Introduction to Deep Learning

“What is deep learning?”

- What would be your answer?

What is Deep Learning?

OpenAI's ChatGPT, a Deep Learning model,
answers the question "what is deep learning" as follows:

what is deep learning?

Deep learning is a type of machine learning that involves using artificial neural networks to learn and make predictions or decisions based on data. It is called "deep" learning because these neural networks typically have many layers, with each layer processing and transforming the data in a different way. Deep learning has been shown to be effective at solving complex problems in fields such as computer vision, natural language processing, and speech recognition. It has also been used to improve the performance of many other machine learning algorithms.

OpenAI ChatGPT, retrieved on December 2022

What is Deep Learning?

what is deep learning?

1 2 4 3

Deep learning is a type of machine learning that involves using artificial neural networks to learn and make predictions or decisions based on data. It is called "deep" learning because these neural networks typically have many layers, with each layer processing and transforming the data in a different way. Deep learning has been shown to be effective at solving complex problems in fields such as computer vision, natural language processing, and speech recognition. It has also been used to improve the performance of many other machine learning algorithms.

Let's examine these four elements
(and see how accurate ChatGPT is)

OpenAI ChatGPT, retrieved on December 2022

What is Deep Learning?

what is deep learning?

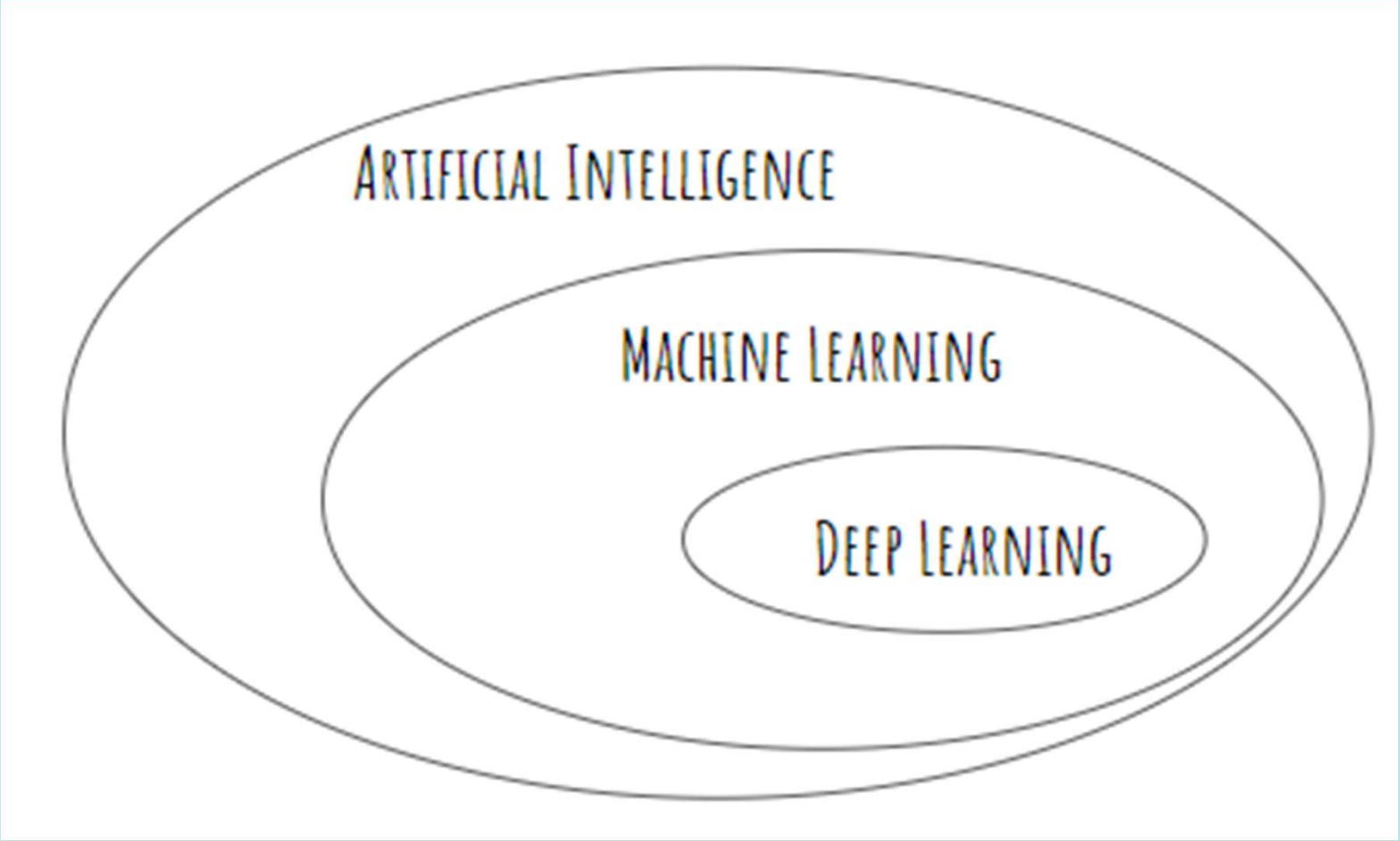
1 2 4 3

Deep learning is a type of machine learning that involves using artificial neural networks to learn and make predictions or decisions based on data. It is called "deep" learning because these neural networks typically have many layers, with each layer processing and transforming the data in a different way. Deep learning has been shown to be effective at solving complex problems in fields such as computer vision, natural language processing, and speech recognition. It has also been used to improve the performance of many other machine learning algorithms.

Let's examine these four elements
(and see how accurate ChatGPT is)

OpenAI ChatGPT, retrieved on December 2022

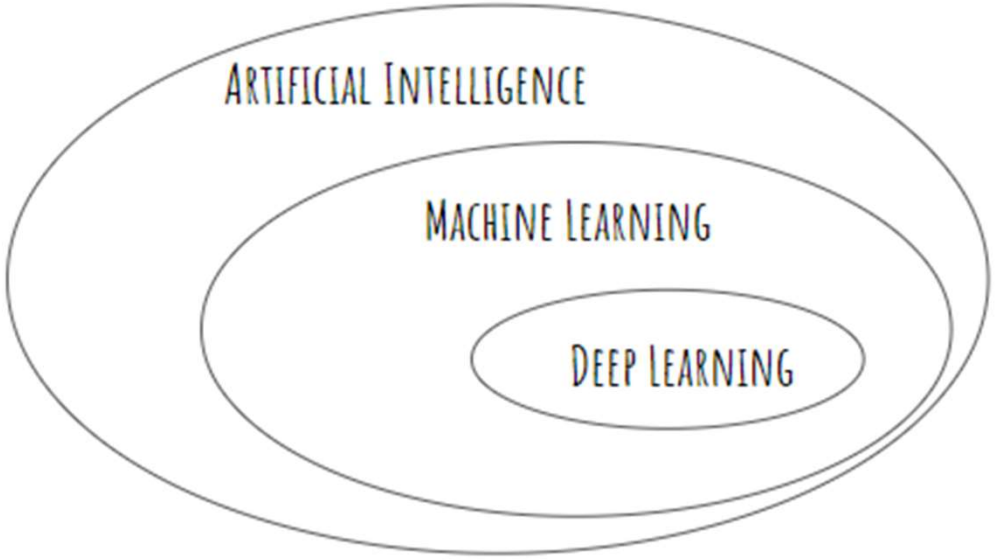
“Deep Learning is a Type of Machine Learning”



But what is Machine Learning?

After Chollet, Chapter 1

Deep Learning is a Type of Machine Learning



Artificial intelligence is intelligence demonstrated by machines

OED | Oxford English Dictionary
The definitive record of the English language

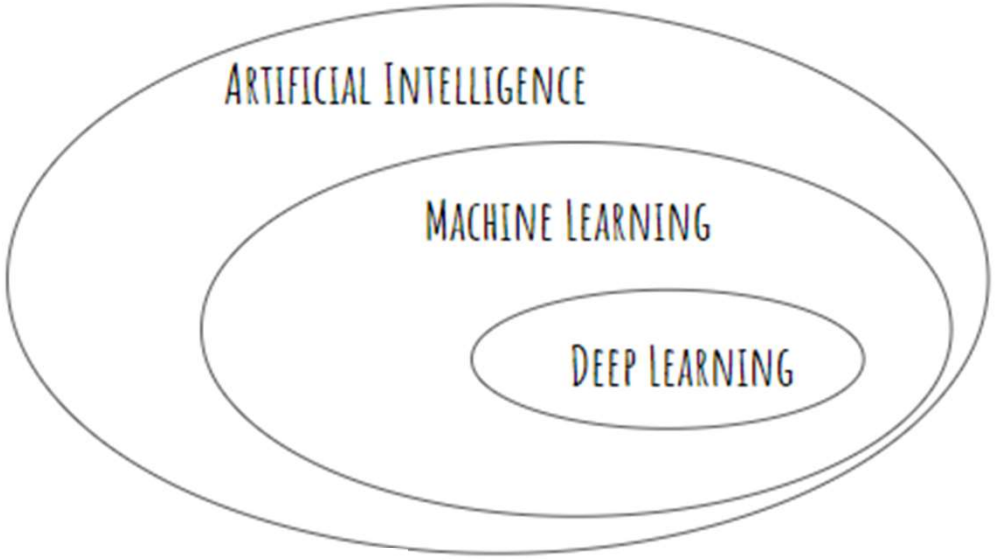
Artificial Intelligence: The capacity of computers or other machines to exhibit or simulate intelligent behavior

WIKIPEDIA
The Free Encyclopedia

Artificial intelligence is *intelligence - perceiving, synthesizing, and inferring information* - demonstrated by machines, as opposed to intelligence displayed by animals and humans.

