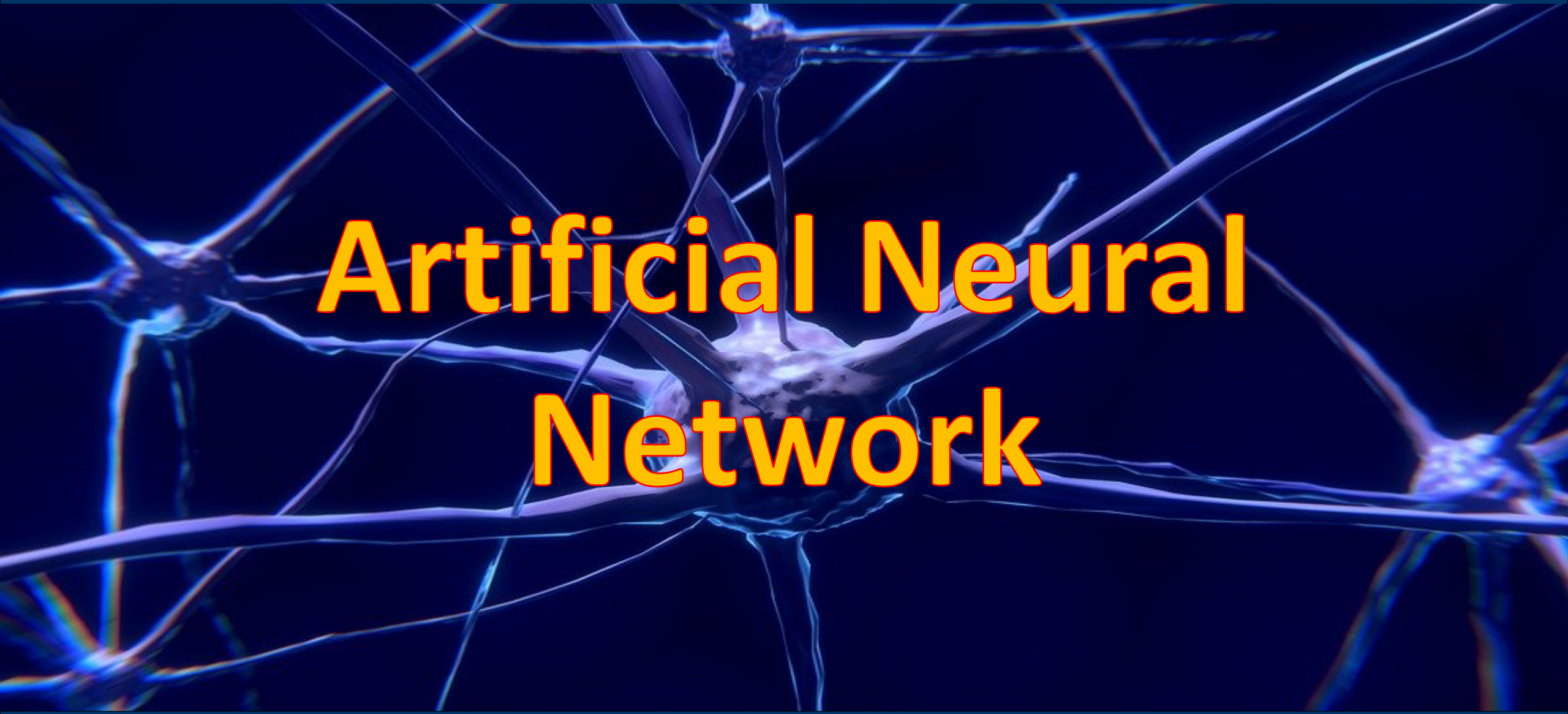
Advanced Machine Learning and Deep Learning

Mohammad Pourhomayoun
Assistant Professor
Computer Science Department
California State University, Los Angeles


Artificial Neural
Network

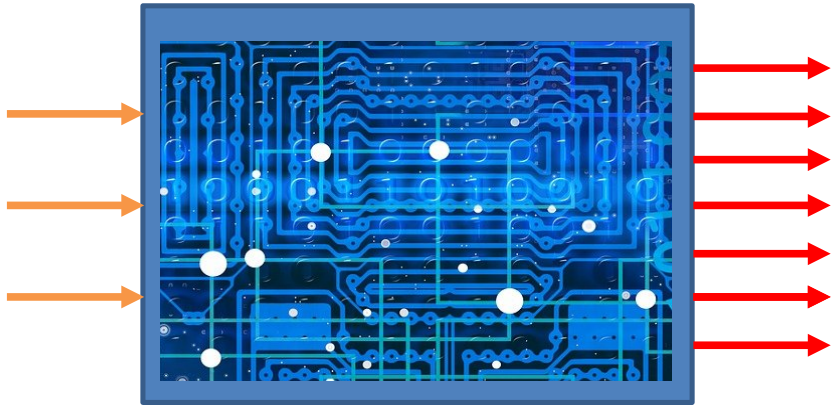
Artificial Neural Networks

- **Artificial Neural Networks (ANNs)**, or simply called **Neural Networks** (**NN**) is a family of Machine Learning systems/models inspired by the human's nervous system, particularly the brain.
- **Neural Network** algorithms try to mimic the brain!

Artificial Neural Networks

- **Neural Networks** are able to learn from data and generate models that get a number of inputs and map them to the desired outputs in an optimal way.

ANN



The Brain

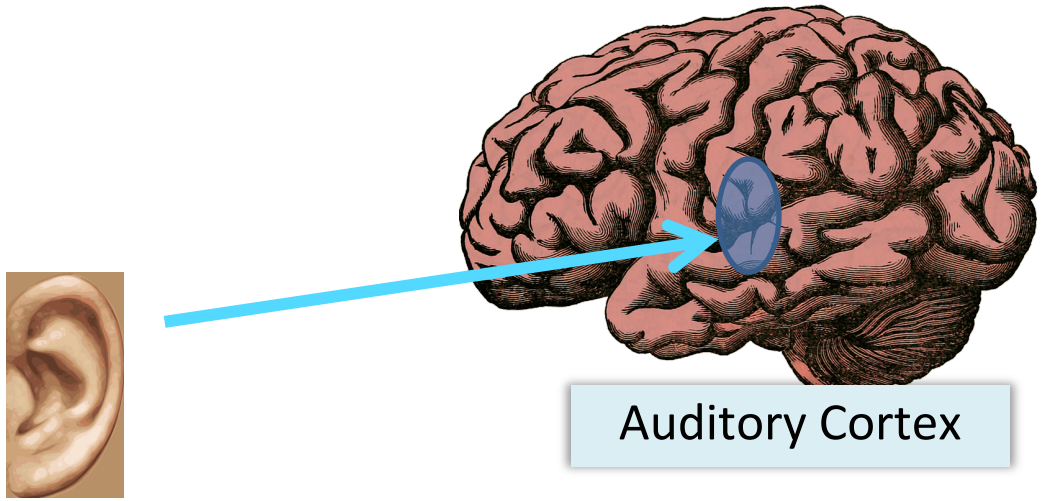
Artificial Neural Networks

- The human brain can **learn so many** different things, makes **so many** amazing decisions, and generates millions of output commands based on various type of input data.

