Neural Networks

E.Milgo

Artificial Neural Networks

- **Neural network:** *information processing paradigm inspired by biological nervous systems, such as our brain*
- Structure: large number of highly interconnected processing elements (*neurons*) working together
- Like people, they learn *from experience* (by example)

“Data processing system consisting of a large number of simple, highly interconnected processing elements (artificial neurons) in an architecture inspired by the structure of the cerebral cortex of the brain”

(Tsoukalas & Uhrig, 1997).

History

- **1943:**Warren McCulloch & Walter Pitts's first mathematical model of NN
- **1949**: Donald Hebb
  - "Hebbian learning rule
- **1950s-1960s**: Rosenblatt
  - Perceptron
- **The "AI Winter" and Resurgence (1960s-1980s):**
- **1974**: Paul Werbos introduced the backpropagation algorithm
- **1980s-1990s**: Yann LeCun & others
  - Convolutional neural networks (CNNs)
- **1982**: John Hopfield
  - Hopfield networks
- **2000s-Present**: Deep learning
  - Advancements in computing power, GPUs,
- **2006**: Geoffrey Hinton
  - deep belief networks
- **2012**: AlexNet
  - a deep convolutional neural network
- **2010s-Present**:
  - Deep learning in various fields

Neural Network

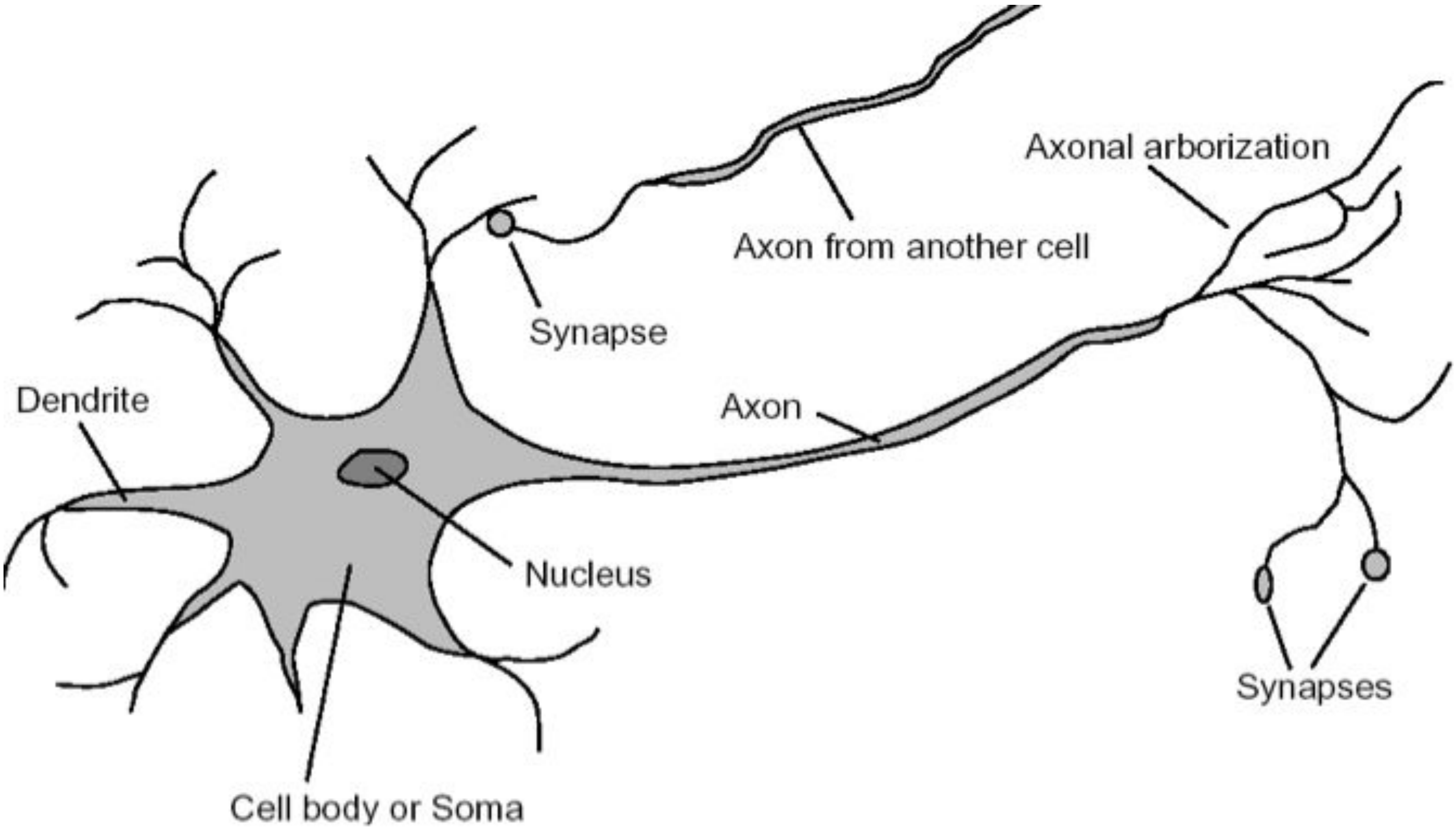
- A neural network is a massively parallel, distributed processor
- made up of simple processing units (artificial neurons).
- It resembles the brain in two respects:
- – Knowledge is acquired by the network from its
- environment through a learning process
- – Synaptic connection strengths among neurons are used to
- store the acquired knowledge.