OED, Retrieved December 2022; Wikipedia, Retrieved December 2022

Deep Learning is a Type of Machine Learning



Machine Learning – methods for achieving intelligence demonstrated by machines, where the intelligence was acquired by learning from experience (data)

WIKIPEDIA
The Free Encyclopedia

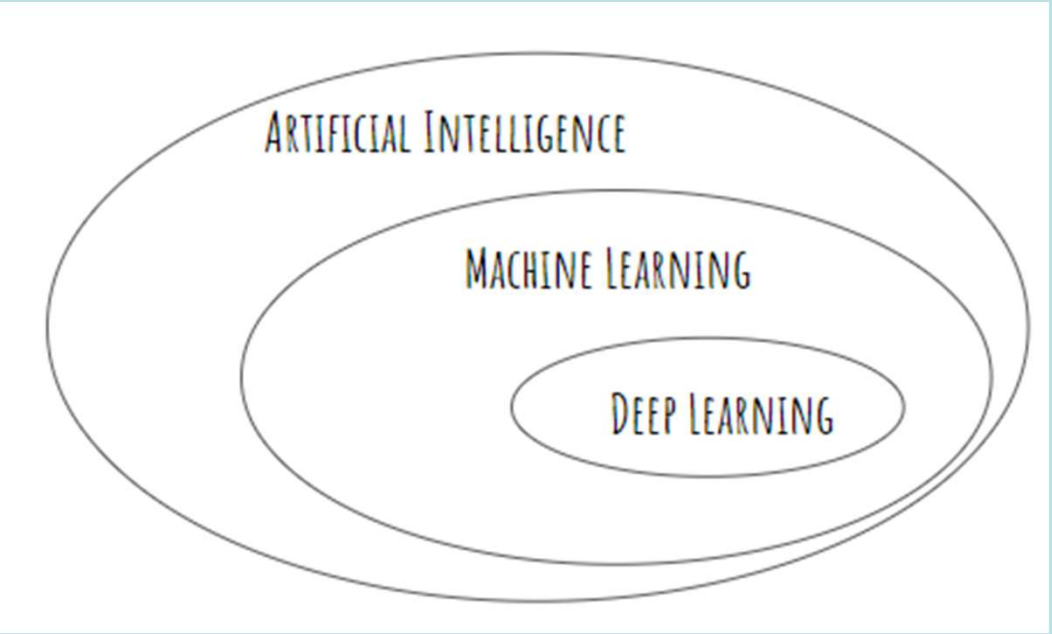
Machine Learning - the study of computer algorithms that improve automatically through experience.

Tom Mitchell's widely quoted definition of machine learning:

Machine Learning - A computer program is set to learn from an experience E with respect to some task T and some performance measure P if its performance on T as measured by P improves with experience E.

Wikipedia, Retrieved December 2022; Russell & Norvig (2003), pp. 649–788.

Deep Learning is a Type of Machine Learning

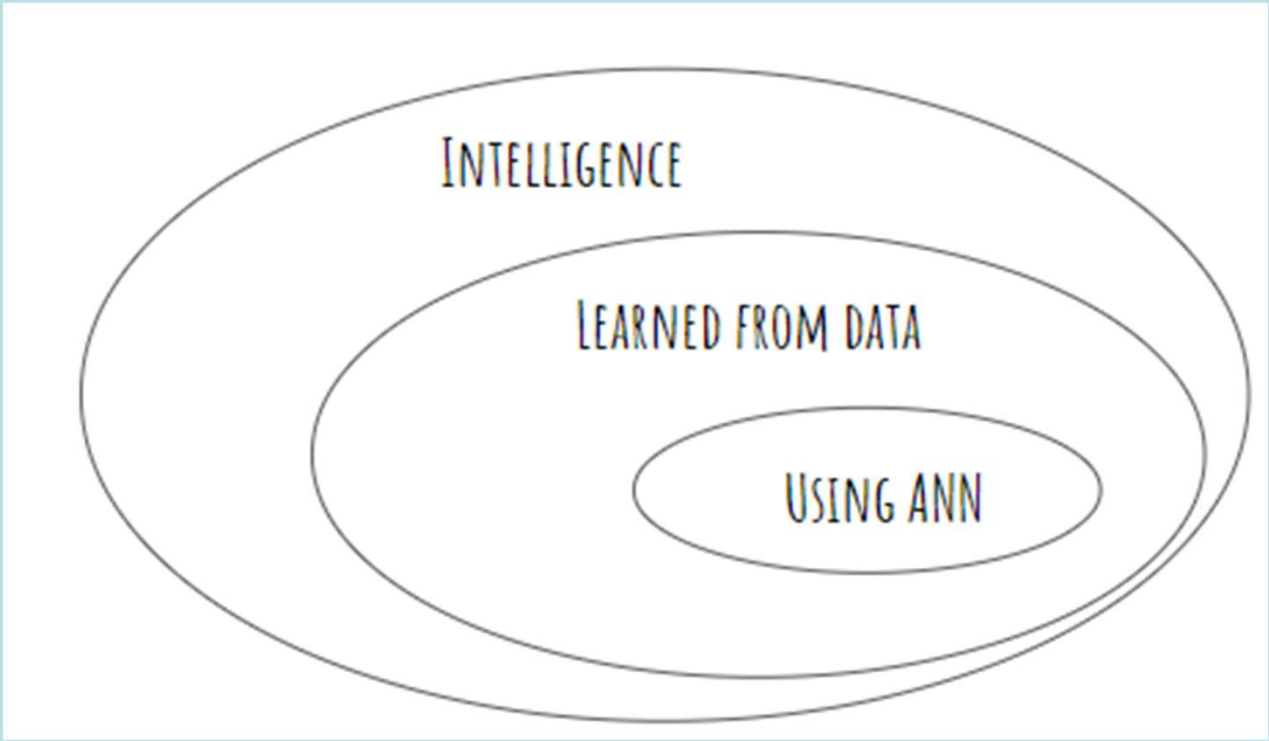


Deep Learning - methods for achieving intelligence demonstrated by machines, where the intelligence was acquired by learning from experience (data), and by using artificial neural networks.

The word “Deep” in “Deep Learning” refers to either:

1. A hierarchy of representations (Chollet, p.7, Goodfellow, p.1)
2. The number of layers in a feedforward-network or an unrolled RNN
3. The length of the credit assignment path (Sugiyama, p.15)

Deep Learning is a Type of Machine Learning



What is Deep Learning?

what is deep learning?

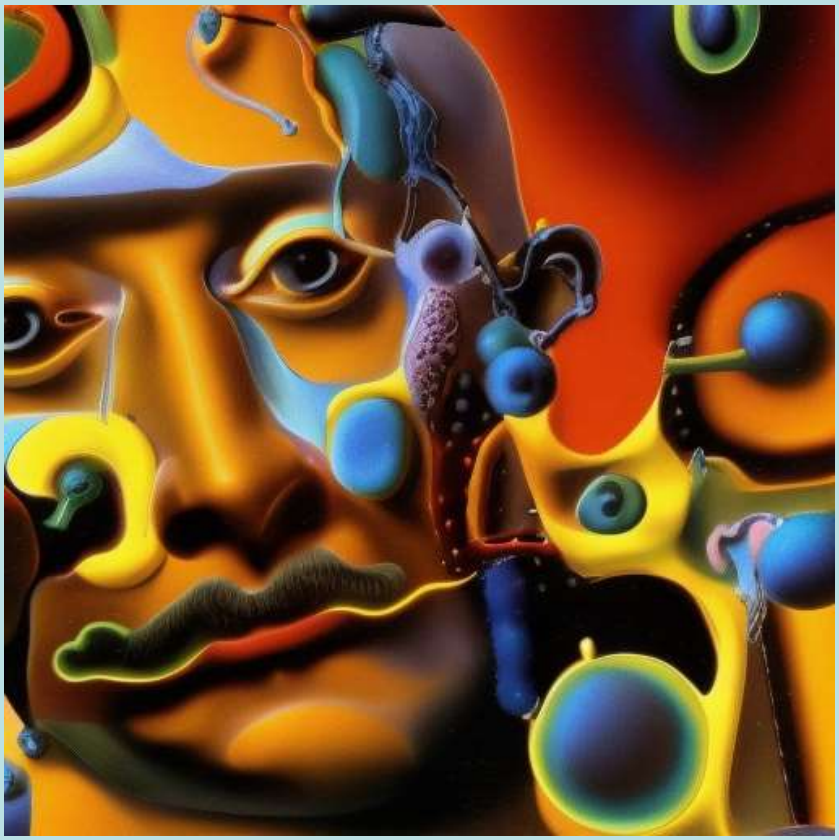
1 2

Deep learning is a type of machine learning that involves using artificial neural networks to learn and make predictions or decisions based on data. It is called "deep" learning because these neural networks typically have many layers, with each layer processing and transforming the data in a different way. Deep learning has been shown to be effective at solving complex problems in fields such as computer vision, natural language processing, and speech recognition. It has also been used to improve the performance of many other machine learning algorithms.