- So, **to mimic the brain**, do we need to write thousands of different algorithms to learn and preform all of those tasks?

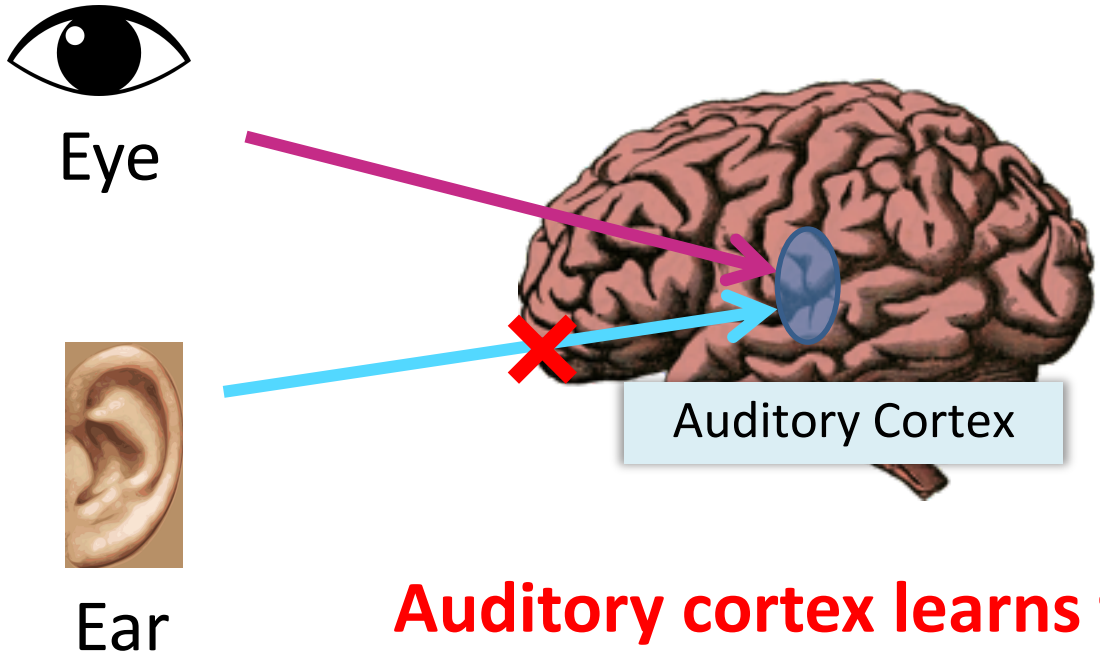
- Well…, scientists discovered that all different things are learned using one single general algorithm! But, how?

The Fascinating Functionality of Brain



Ref: Roe et al., 1992
Ref: Andrew Ng, Machine Learning, Stanford University

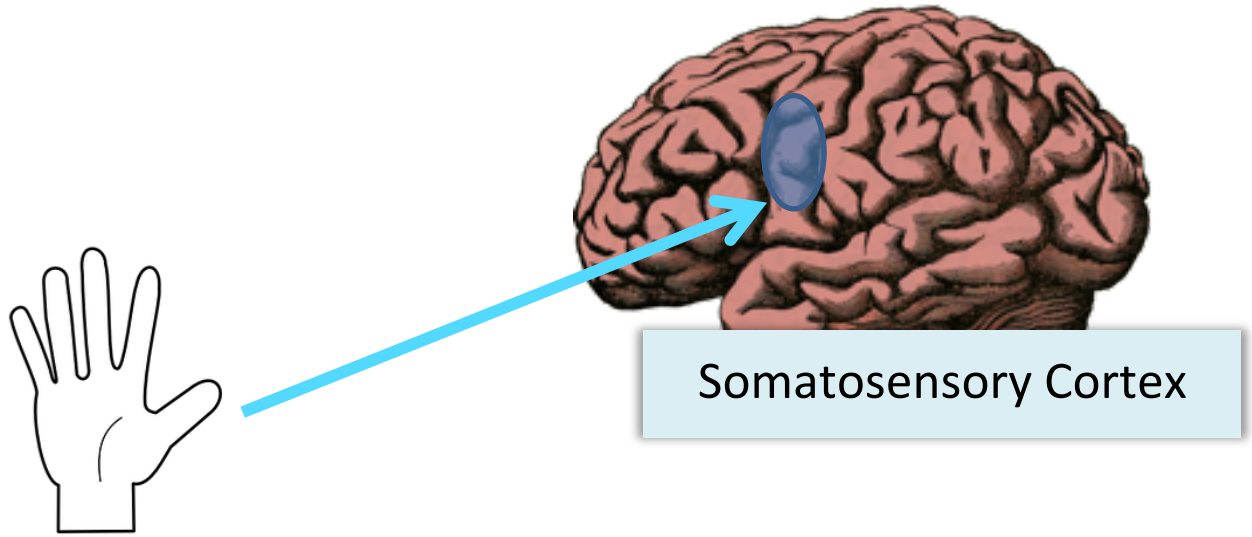
The Fascinating Functionality of Brain



Auditory cortex learns to see!!!

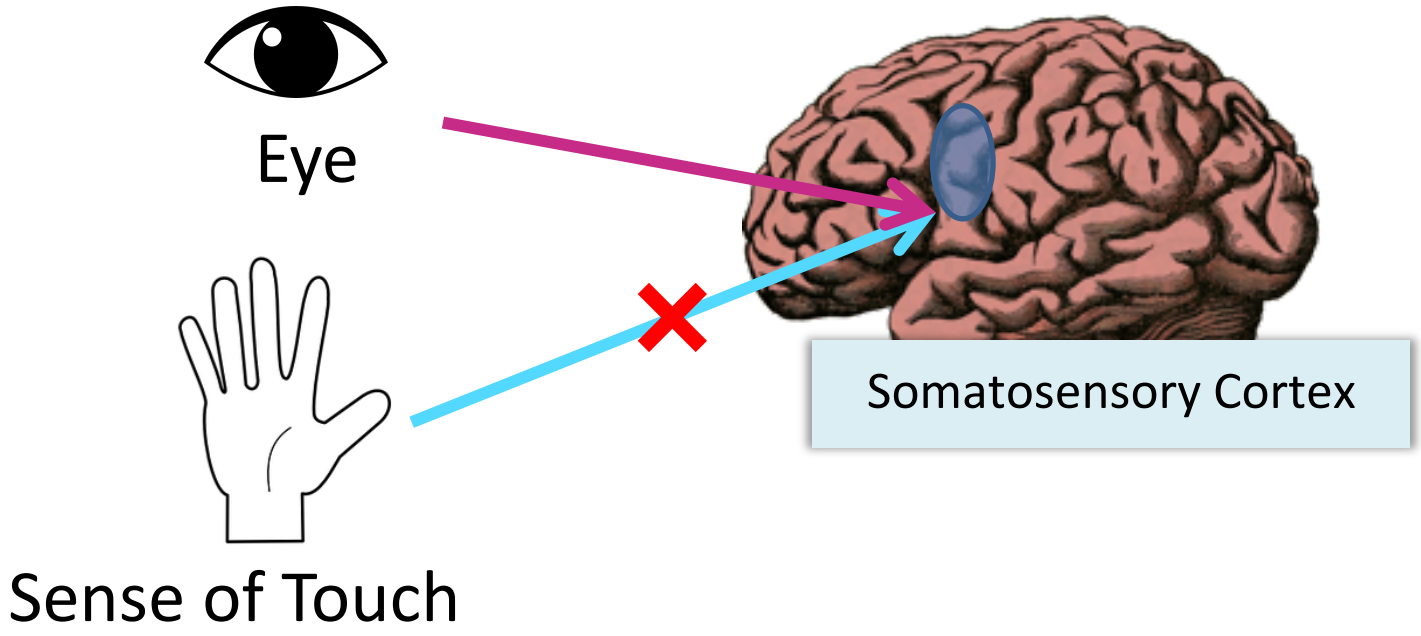
Ref: Roe et al., 1992
Ref: Andrew Ng, Machine Learning, Stanford University

