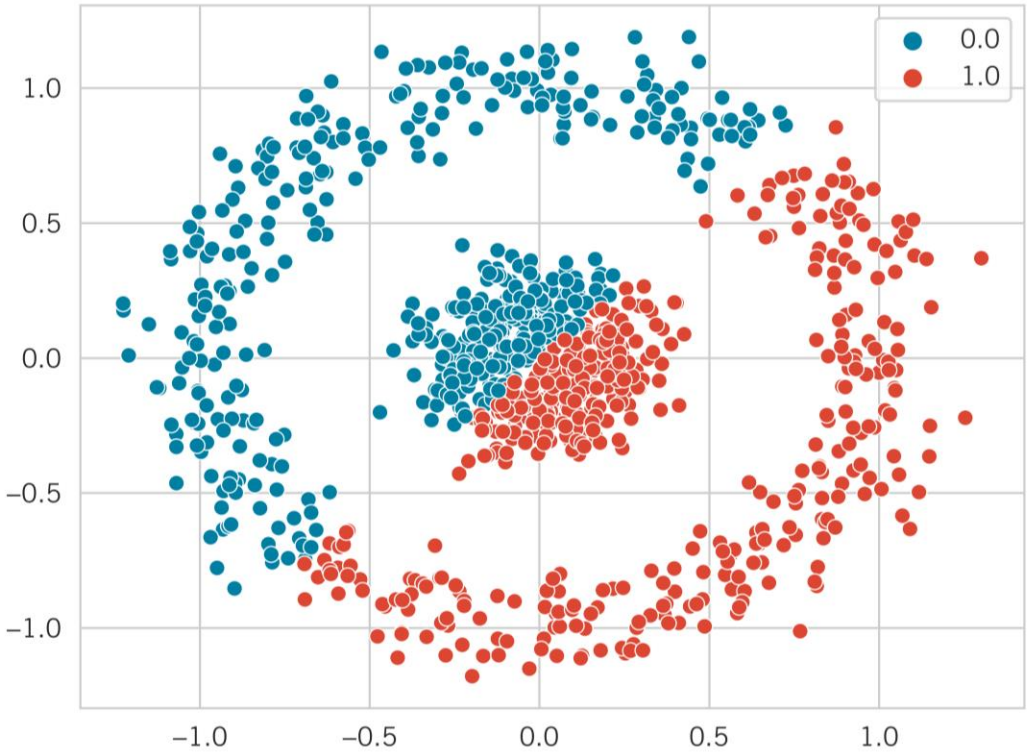
- 2000s – Dawn of the Deep Learning Era
  - The term “deep learning” began to circulate, reflecting a new focus on deeper, multi-layered neural networks
  - Advances in hardware, datasets, and training techniques allowed the development of much more sophisticated networks
- 2010s – Breakthroughs and Wide Adoption
  - With the success of AlexNet, Convolutional Neural Networks gained prominence and became a go-to method for image tasks
  - Recurrent Neural Networks show impressive results in natural langauge understanding
  - Tech giants begin to heavily invest in deep learning technology

History of NNs

- 2020s – Transformers and the Era of Large Language Models
  - The Transformer model, introduced in the paper “Attention is All You Need”, starts demonstrating state-of-the-art performance in language tasks
  - An increasing focus on large-scale models with billions, or even trillions, of parameters begins, leading to unprecedented performance…
  - …but also raising questions about computational efficiency, environmental impact, and accessibility.

Why do we want non-linearity?