3 4

OpenAI's ChatGPT, a Deep Learning model, answers the question "what is deep learning?".

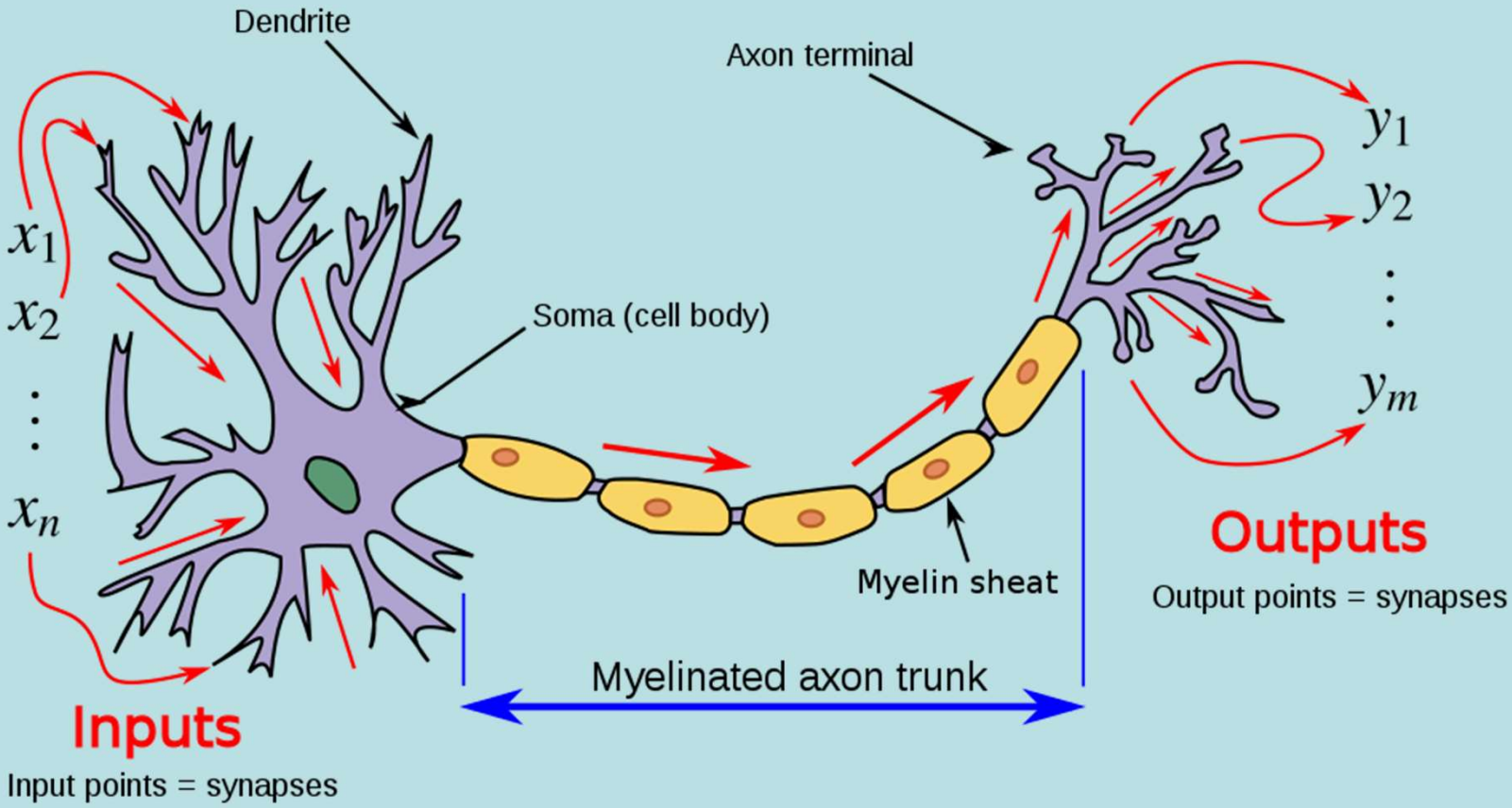
OpenAI ChatGPT, retrieved on December 2022



“Artificial neural network in the style of Salvador Dalí”
Produced with DreamStudio, by StableAI