The Fascinating Functionality of Brain



Ref: Roe et al., 1992
Ref: Andrew Ng, Machine Learning, Stanford University

The Fascinating Functionality of Brain



Somatosensory cortex learns to see!!!

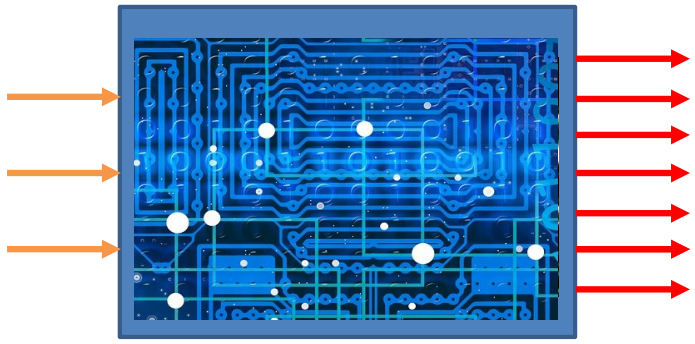
Ref: Roe et al., 1992
Ref: Andrew Ng, Machine Learning, Stanford University

- The Brain learns by wiring **Billions of neurons** together to build an extremely complicated **neural network**.
- **Electrical Signals**: Input Sensor Signals, Output Command Signals, Thoughts, Dreams, …!

How To Mimic The Human Brain?

- **Approach 1:** By extracting a blue print of the whole human brain, and then trying to reconstruct and simulate it.

- **Approach 2:** By simulating a **Neuron** as the brain simplest element first, and then generate a complex Artificial Neural Networks.



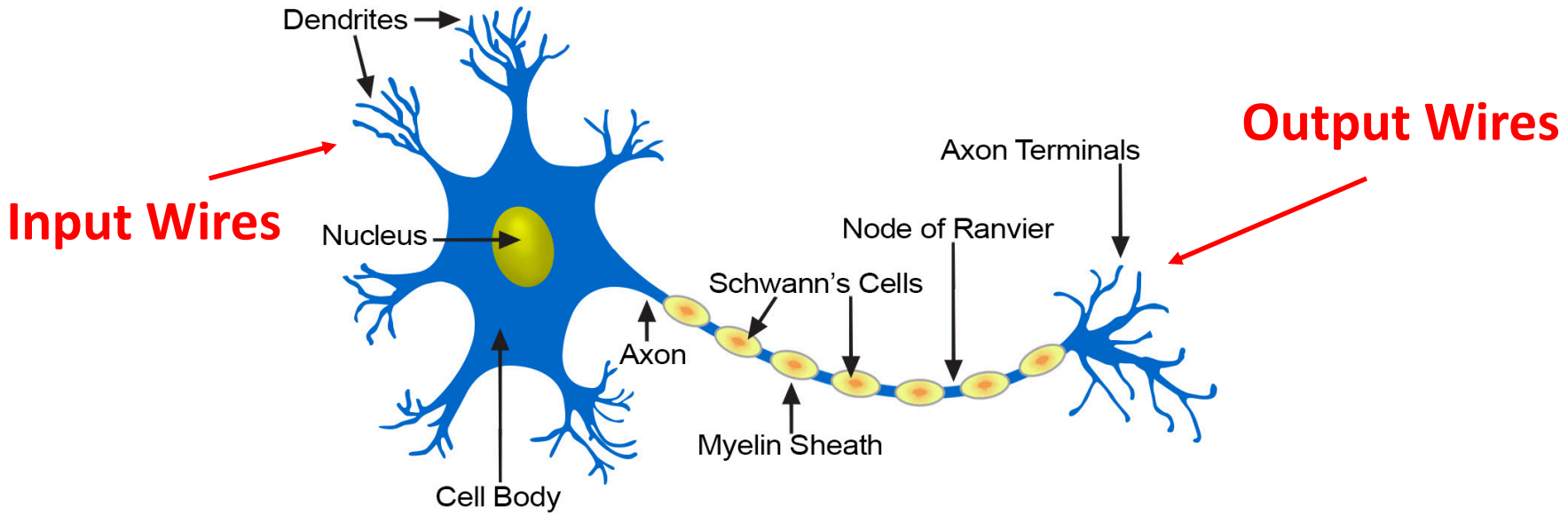
The Blue and Human Brain Projects

- **The Human Brain Projects (HBP)**
  - A large ten-year scientific research project stared in 2013:
  - https://www.youtube.com/watch?v=hm4XK02dFlU

- **Prof. Henry Markram's Talk (Blue Brain Project started 2005):**
  - https://www.youtube.com/watch?v=a1XcY-xAvos

Neurons

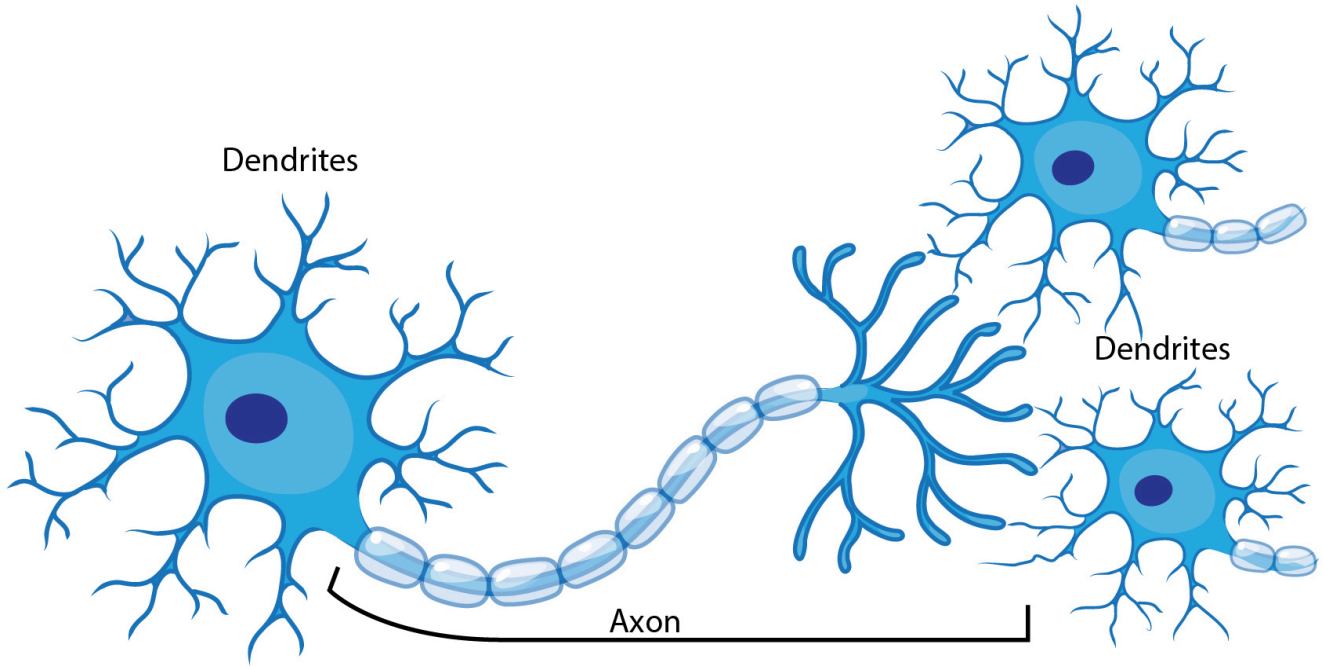
- **Artificial Neural Networks** are developed by **simulating of Neurons** (nerve cells), and then **Network of Neurons** in the human nervous system.