- We are born with about 100 billion neurons
- A neuron may connect to as many as 100,000 other neurons
- Signals “move” via electrochemical signals

- The synapses release a chemical transmitter – the sum of which can cause a threshold to be reached – causing the neuron to “fire”

- Synapses can be inhibitory or excitatory

Human Biological Neural



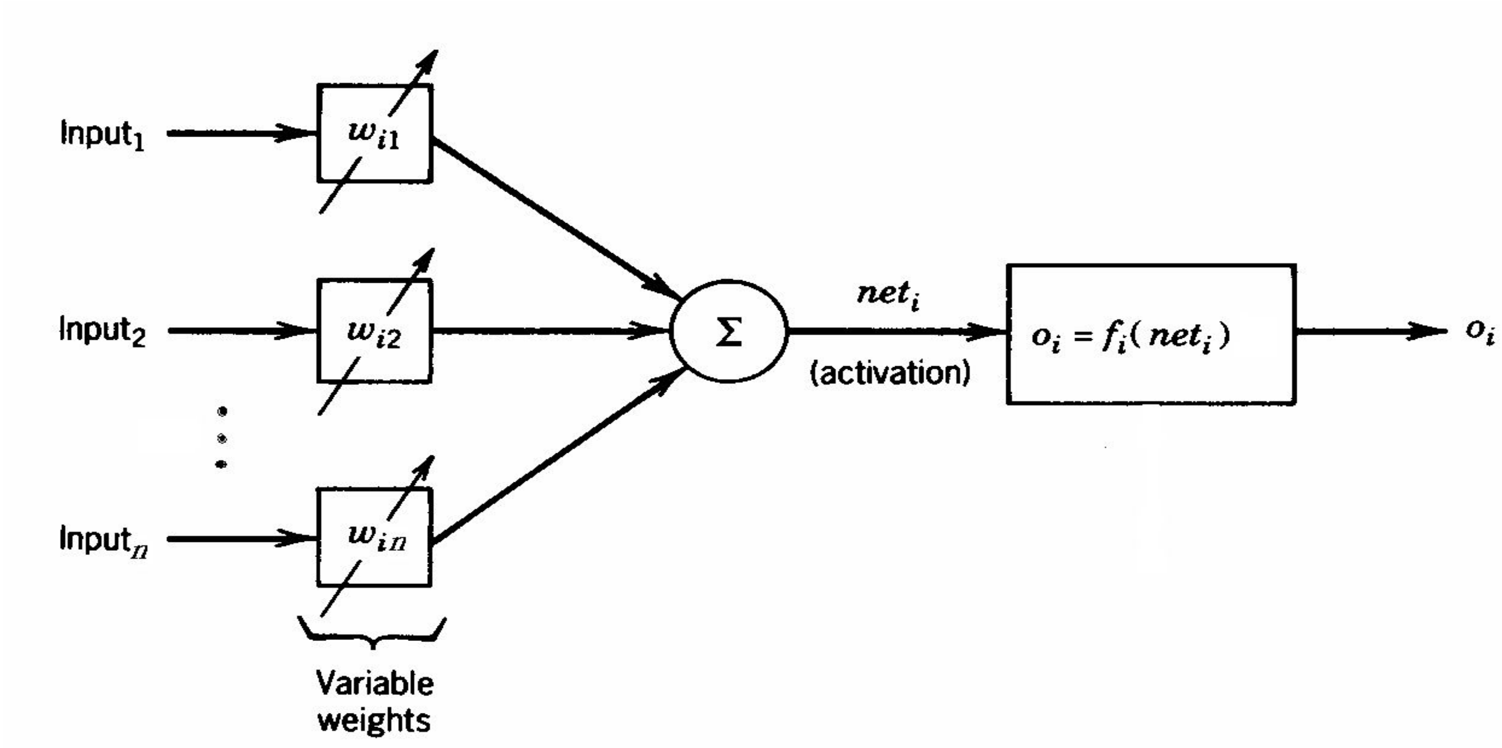
The Biological Neuron

- A biological neuron has three types of main components; **dendrites, soma** (or cell body) and **axon**.
- Dendrites receives signals from other neurons.
- The soma, sums the incoming signals. When sufficient input is received, the cell fires; that is it transmit a signal over its axon to other cells.

- A neuron only fires if its input signal exceeds a certain amount (the **threshold**) in a short time period.
- Synapses vary in strength
  - Good connections allowing a large signal
  - Slight connections allow only a weak signal.