The Biological Neuron Analogy

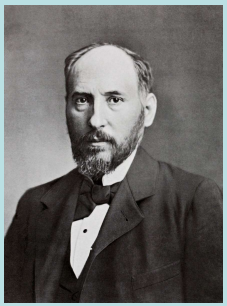
A Biological Neuron



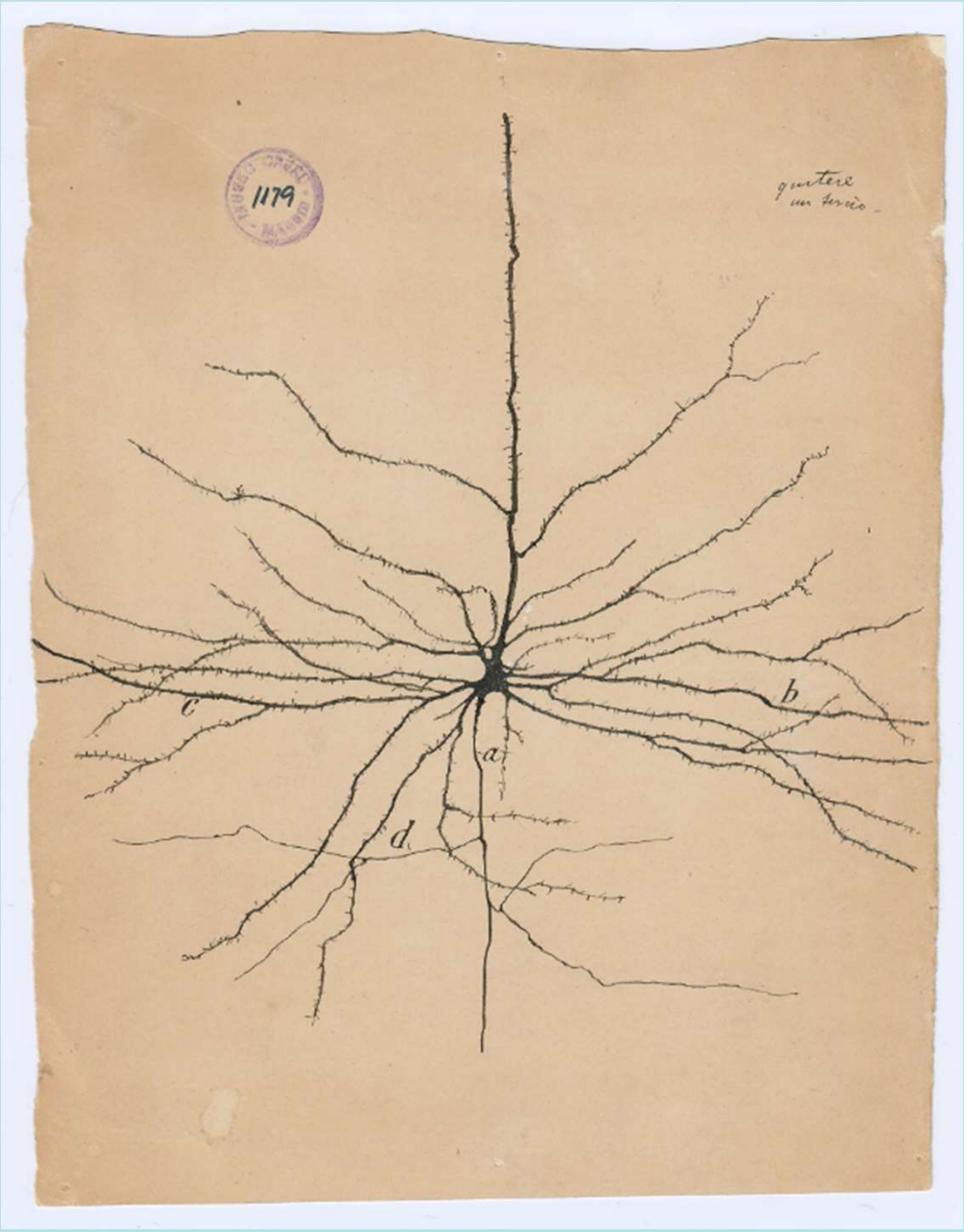
By Egm4313.s12 , Wikipedia

The Biological Neuron Analogy

A Biological Neuron



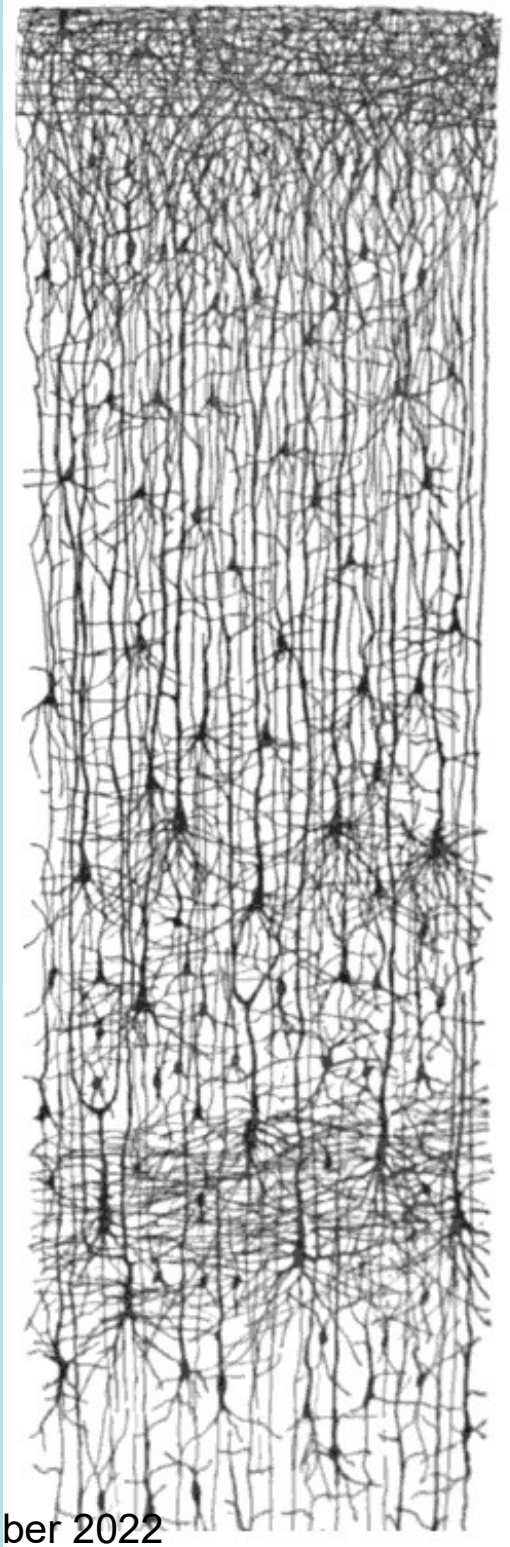
Wikipedia



Santiago Ramón y Cajal, The pyramidal neuron of the cerebral cortex, 1904, Credit: Cajal Institute (CSIC), Madrid

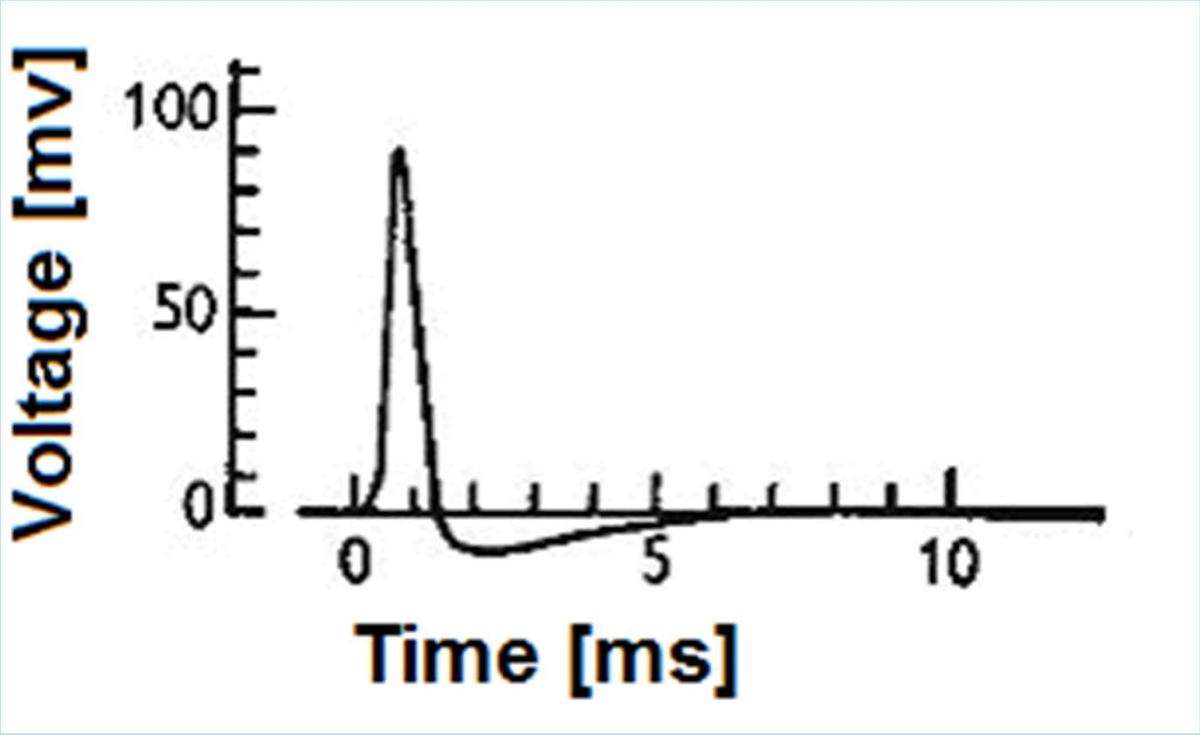
The Biological Neuron Analogy

- Drawing of cortical lamination by Santiago Ramon y Cajal,
- A vertical cross-section, with the surface of the cortex at the top.
- (this is a Golgi-stain which shows the dendrites and axons of a random subset of neurons).



The Biological Neuron Analogy

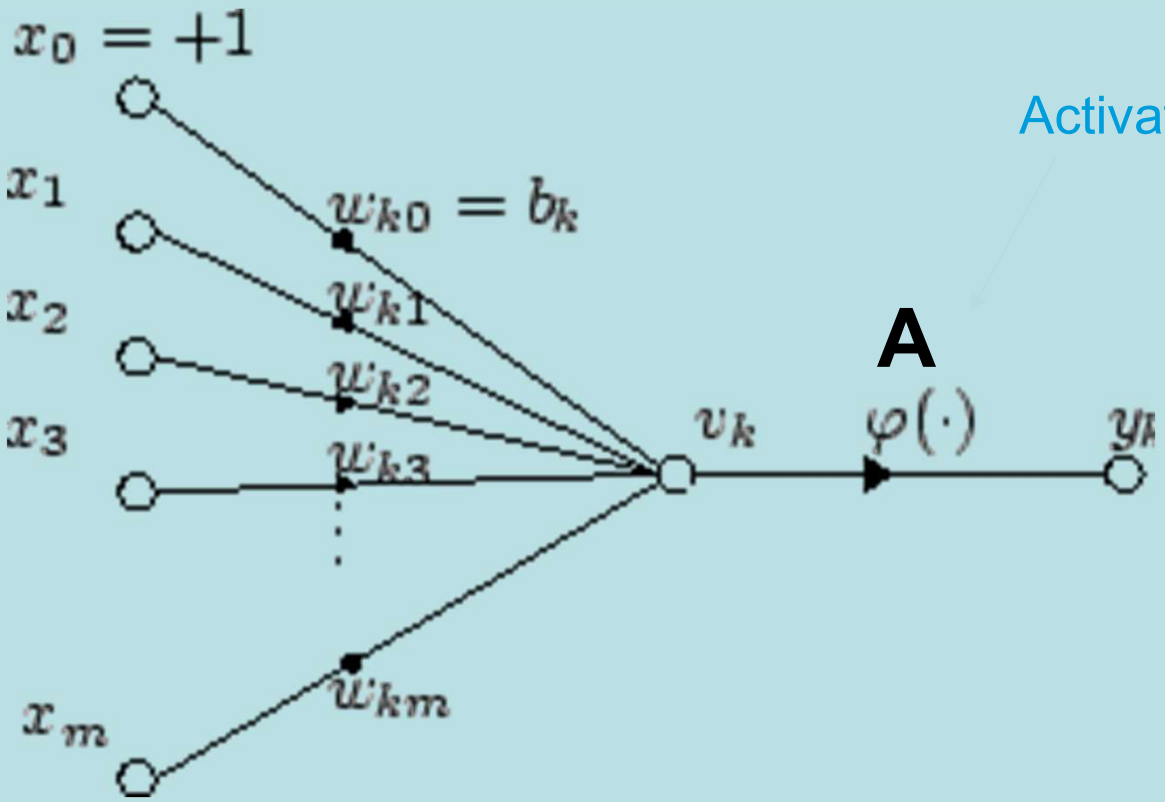
- All-or-Nothing
- Integrate-and-Fire



Time course of neuronal action potential ("spike").