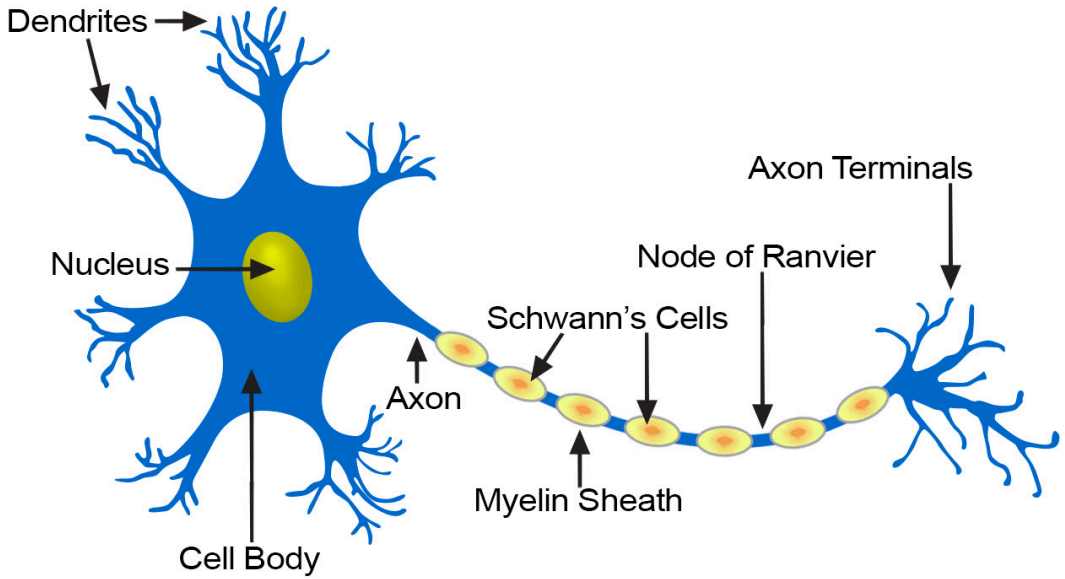
[Image ref]: National Institute of Health (NIH)

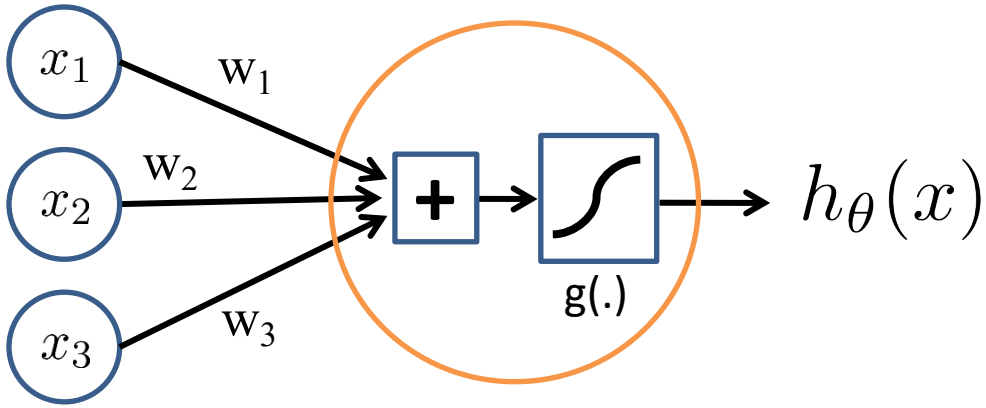
Neurons Communications

- The average human brain includes about **100 billion neurons:**



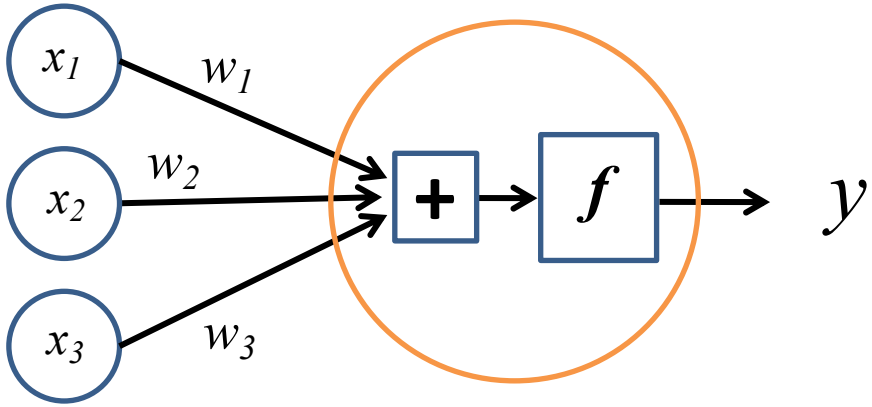
First Step in ANN: Simulating a Neuron





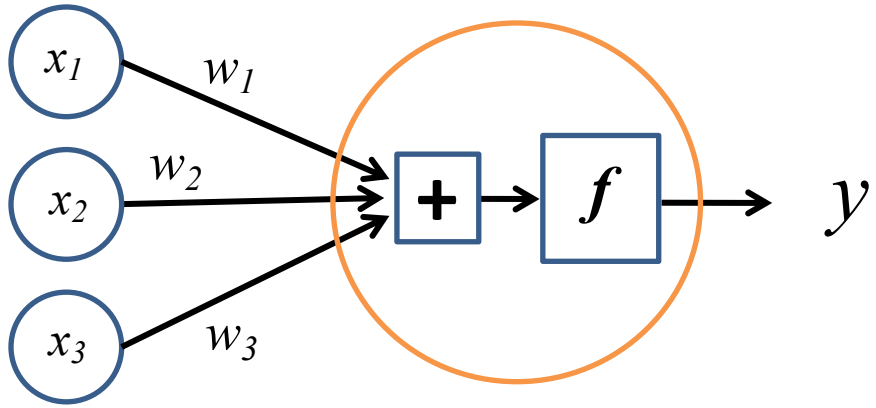
Neural Network Unit

- **Neural Network Unit** (also called **Node** or **Artificial Neuron**) is the smallest element of NN that tries to mimic a brain neuron:



Neural Network Unit

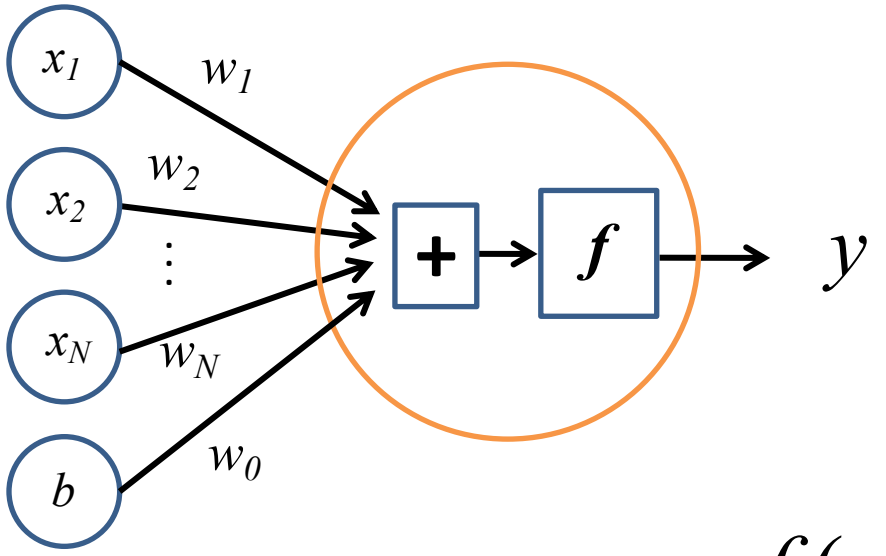
- The early model of an **artificial neuron** is introduced by Warren McCulloch and Walter Pitts. This model includes two main parts: In the first stage, the ***linear combination*** of inputs is calculated. In the second stage, the sum-of-product is passed into a ***non-linear function (f)*** called ***activation function*** which generates the output from the neuron.



$$y = f(x_1 w_1 + x_2 w_2 + x_3 w_3 + \ldots)$$

Neural Network Unit

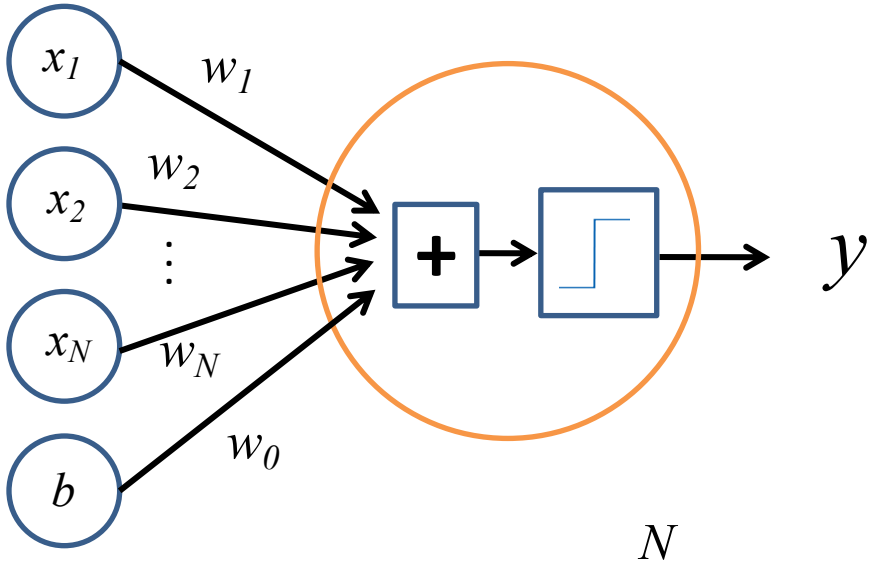
- In general, we may have **many inputs** (x_1 ,…, x_N).
- In addition to the weighted inputs, we can add an extra element that represents **Bias** (b).

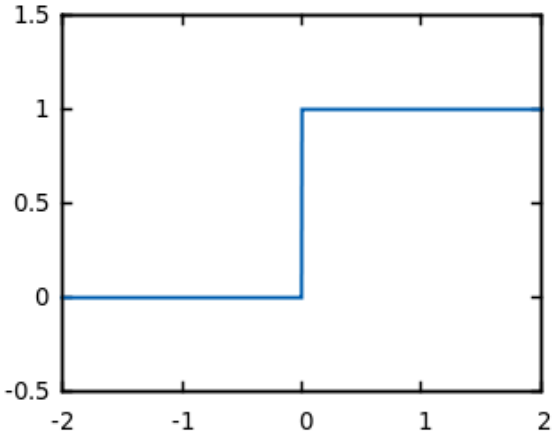


$$y = f(w_0 b + \sum_{i=1}^{N} x_i w_i)$$

Neural Network Unit

- f should be a non-linear function. The simplest function that can be used is a ***step function***.



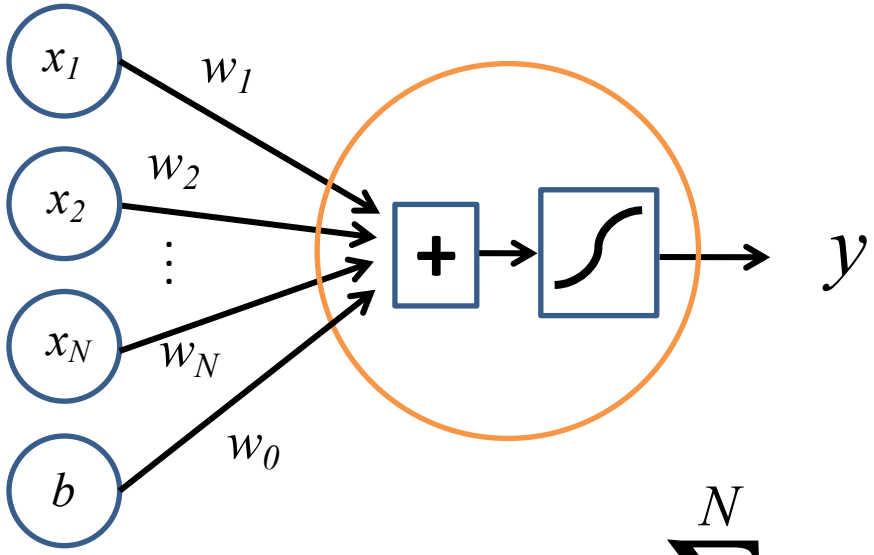


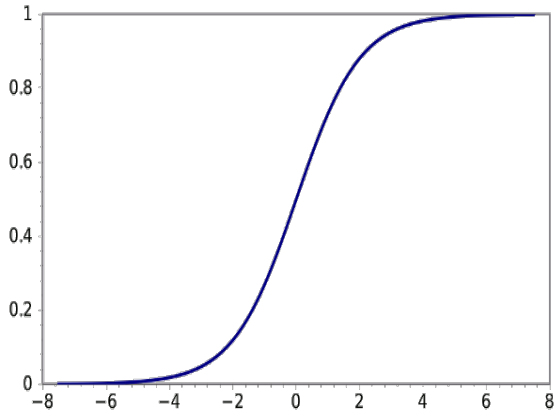
$y = f(x)$

$$y = f(w_0 b + \sum_{i=1}^{N} x_i w_i)$$

Neural Network Unit

- A popular choice for f is ***logistic sigmoid function***.