- Each neuron has a threshold value
- Each neuron has weighted inputs from other neurons
- The input signals form a weighted sum
- If the activation level exceeds the threshold, the neuron “fires”

The Artificial Neuron



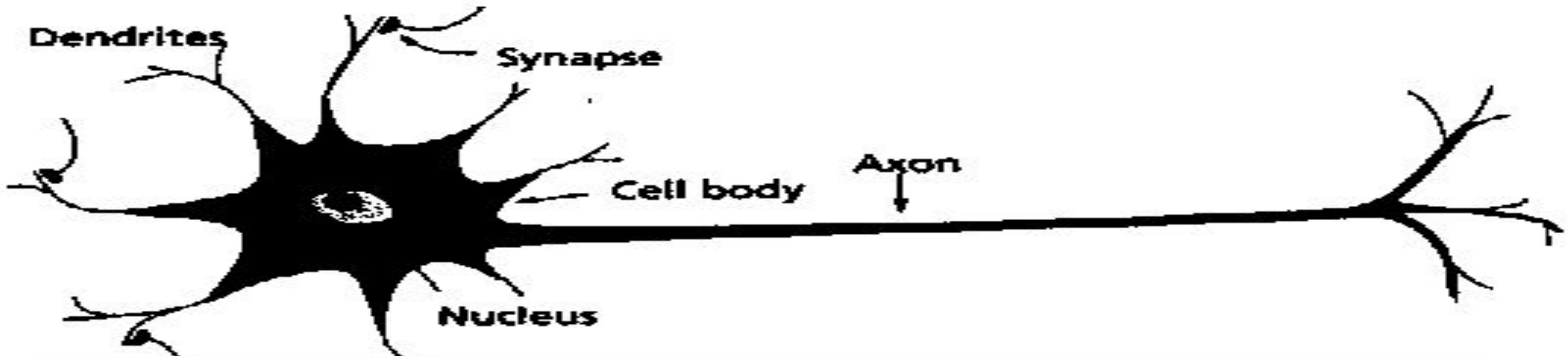
Artificial Neuron

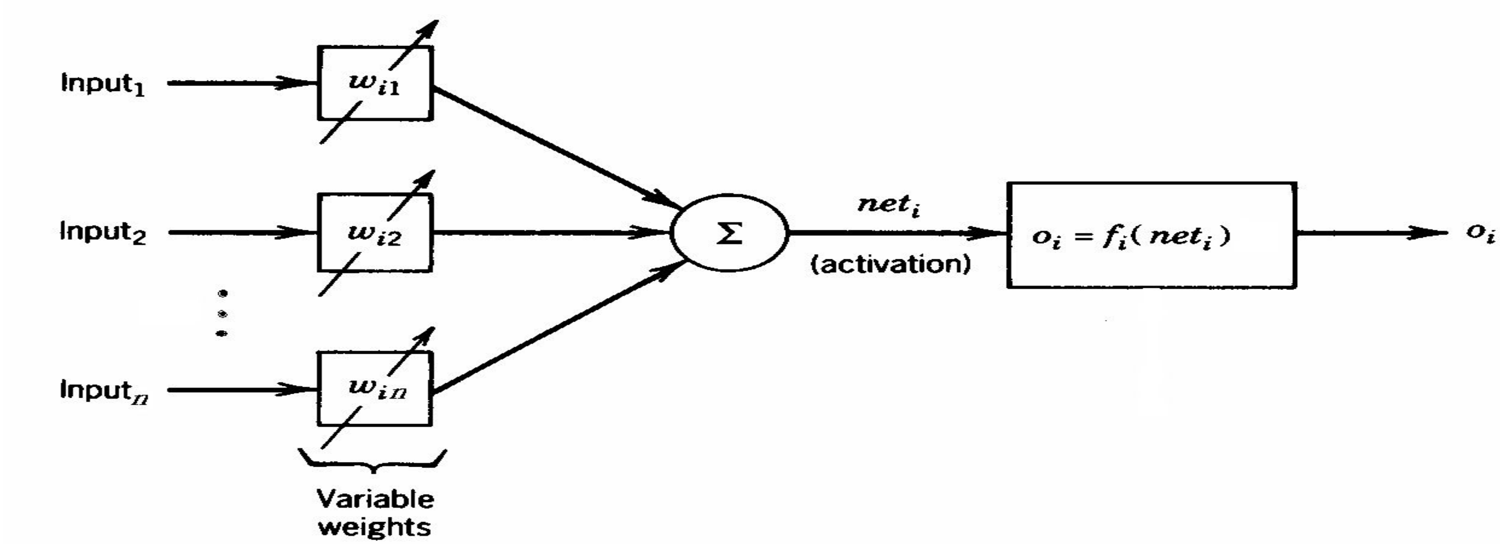
- Each hidden or output neuron has weighted input connections from each of the units in the preceding layer.
- The unit performs a weighted sum of its inputs, and subtracts its threshold value, to give its activation level.
- Activation level is passed through a sigmoid activation function to determine output.