By Nir.nossenson - Wikipedia

Modeling a Single Neuron

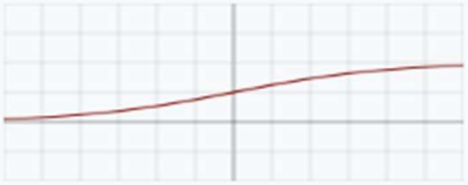
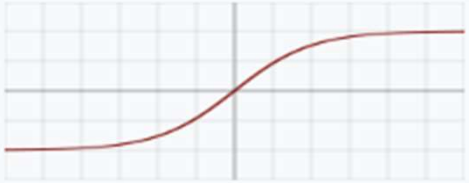


$$z = w_0 * x_0 + w_1 * x_1 ... + w_m * x_m = \sum_{j=0}^{m} w_j x_j$$

$$y = A(z)$$

Wikipedia

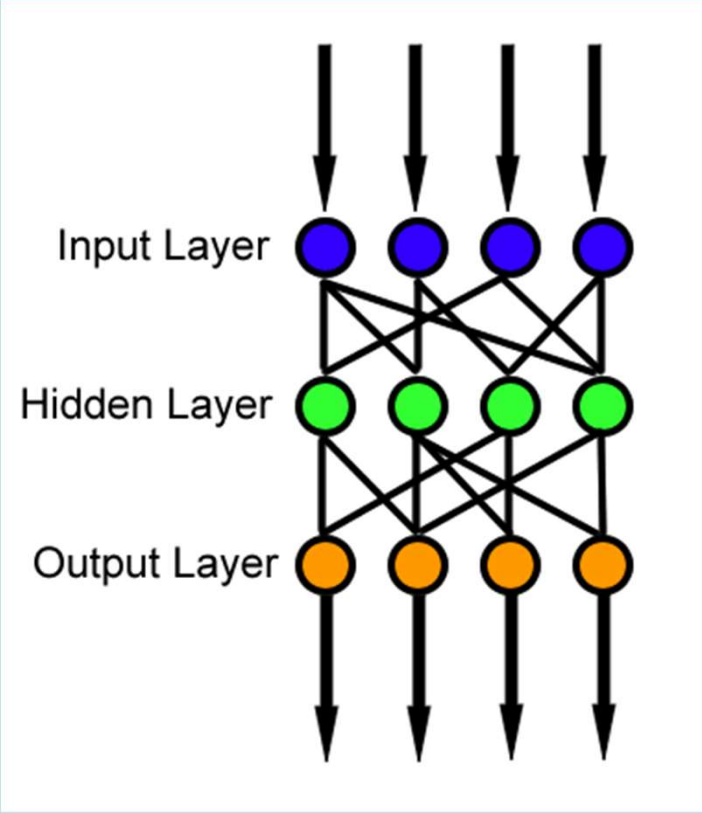
Activation Functions

Name	Plot	Function, $g(x)$
Identity		x
Binary step		$\begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$
Logistic, sigmoid, or soft step		$\sigma(x) \doteq \frac{1}{1+e^{-x}}$
Hyperbolic tangent (tanh)		$\tanh(x) \doteq \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$
Rectified linear unit (ReLU)[8]		$(x)^{+} \doteq \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } x > 0 \end{cases}$ $= \max(0,x) = x\mathbf{1}_{x>0}$

Wikipedia

Two Types of Neural Networks

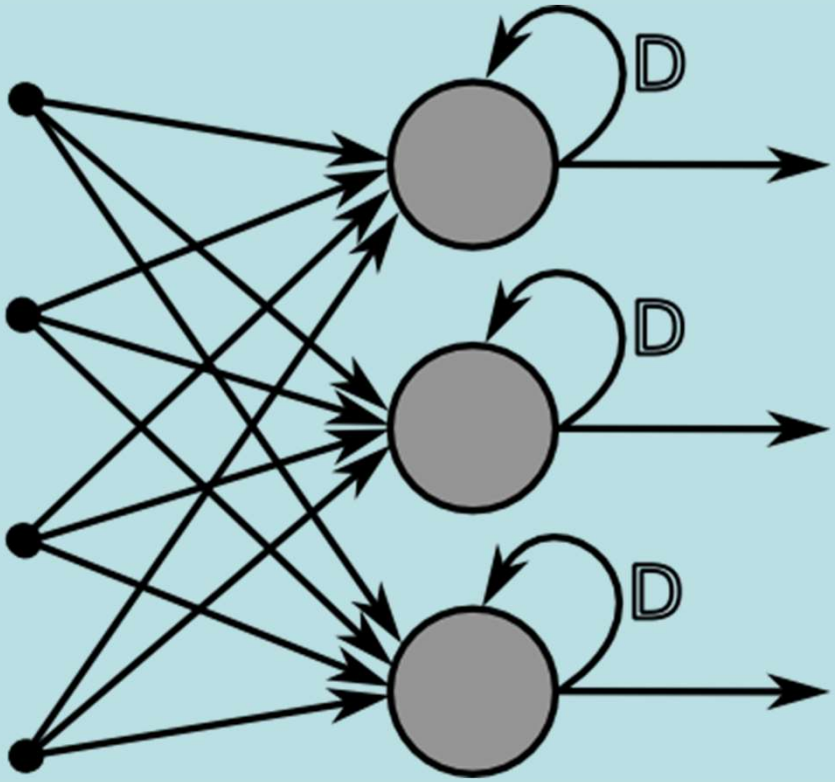
Feed-Forward



Directed Acyclic Graph
(DAG)

Wikipedia

Recurrent



Directed Cyclic Graph

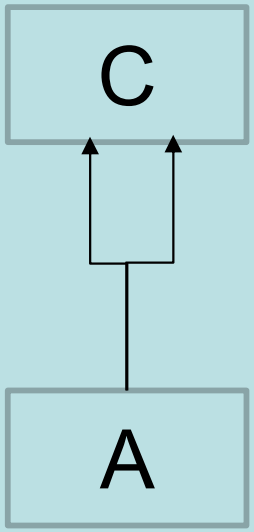
Wikipedia

Let’s look at the simplest neural networks

McCulloch-Pitts Neuron

Binary step		$\begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$

$$f(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{w} \cdot \mathbf{x} + b > 0, \\ 0 & \text{otherwise} \end{cases}$$



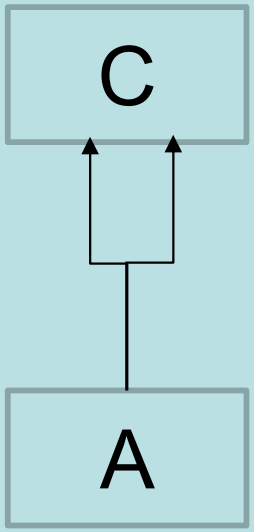
A	C
0	?
1	?

What is the output of the C neuron?

- W - The number of solid arrows represents the weight
- B – the bias is -1 (that is b= -1)

Geron p.303

$$f(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{w} \cdot \mathbf{x} + b > 0, \\ 0 & \text{otherwise} \end{cases}$$



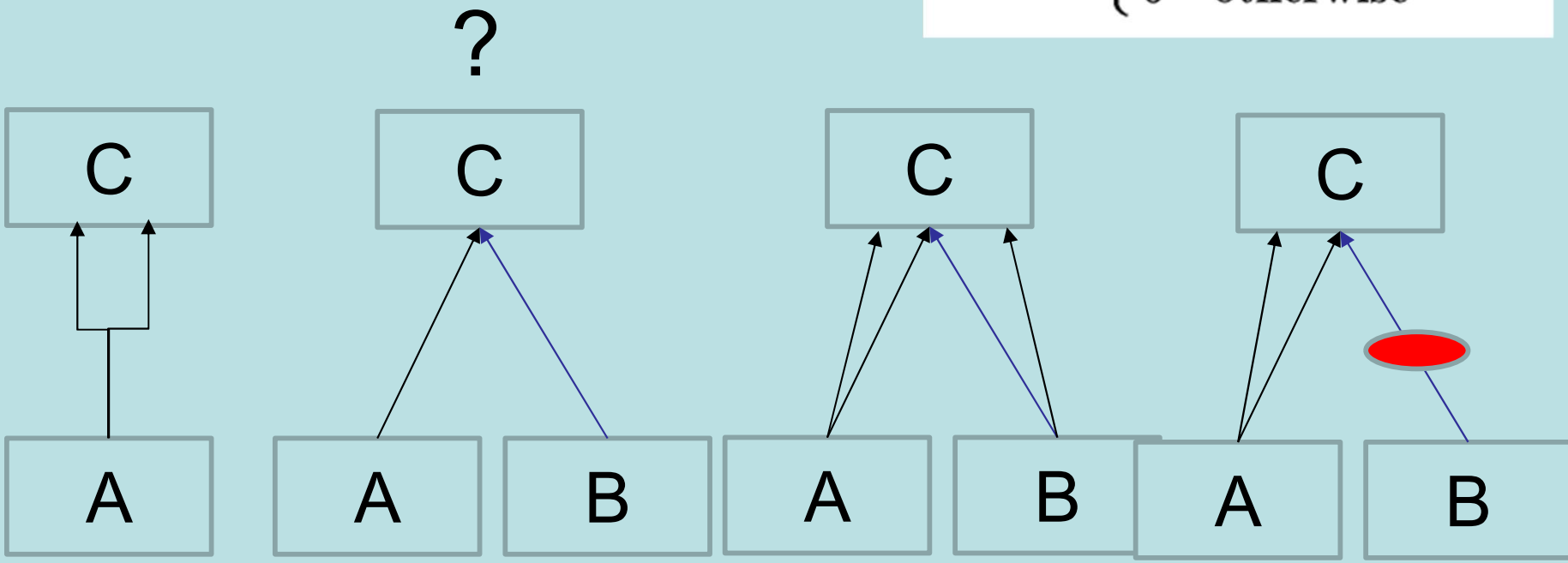
A	C
0	0
1	1

- W - The number of solid arrows represents the weight
- B – the bias is -1 (that is b= -1)