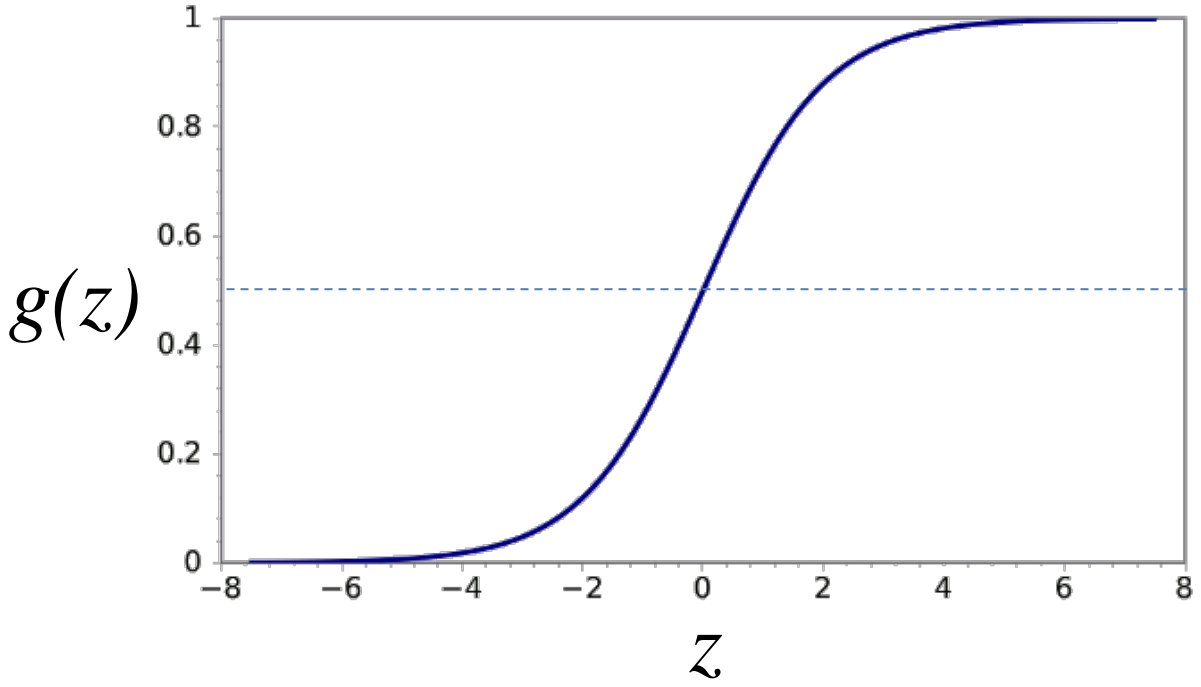
$y = g(x)$

$$y = g(w_0 b + \sum_{i=1}^{N} x_i w_i)$$

Sigmoid Function

- ***Sigmoid Function***: As the input z tends to large positive value, the output value $g(z)$ approaches to 1. Similarly, the output gets close to 0 as z goes negative. So, it limits the output in the range of (0,1).

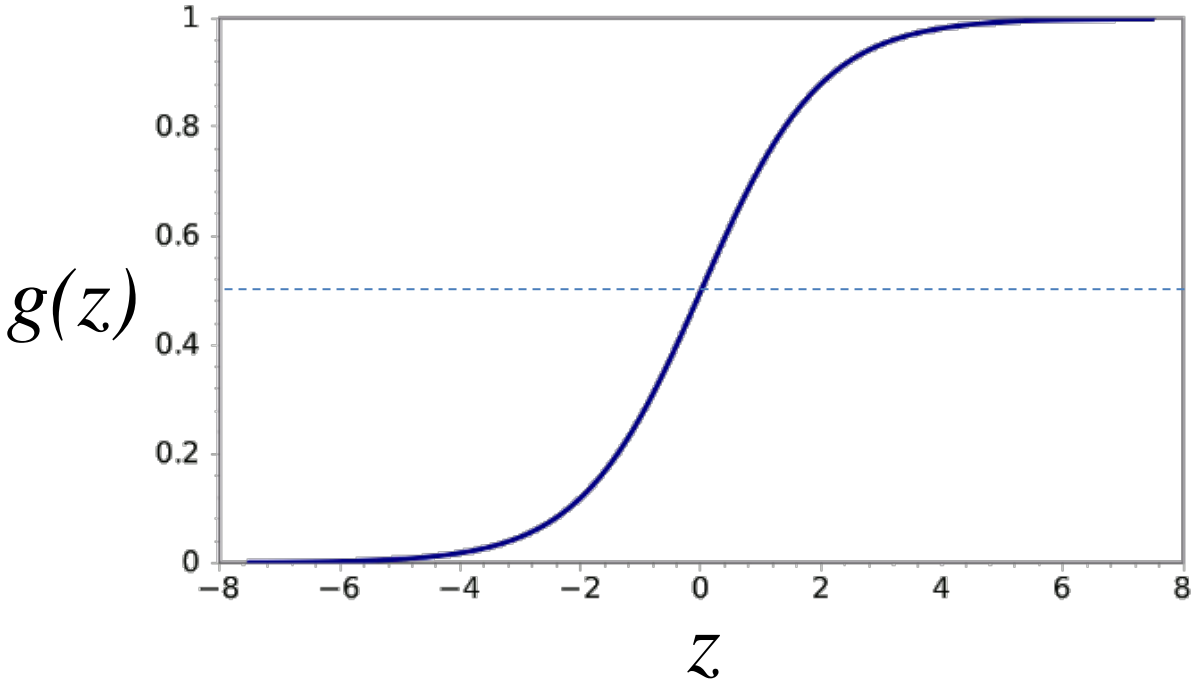
$$g(z) = \frac{1}{1 + e^{-z}}$$



Sigmoid Function

- The sigmoid function adaptively describes the **closeness** to the threshold point by the curve slope. As z approaches to - *infinity* or *+ infinity* , the slope is **zero**; the slope gets larger as z approaches to 0.