Weights in NN

- The strength of connection between the neurons is stored as a weight-value for the specific connection.
- Learning the solution to a problem involves changing the connection weights

Biological Vs Artificial Neuron





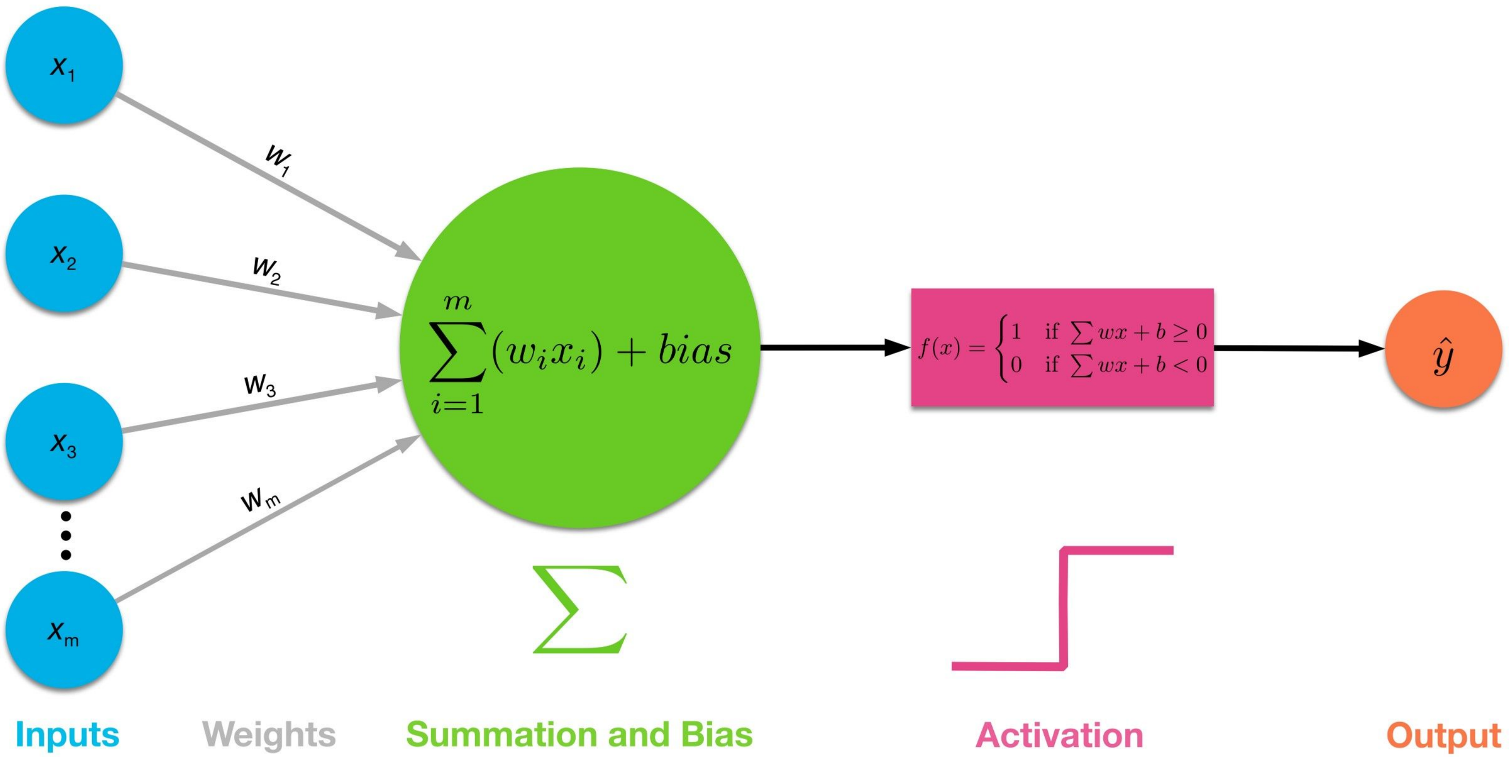
Neuron Net

- A neural net consists of a large number of simple processing elements called **neurons, units, cells or nodes.**

- Each neuron is connected to other neurons by means of directed communication links, each with **associated weight**.

- The weight represent information being used by the net to solve a problem.