Geron p.303

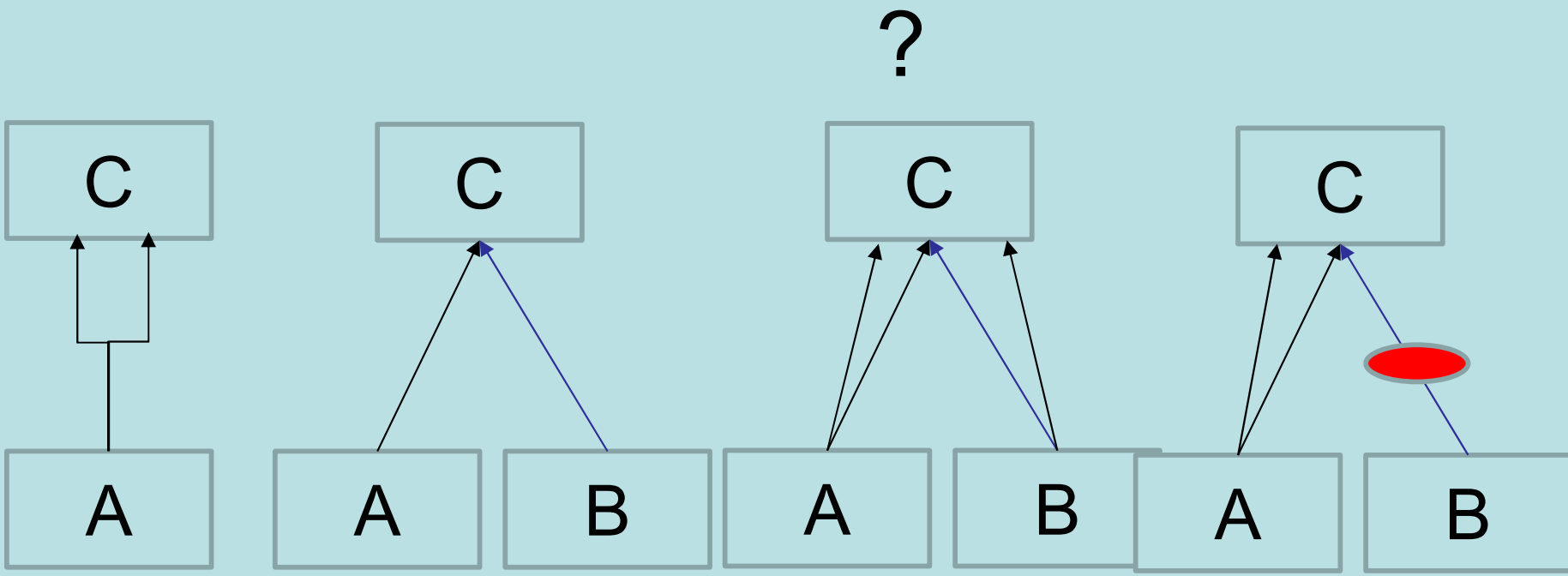
Bias of neuron C = -1 (b = -1)

$$f(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{w} \cdot \mathbf{x} + b > 0, \\ 0 & \text{otherwise} \end{cases}$$



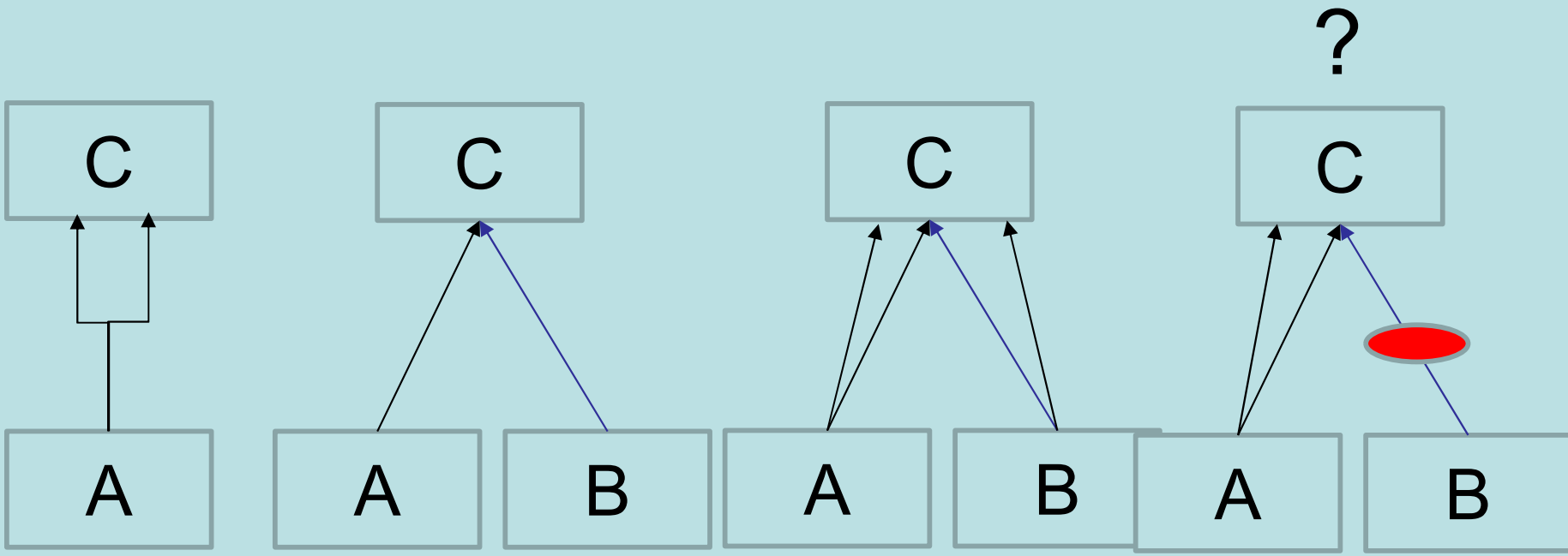
[A, B]	C	C	C	C
[0, 0]	0			
[0, 1]	0			
[1, 0]	1			
[1, 1]	1			