$$g(z) = \frac{1}{1 + e^{-z}}$$



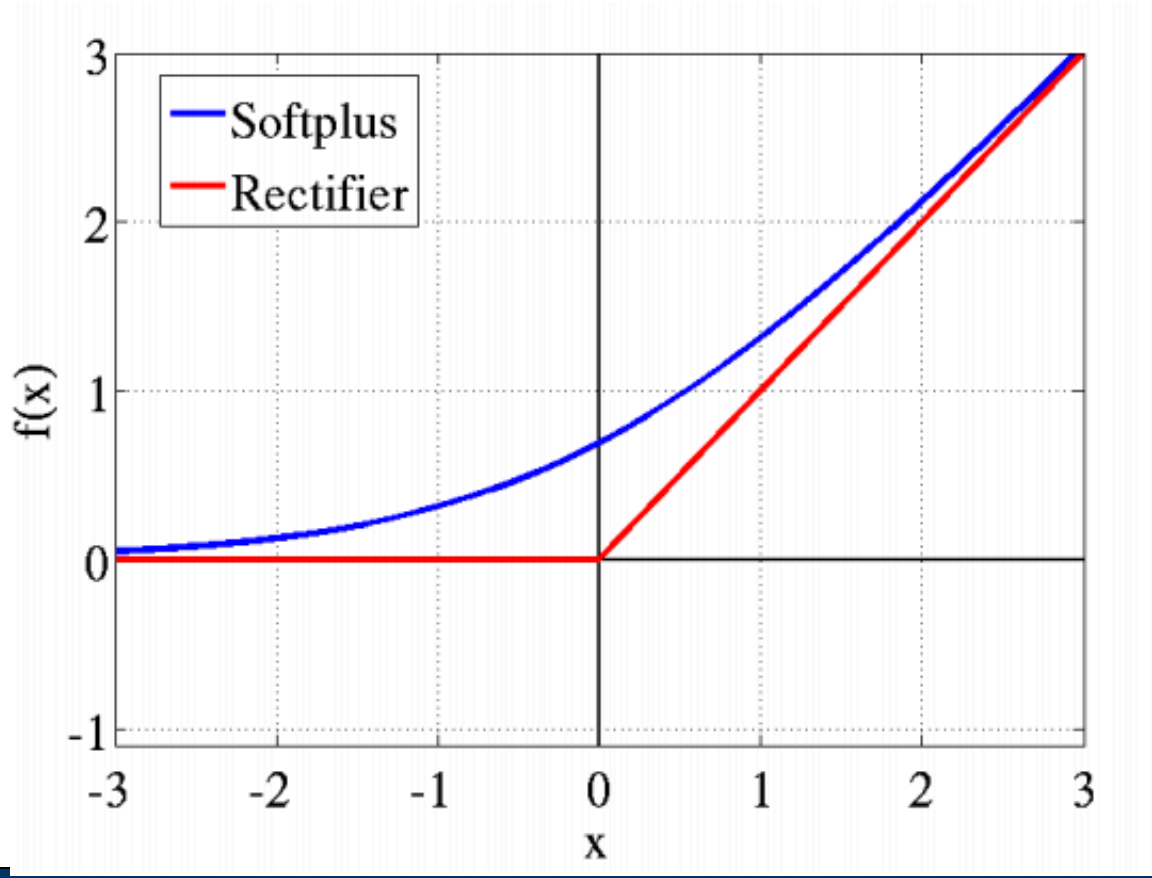
Rectified Linear Unit (ReLU) Function

- Another popular activation functions specially in Deep Neural Networks is **ReLU** (also called Rectifier, or Rectified Linear Unit), and also the smooth approximation of ReLU, which is called **Softplus**.

$$RELU(x) = \begin{cases} 0 \ if \ x < 0 \\ x \ if \ x >= 0 \end{cases}$$

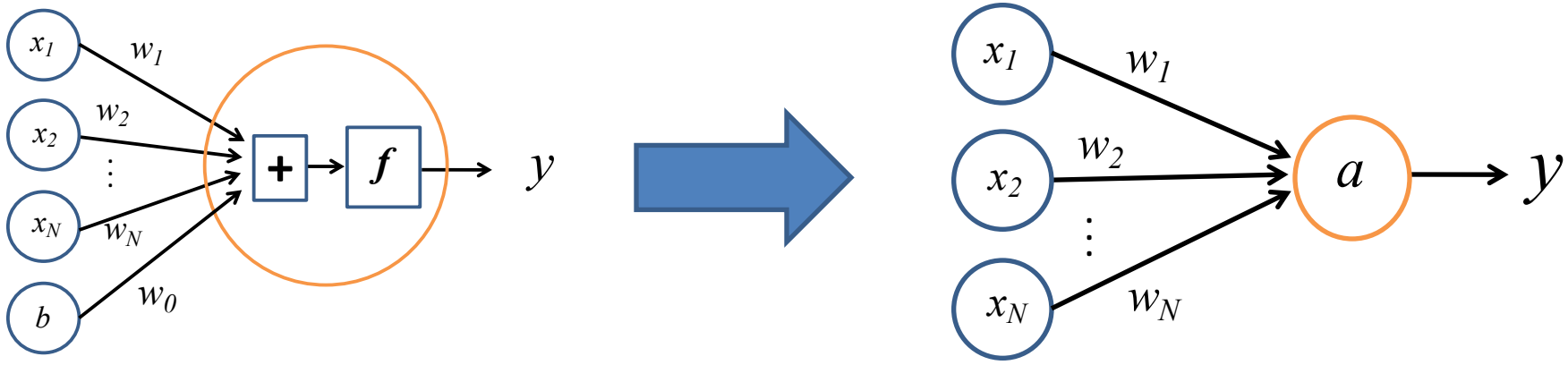
$$\text{RELU}(x) = \max(0, x)$$

$$\text{Softplus}(x) = \log(1+\exp(x))$$



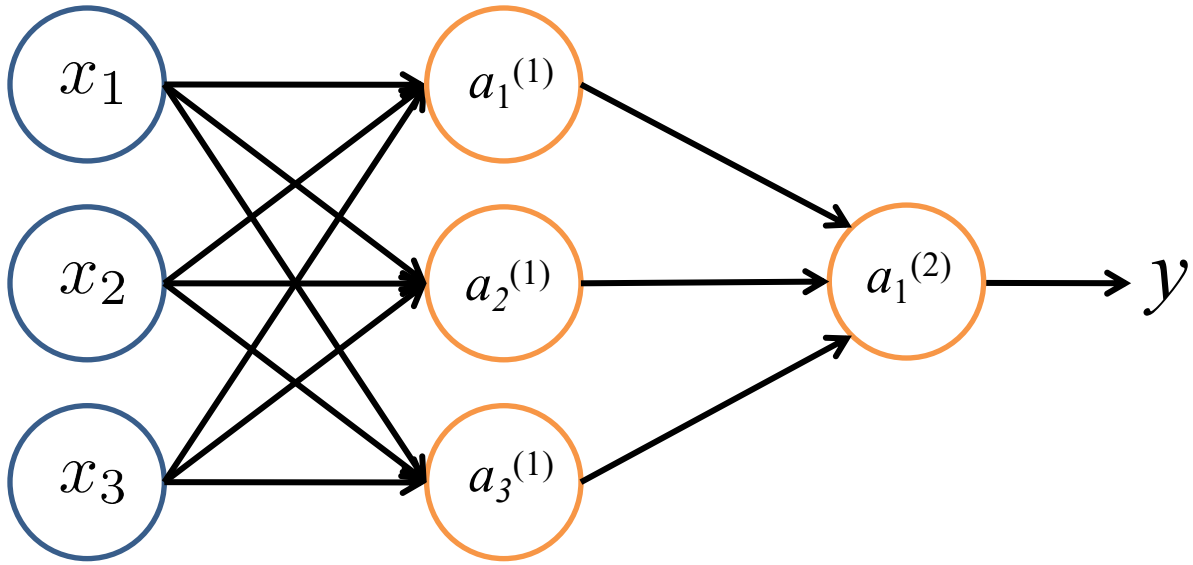
Representation

- **From now on, for the sake of simplicity, we use the following figure to represent a Unit (a neuron node):**