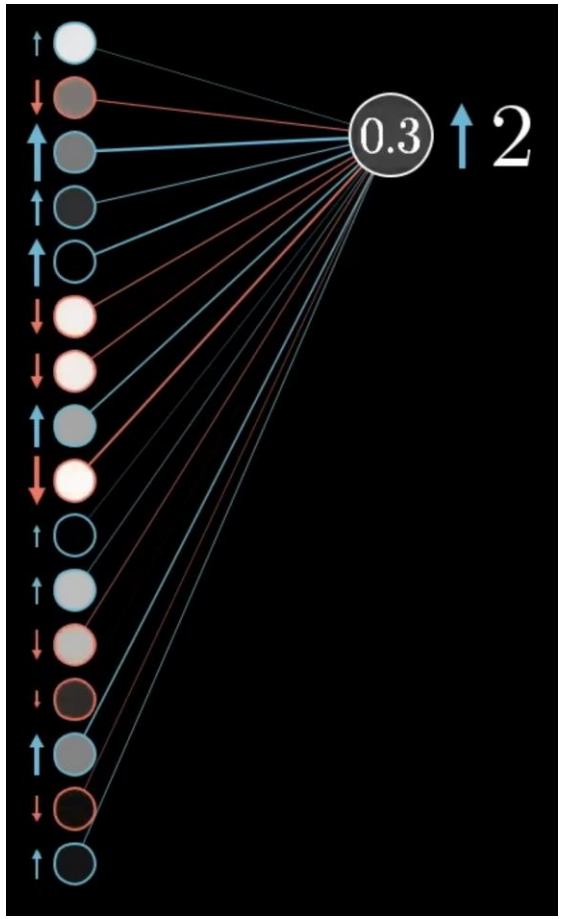


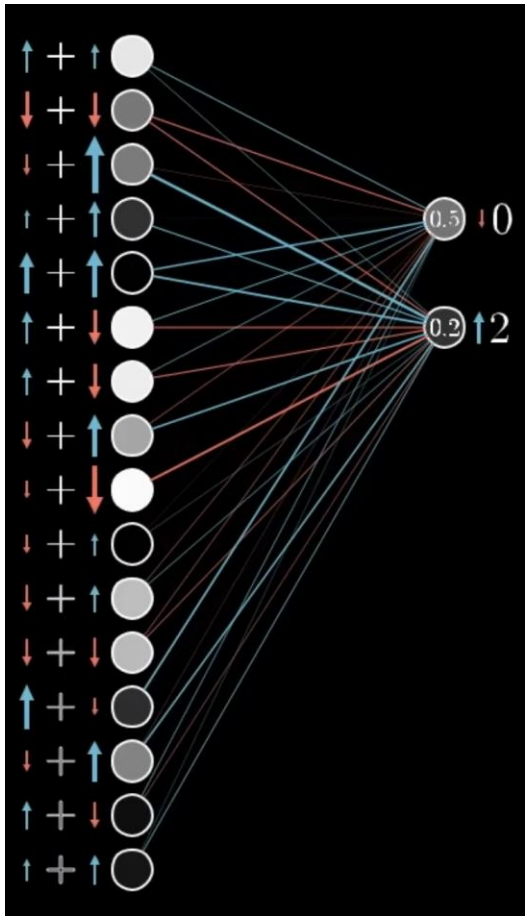


Cannot apply a linear classifier!