Geron p.303



[A, B]	C	C	C	C
[0, 0]	0	0		
[0, 1]	0	0		
[1, 0]	1	0		
[1, 1]	1	1		

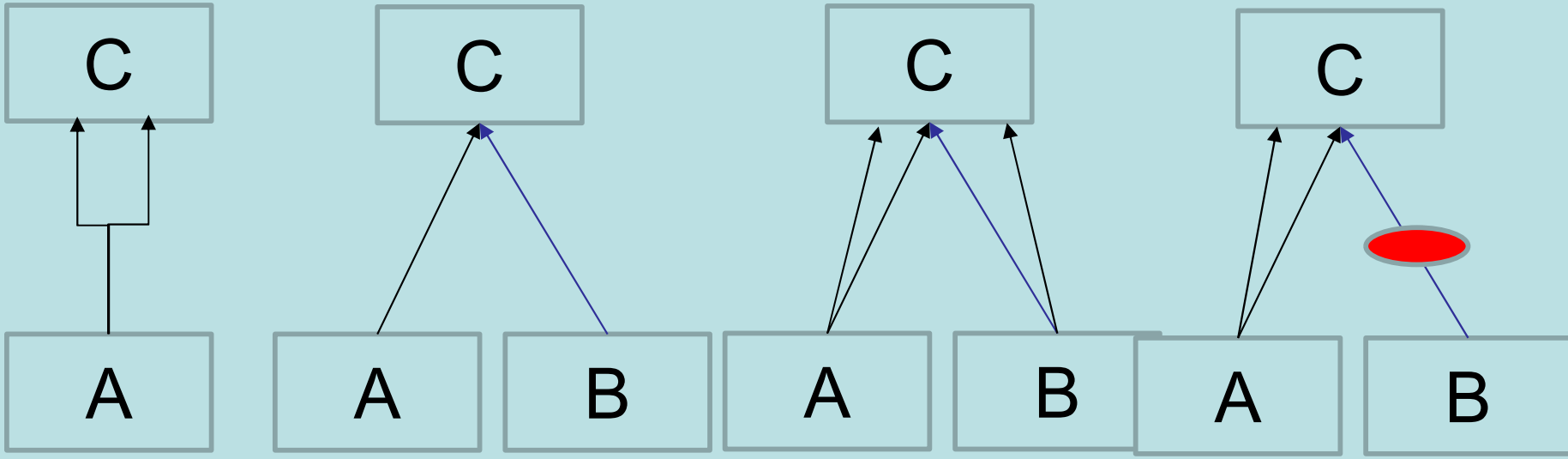
Geron p.303



[A, B]	C	C	C	C
[0, 0]	0	0	0	
[0, 1]	0	0	1	
[1, 0]	1	0	1	
[1, 1]	1	1	1	

Geron p.303

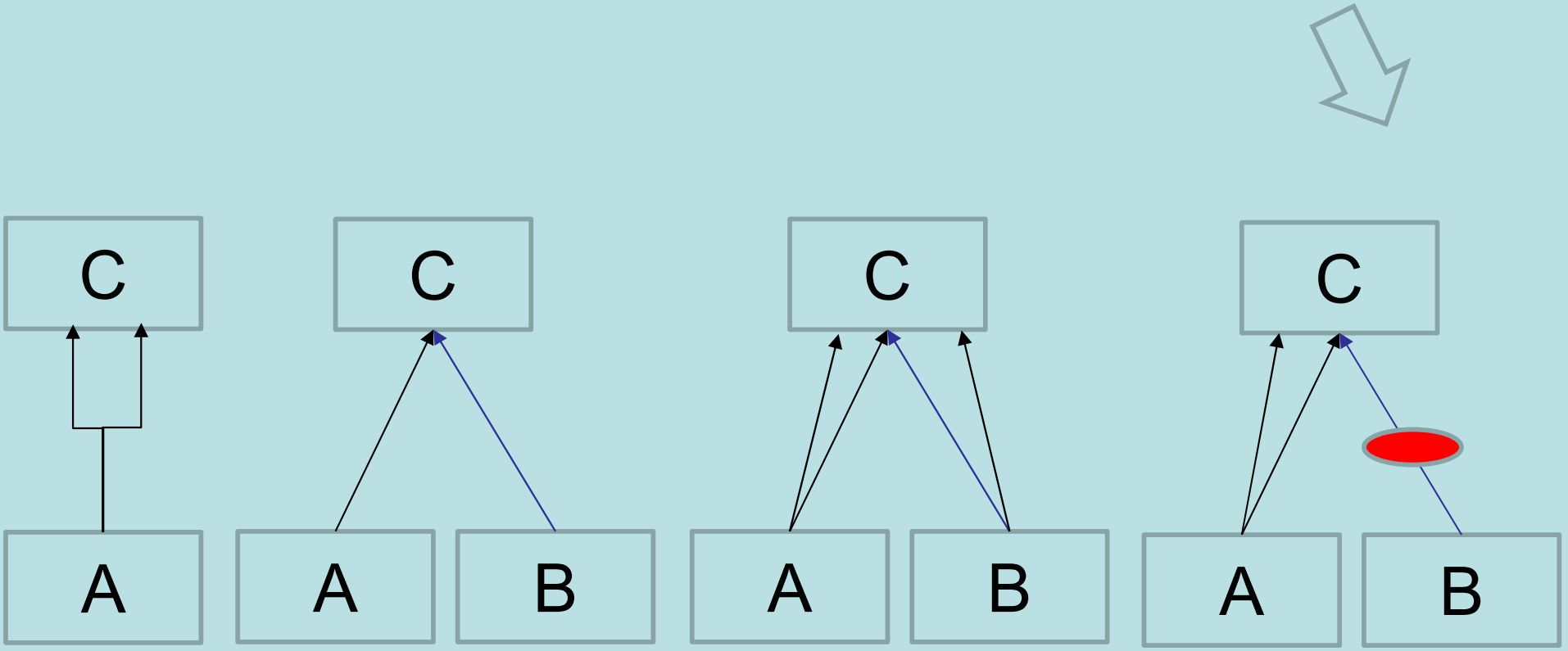
Which logic gate (“And, Or, Not”) does each of these networks represent?



[A, B]	C	C	C	C
[0, 0]	0	0	0	0
[0, 1]	0	0	1	0
[1, 0]	1	0	1	1
[1, 1]	1	1	1	0

Geron p.303

Neuronal Logic Gates



C=A

C=A **and** B

C=A **or** B

C=A **not** B

Geron p.303

Artificial Neural Network

Artificial neural networks (ANNs) -are computing systems inspired by the biological neural networks that constitute animal brains.

An ANN is based on a collection of connected units or **nodes** called artificial neurons, which loosely model the neurons in a biological brain. Each **connection**, like the synapses in a biological brain, can transmit a signal to other neurons. An artificial neuron receives signals then processes them and can signal neurons connected to it. The "signal" at a connection is a **real number**, and the output of each neuron is computed by some **non-linear function** of the sum of its inputs. The connections typically have a **weights** that adjusts as learning proceeds. The weight increases or decreases the strength of the signal at a connection. Neurons may have a **threshold** such that a signal is sent only if the aggregate signal crosses that threshold.

Wikipedia

The universal approximation theorem

A feedforward network with a linear output layer and at least one hidden layer with any “squashing” activation function (such as a sigmoid) can approximate ***any continuous function*** with any desired nonzero amount of error, provided that the network is given enough hidden units.

Goodfellow et al., Deep Learning, Section 6.4.1, p. 194-197

So far we have reviewed what neural networks are.

We also indicated that neural nets are very capable *theoretically*. However, the required architecture, training methodology and data are not given by the mathematical theorems.

And this is what we are going to look into next.

Definition of Deep Learning

what is deep learning?

1 2

Deep learning is a type of machine learning that involves using artificial neural networks to learn and make predictions or decisions based on data. It is called "deep" learning because these neural networks typically have many layers, with each layer processing and transforming the data in a different way. Deep learning has been shown to be effective at solving complex problems in fields such as computer vision, natural language processing, and speech recognition. It has also been used to improve the performance of many other machine learning algorithms.

4

3

OpenAI's ChatGPT, a Deep Learning model, answers the question "what is deep learning?".

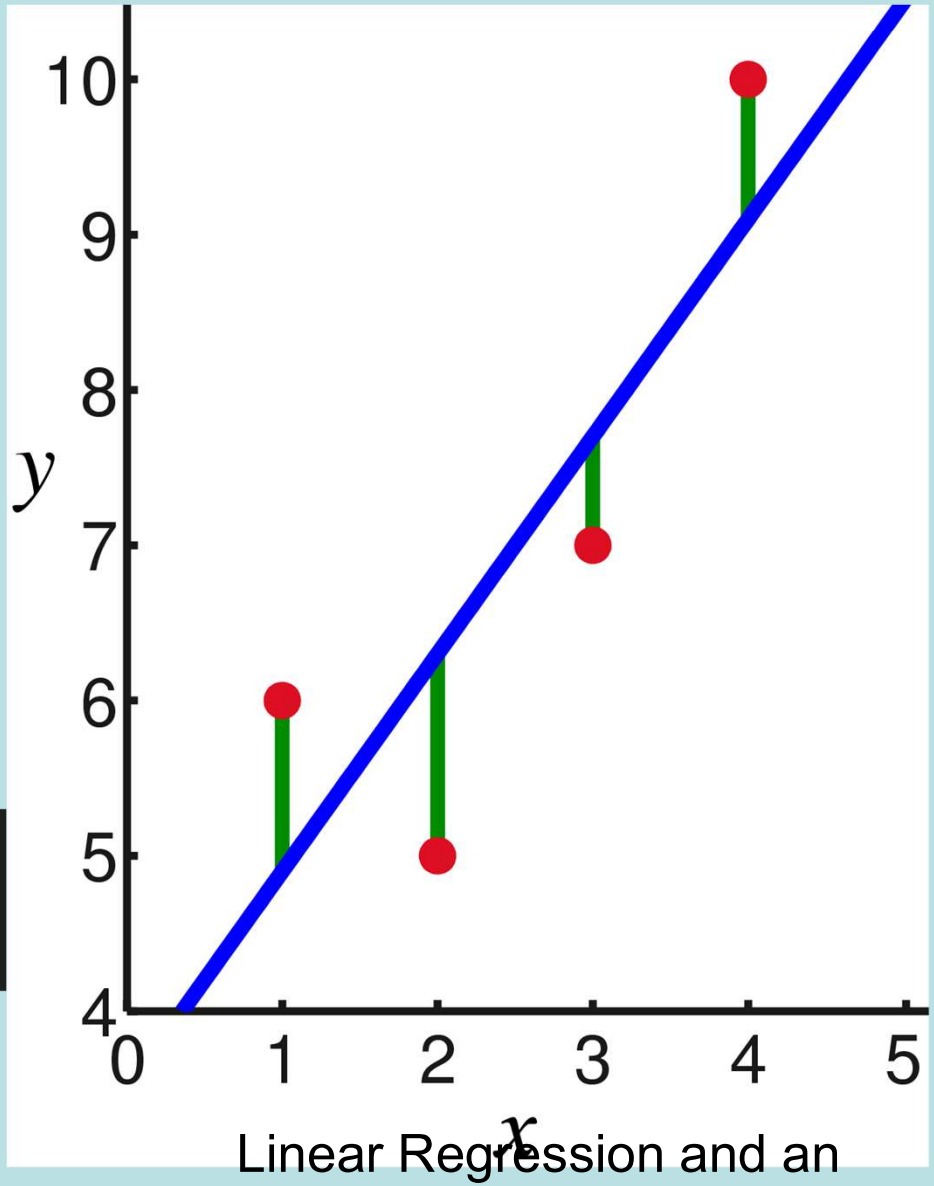
OpenAI ChatGPT, retrieved on December 2022

Loss Function

A loss function (or cost function) - is a function that maps values of one or more variables onto a real number intuitively representing some "cost" associated with the event, that is negatively correlated with a success measure (such as accuracy). An optimization problem seeks to minimize a loss function.

Mean Squared Error (MSE):

$$L = \frac{1}{m} * \sum_{n=1}^{m} (y_{predicted,n} - y_n)^2$$



Linear Regression and an associated Loss Function (in green)

https://en.wikipedia.org/wiki/Mean_squared_error

By Oleg Alexandrov Wikipedia

Definition of Deep Learning

what is deep learning?

1 2 4 3

Deep learning is a type of machine learning that involves using artificial neural networks to learn and make predictions or decisions based on data. It is called "deep" learning because these neural networks typically have many layers, with each layer processing and transforming the data in a different way. Deep learning has been shown to be effective at solving complex problems in fields such as computer vision, natural language processing, and speech recognition. It has also been used to improve the performance of many other machine learning algorithms.