A Neuron Model



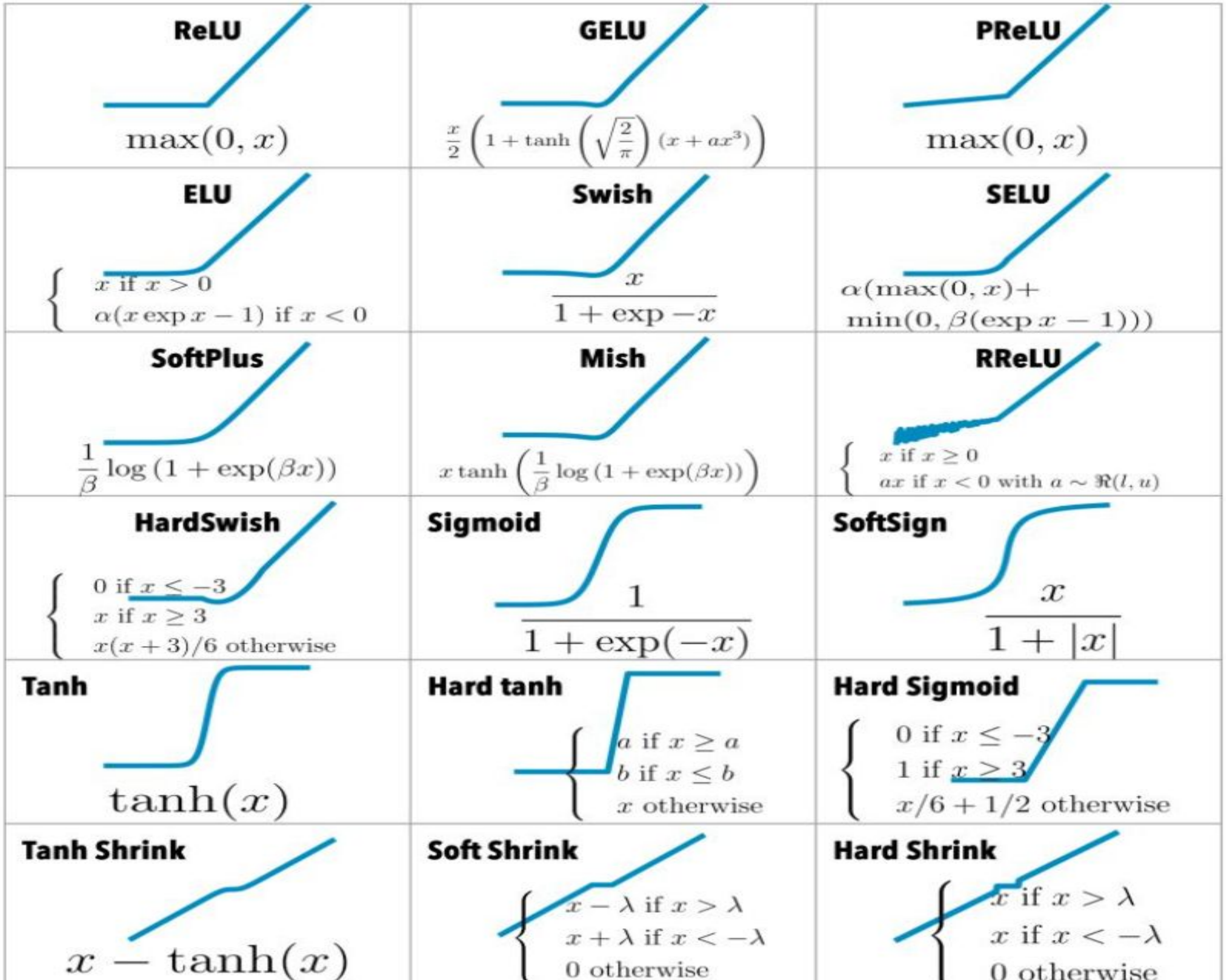
Activation Function

Neural networks use non-linear activation functions, which help the network learn complex data, compute and learn almost any function, and provide accurate predictions.

Common Activation Functions

1. **Sigmoid:** Outputs values between 0 and 1, often used in the output layer for binary classification.
2. **Tanh:** Outputs values between -1 and 1, zero-centered, which can help with faster convergence during training.
3. **ReLU (Rectified Linear Unit):** Outputs the input directly if positive, otherwise outputs 0. A popular choice for hidden layers due to its simplicity and ability to mitigate the vanishing gradient problem.
4. **Softmax:** Converts a vector of numbers into a probability distribution, commonly used in the output layer for multi-class classification.

Neural Network Activation Functions: a small subset!



source:TheAiEdge,io

Role of Activation Functions

- **Introduce non-linearity:** transform the weighted sum of inputs into a usable output, enabling the network to learn non-linear relationships in data.
- **Control neuron activation:** determine whether a neuron should "fire"
- **Normalize output:** Some activation functions, like sigmoid and tanh, squash the output into a specific range (e.g., 0-1 or -1 to 1), which can help with stability during training.