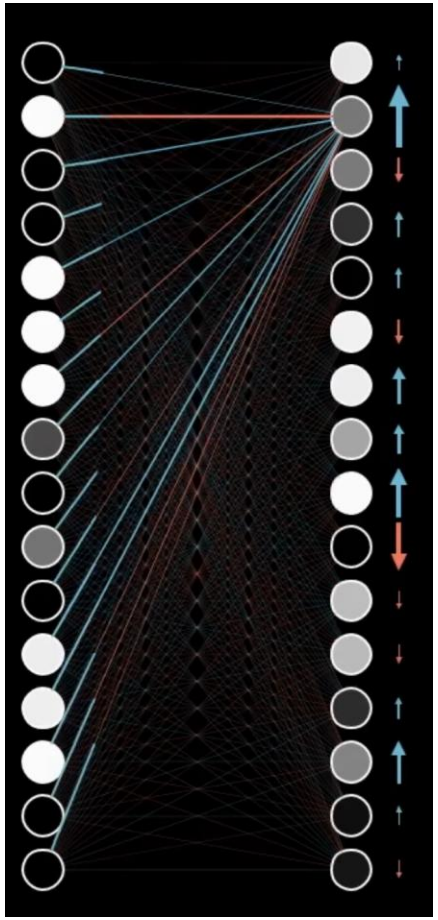
Training a network

Training a network





Training a network

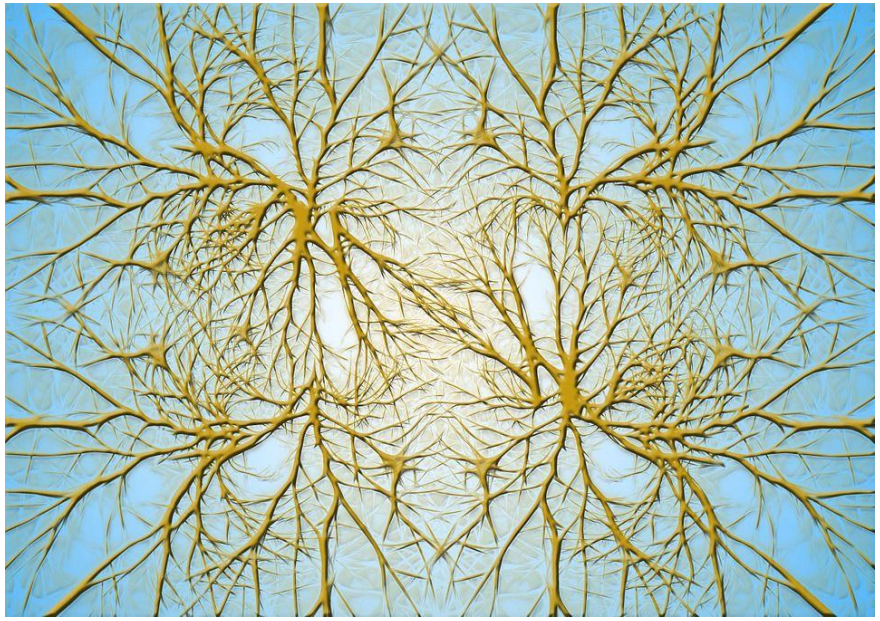


The Neuron Metaphor

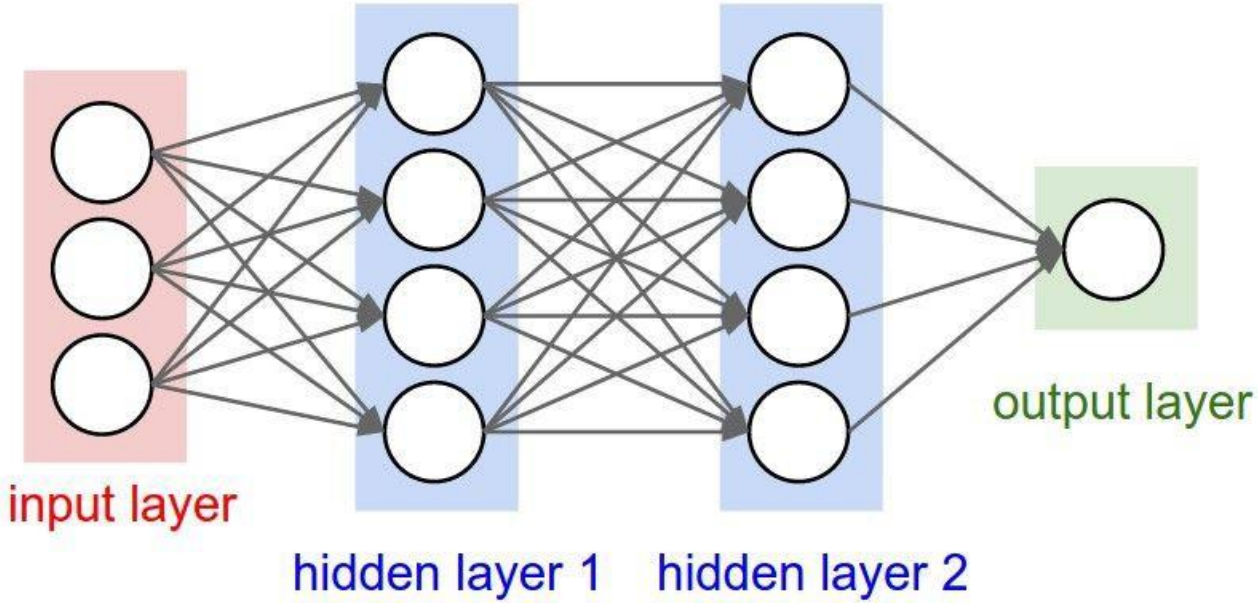
- Neural networks were inspired by our understanding of the brain and how neurons interact.
- An artificial neuron in a neural network takes in multiple inputs, applies a function to them, and generates an output — mirroring the basic functionality of a biological neuron.
- This analogy has been extremely useful for explaining and visualizing how these artificial structures work.

The Metaphor Breaks Down

Biological Neurons:
Complex connectivity patterns



Neurons in a neural network:
Organized into regular layers for computational efficiency



The Metaphor Breaks Down

- Biological neurons are vastly more complex: they use a mixture of electrical and chemical signals, have complex temporal dynamics, and can restructure their own connections.
- The brain is not just a feed-forward network: it has many complex feedback loops, which are not typically found in artificial neural networks.
- The brain isn't easily divided into distinct layers, as we do in artificial neural networks.

The Metaphor Breaks Down

- Over-reliance on the analogy can lead to misunderstandings about how neural networks function and their capabilities.
- This can lead to unrealistic expectations about what neural networks can do, or to overgeneralizations about their functioning.
- For instance, claiming a neural network "thinks" or "understands" like a human brain is misleading.
- To further progress, it's important to view artificial neural networks as mathematical/statistical tools, and not overstate the comparison to the human brain.

Neural Network Playground

https://playground.tensorflow.org