OpenAI's ChatGPT, a Deep Learning model, answers the question "what is deep learning?".

OpenAI ChatGPT, retrieved on December 2022

Learning with Gradient Descent

By U.S. Forest Service - Okanogan-Wenatchee National Forest, morning fog shrouds trees, Wikipedia

Learning with Gradient Descent

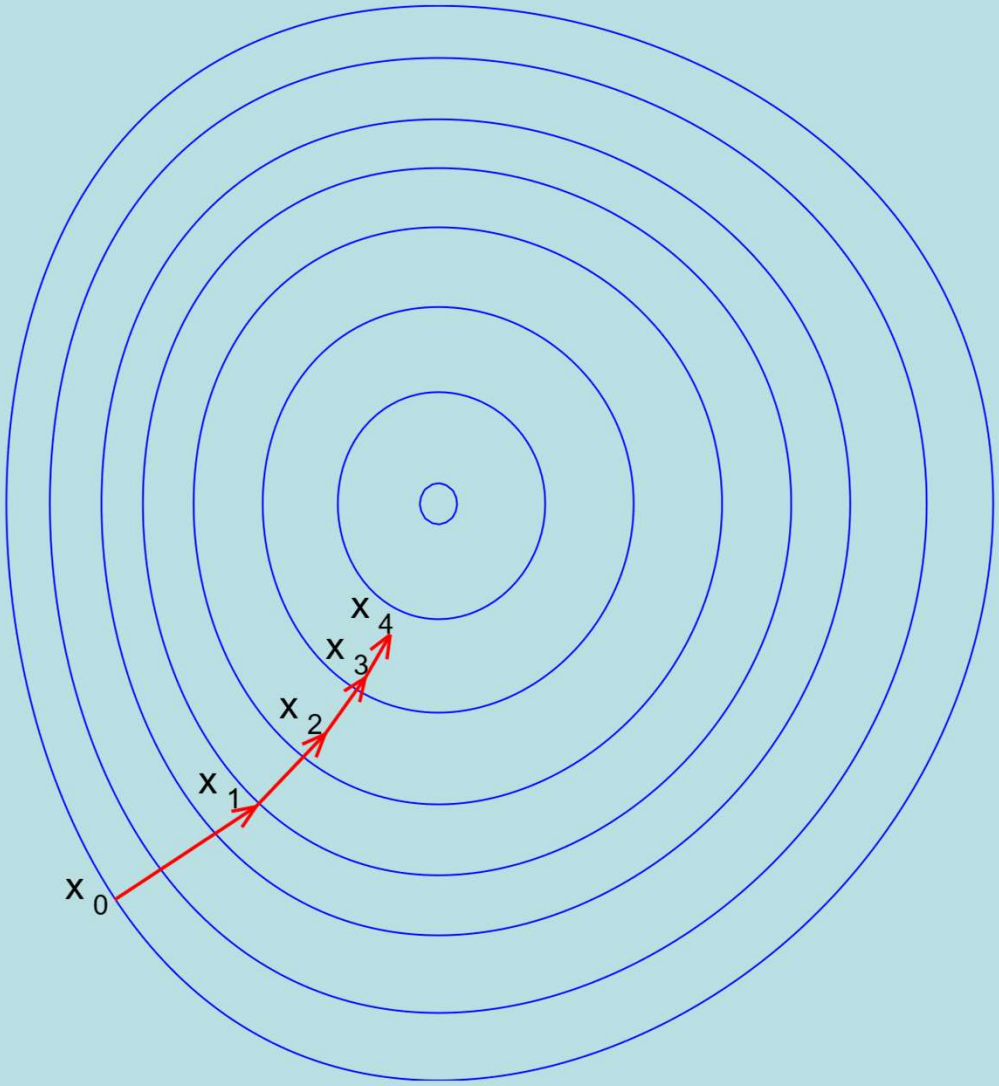
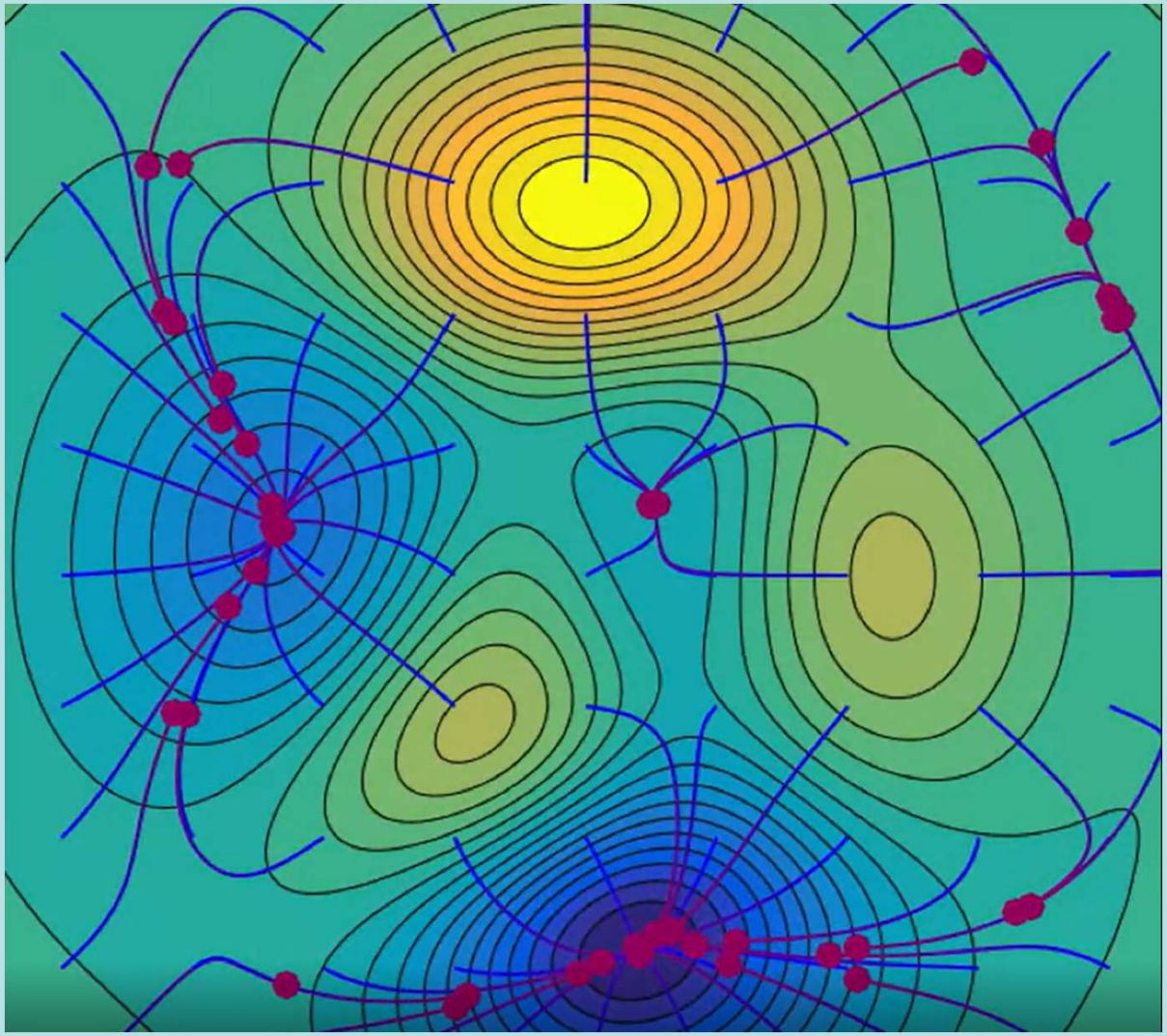


Illustration of gradient descent on a series of level sets of a 2D -> 1D convex function

Wikipedia

Learning with Gradient Descent



Video: https://upload.wikimedia.org/wikipedia/commons/4/4c/Gradient_Descent_in_2D.webm

Definition of Deep Learning

what is deep learning?

Deep learning is a type of machine learning that involves using artificial neural networks to learn and make predictions or decisions based on data. It is called "deep" learning because these neural networks typically have many layers, with each layer processing and transforming the data in a different way. Deep learning has been shown to be effective at solving complex problems in fields such as computer vision, natural language processing, and speech recognition. It has also been used to improve the performance of many other machine learning algorithms.

Take-Aways: Definitions

- Deep Learning (DL) is a type of Machine Learning.
- In the general sense, DL is distinguished from ML models by having many intermediate representations (Goodfellow and Chollet).
- In practice, DL models are artificial neural networks (ANN), trained based on gradient descent algorithms.

Take-Aways: Single Neuron

- An Artificial neuron sums up all of its input, and feeds it into an activation function
- The activation function might be linear ('pass-through'), or non-linear (e.g. sigmoid or a ReLU). Its output is the neuron's output

Take-Aways: Loss Function

- In DL, Learning from data is equivalent to training of a neural network
- During training, a neural network requires an objective by which its performance can be numerically measured, and this is the loss function
- A common loss function for regression problems in the mean squared error (MSE)

Take-Aways: Gradient Descent

- Training a neural network is done by minimizing the value of its loss function for a given dataset.
- This is done by taking small steps, in the direction opposite to the direction of the gradient. This is called gradient descent (a detailed example will appear soon).