Choosing an Activation Function

The choice of activation function depends on the specific task

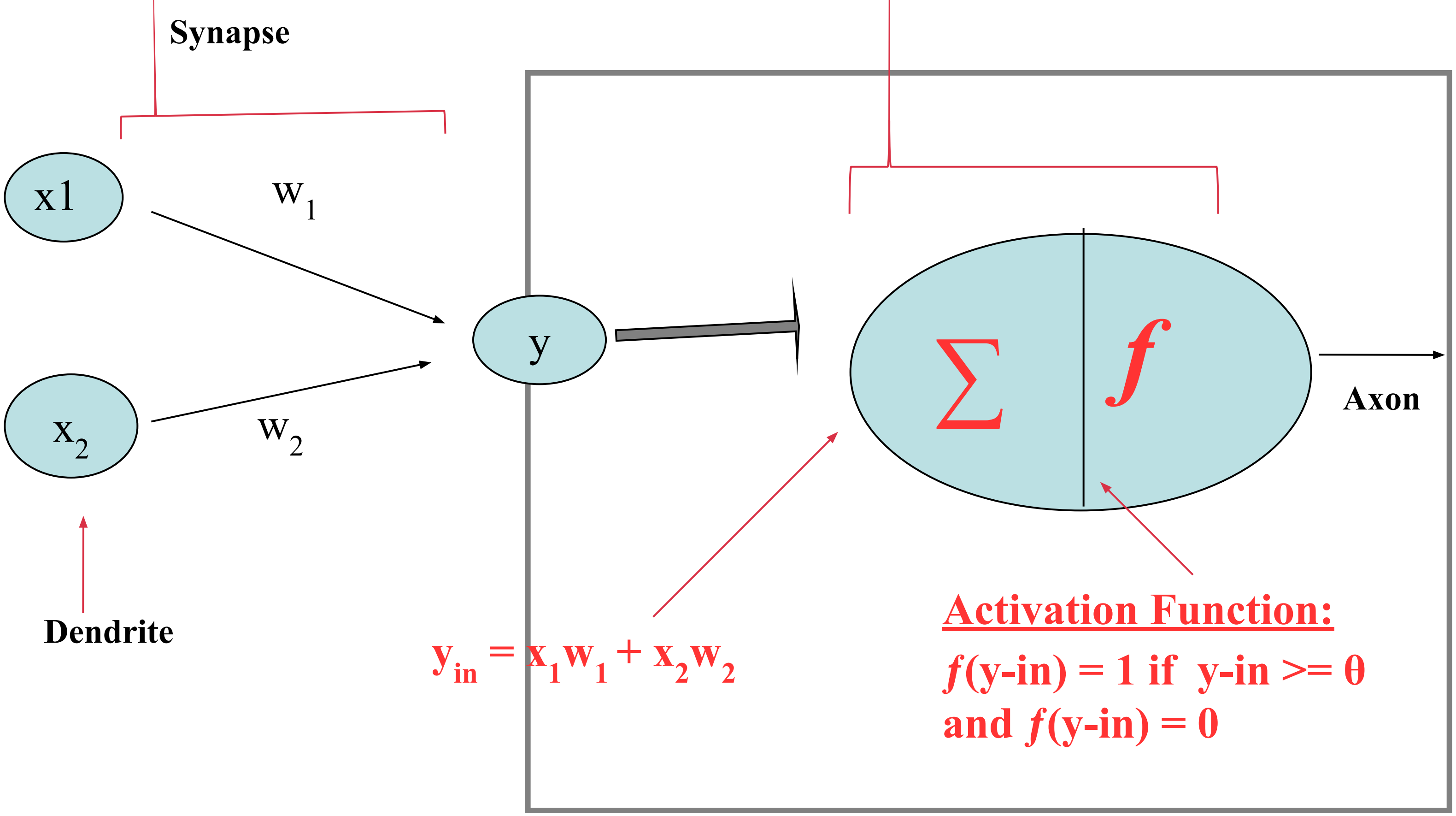
- **Sigmoid** and **tanh** are often used in shallow networks or for specific tasks like binary classification,
- **Rectified Linear Unit (ReLU)** is a popular activation functions in deep learning models. The ReLU function is a piecewise linear function that outputs the input directly if it is positive; otherwise, it outputs zero. ReLU helps mitigate the **vanishing gradient problem**, which can hinder the training of deep neural networks.
- **Softmax** is used for multi-class classification tasks, while linear activation functions are suitable for linear regression problems.

- Each neuron has an internal state, called its **activation or activity level**, which is a function of the inputs it has received. Typically, a neuron sends its activation as a signal to several other neurons.

- It is important to note that a neuron can send only one signal at a time, although that signal is broadcast to several other neurons.

- Neural networks are configured for a specific application, such as pattern recognition or data classification, through a learning process
- In a biological system, learning involves adjustments to the synaptic connections between neurons same for artificial neural networks

Artificial Neural



- A neuron receives input, determines the strength or the weight of the input, calculates the total weighted input, and compares the total weighted with a value (threshold)
- The value is in the range of 0 and 1
- If the total weighted input is greater than or equal the threshold value, the neuron will produce the output
- If the total weighted input is less than the threshold value, the output will be zero