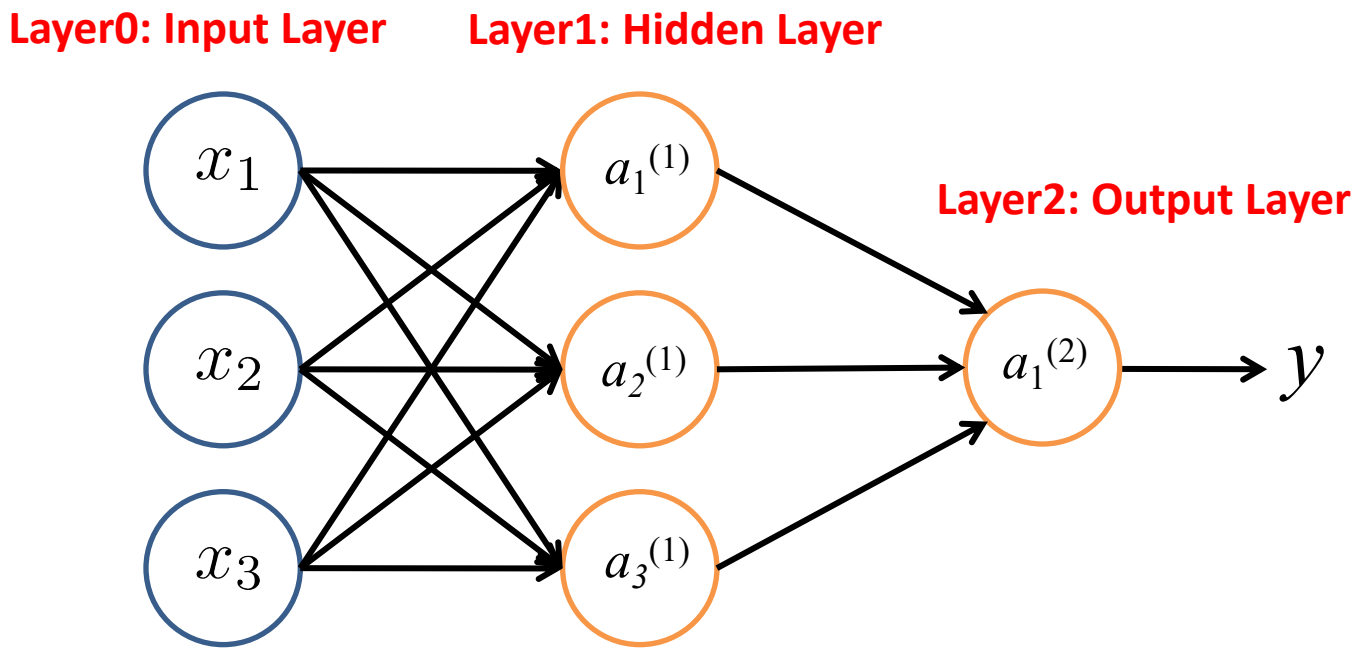
Neural Networks

- Similar to the human brain, an artificial neural network is formed by connecting multiple artificial neurons.
- The most common structure of connecting neurons into a network is by **layers**. The simplest form of layered network is shown in this figure:



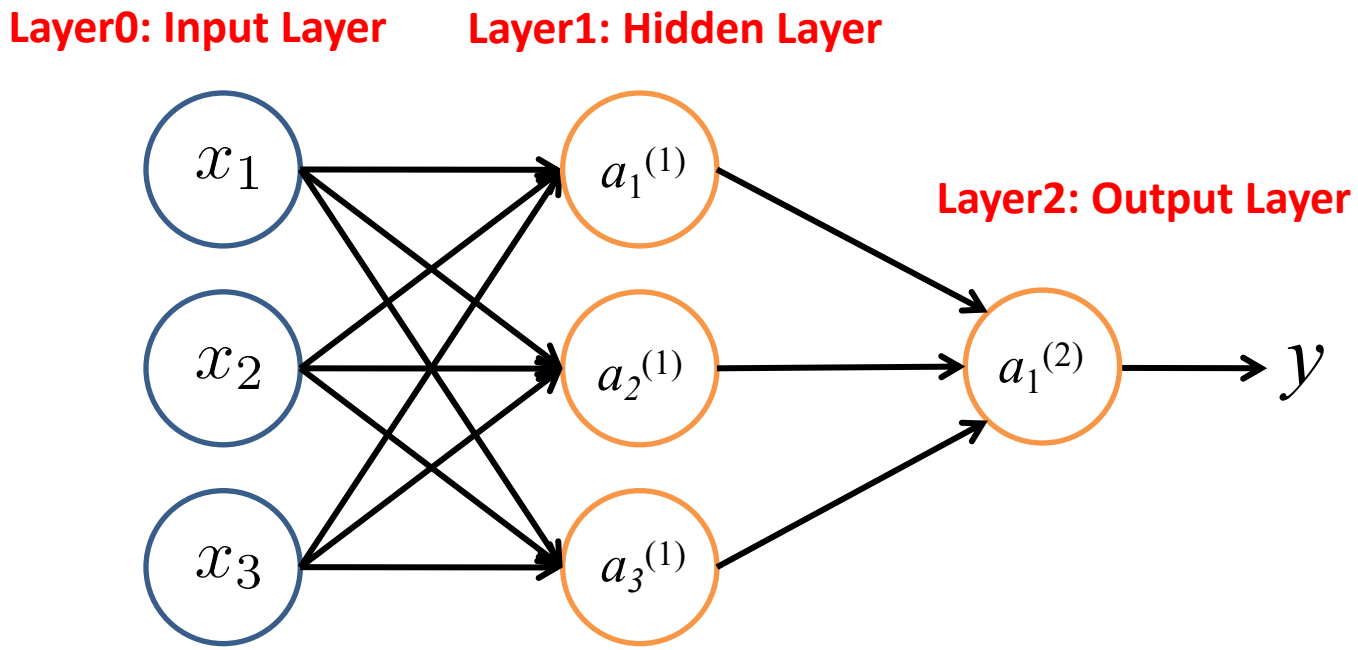
Neural Networks

- The first layer is called **Input Layer**. The input layer only passes and distribute the inputs and perform **no computation**.
- The last layer is called **Output Layer**.
- The middle layers are called **Hidden Layers**.



Neural Networks

- Each of the inputs is connected to every artificial neuron through a ***weighted*** connection.
- Although the presented network is ***fully connected***, the true biological neural network may not have all possible connections. In our model, the weight value of **zero** can be represented as "no connection".



Thank You!

Questions?