Types of NN

Single layer Perceptron

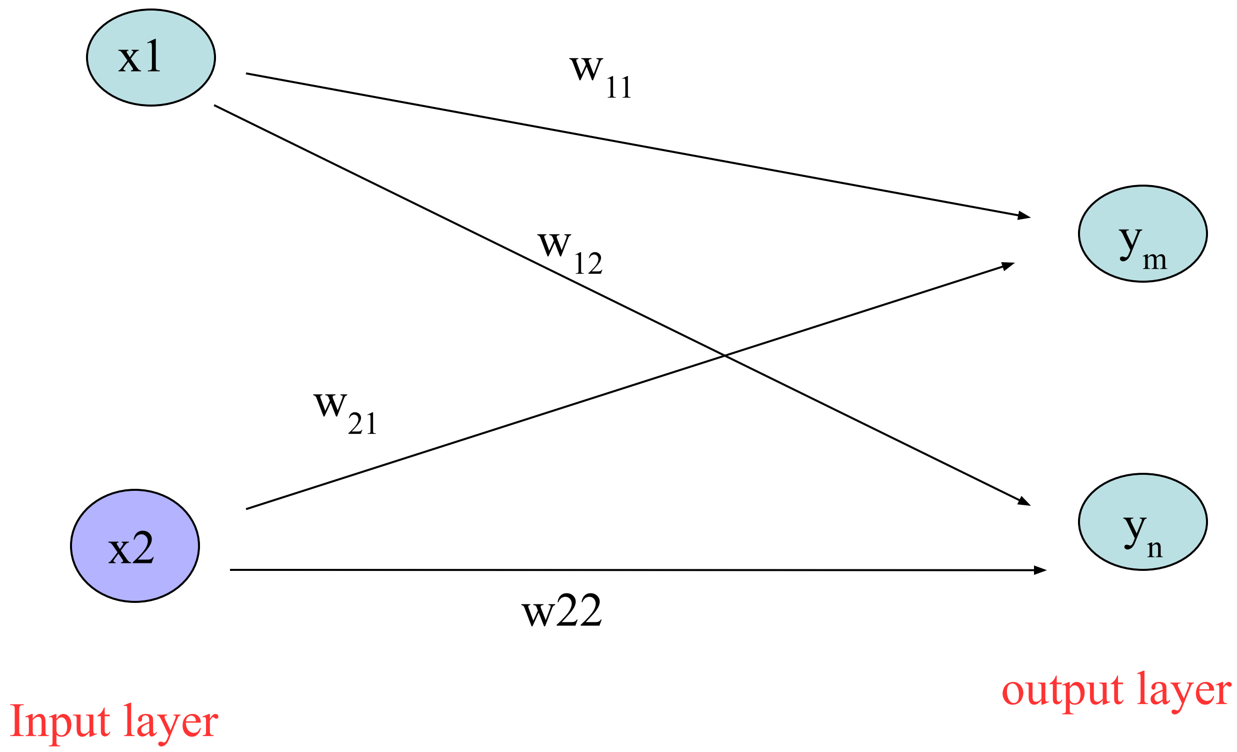
Multi Layer perceptron

Self Organizing maps

Deep learning

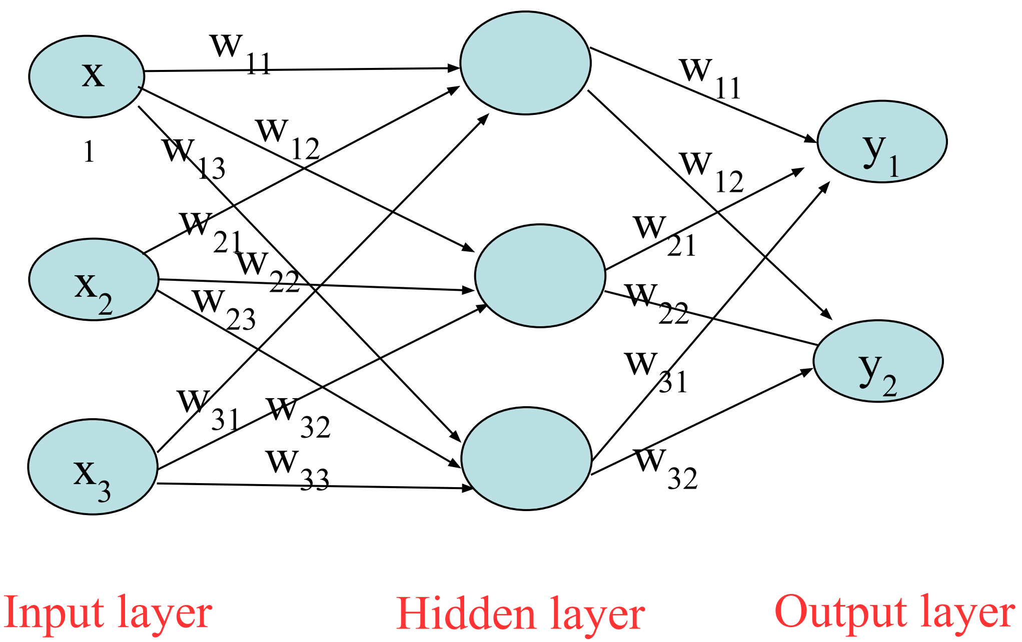
Single Layer Perceptron

Input layer projecting into the output layer



Multi Layer Perceptron

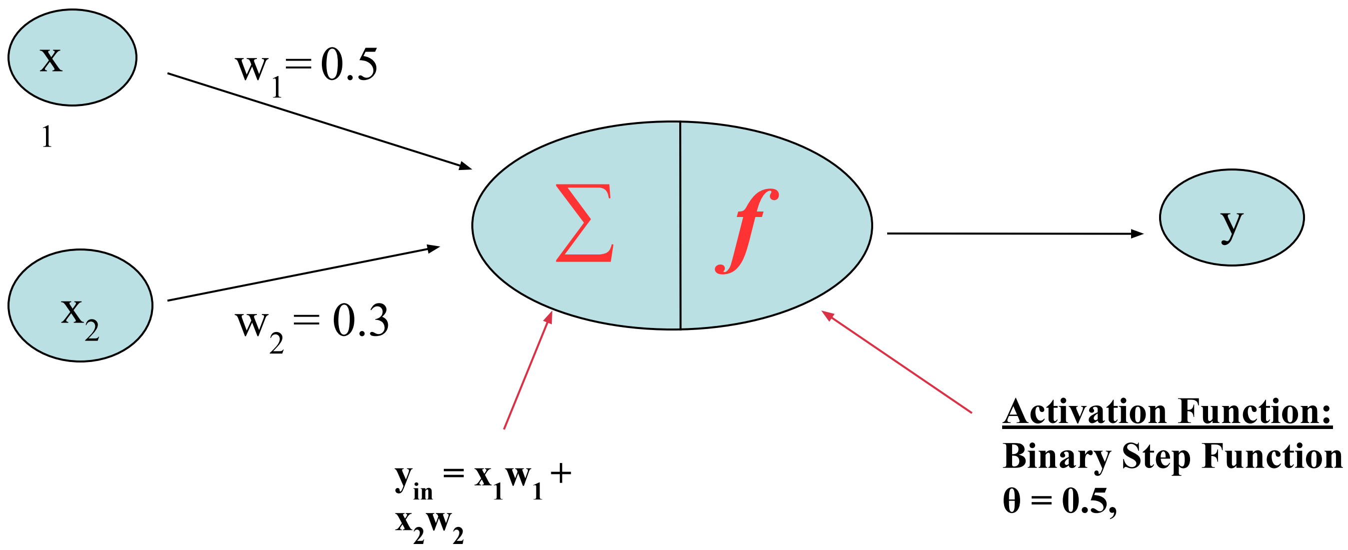
- One or more hidden layers.
- fully connected network



Types of Layers

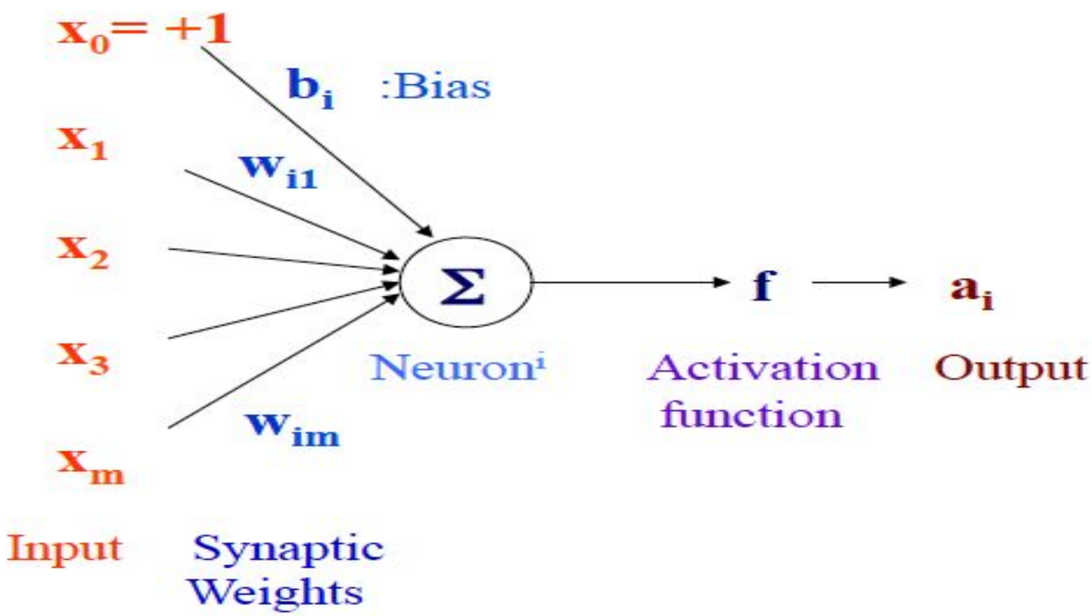
- The **input layer:** Introduces input values into the network
    - No activation function or other processing
- The **hidden layer(s):** Perform classification of features
    - Two hidden layers are sufficient to solve any problem
    - Features imply more layers may be better
- **The output layer:** Functionally just like the hidden layers
    - Outputs are passed on to the world outside the neural network.

Example



x_1
$w_1 = 0.5$
x_2
$w_2 = 0.3$
Σ
f
y
$y_{in} = x_1w_1 + x_2w_2$
<u>Activation Function:</u>
Binary Step Function
$\theta = 0.5$,

NN with bias



$x_0 = +1$
b_i :Bias
x_1
w_{i1}
x_2
Σ
f
a_i
x_3
Neuron^i
Activation function
Output
w_{im}
x_m
Input
Synaptic Weights

Calculating the Bias

An artificial neuron:

- computes the weighted sum of its input (called its net input)

- adds its bias

- passes this value through an activation function

We say that the neuron "fires" (i.e. becomes active) if its output is

NN design

- Number of layers
  - Apparently, three layers is almost always good enough and better than four layers.
  - Also: fewer layers are faster in execution and training
- How many hidden nodes?
  - Many hidden nodes allow to learn more complicated patterns
  - Because of overtraining, almost always best to set the number of hidden nodes too low and then increase their numbers.

NN Architecture

Even for a basic Neural Network, there are many design decisions to make:

1. # of hidden layers (depth)
2. # of units per hidden layer (width)
3. Type of activation function (nonlinearity)
4. Form of objective function

NN application

- Signal processing
- Pattern recognition, e.g. handwritten characters or face identification.
- Diagnosis or mapping symptoms to a medical case.
- Speech recognition
- Human Emotion Detection
- Educational Loan Forecasting