COM 424 E: NEURAL NETWORKS

Lecture 01:Introduction to Artificial Neural Networks (ANN)

E.M

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)

Definition of ANN

"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: McCulloch and Pitts model neural networks based on their understanding of neurology.

- 1950s: Farley and Clark
  - » IBM group that tries to model biological behavior

- Perceptron (Rosenblatt 1958)
  - Association units A_1, A_2, … extract features from user input
  - Output is weighted and associated
  - Function fires if weighted sum of input exceeds a threshold.

History

- Back-propagation learning method (Werbos 1974)
  - Three layers of neurons
    - Input, Output, Hidden
  - Better learning rule for generic three layer networks
  - Regenerates interest in the 1980s
- Successful applications in medicine, marketing, risk management, … (1990)
- In need for another breakthrough.

Neural Networks

- 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.

Neural Networks

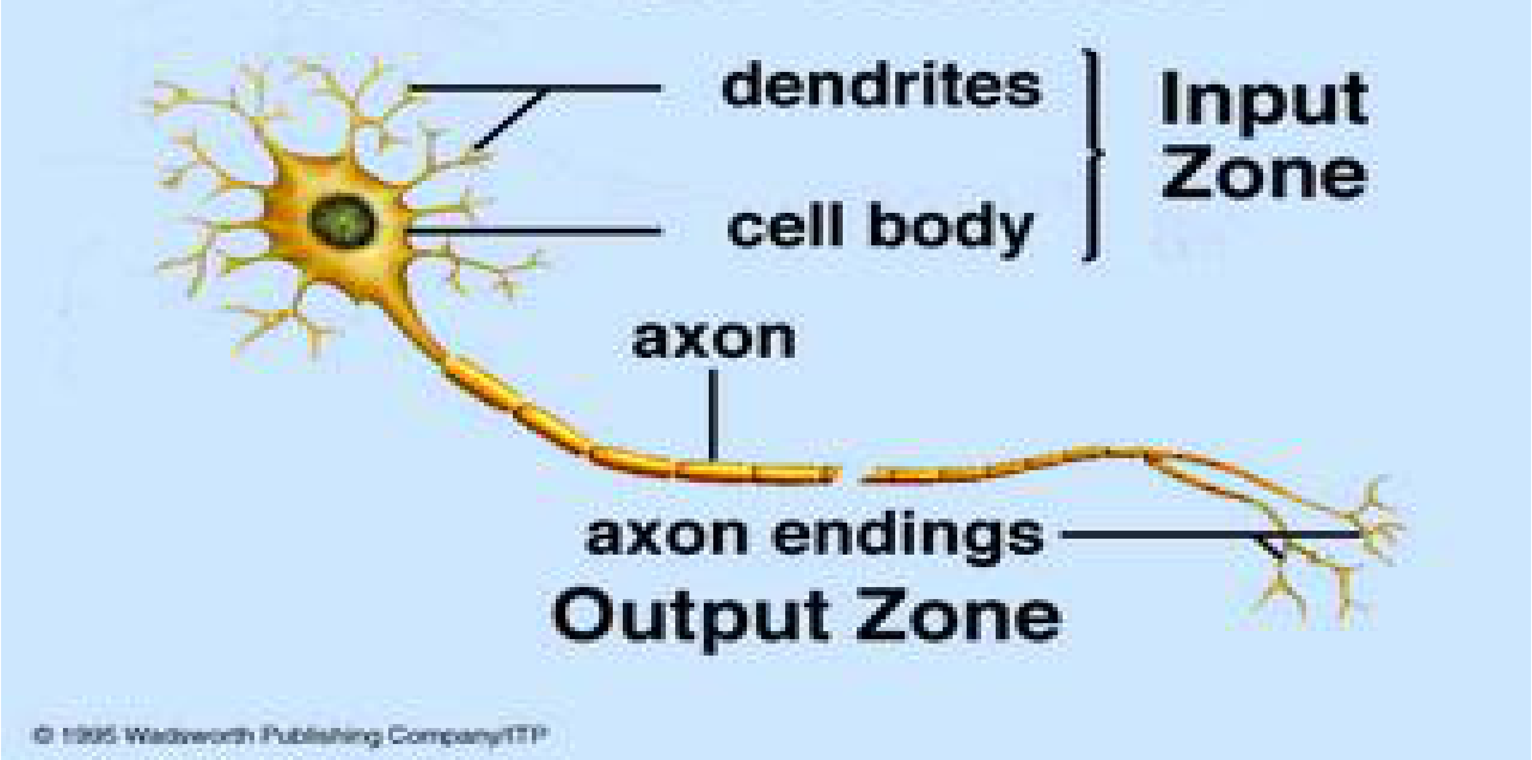
- We are born with about 100 billion neurons
- A neuron may connect to as many as 100,000 other neurons
- Signals “move” via electrochemical signals

- 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

Inspiration from Neurobiology

Human Biological Neuron



The Structure of Neurons

A neuron has a cell body, a branching **i**nput structure (the dendr**I**te) and a branching **o**utput structure (the ax**O**n)

- Axons connect to dendrites via synapses.
- Electro-chemical signals are propagated from the dendritic input, through the cell body, and down the axon to other neurons

The Structure of Neurons

- 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"

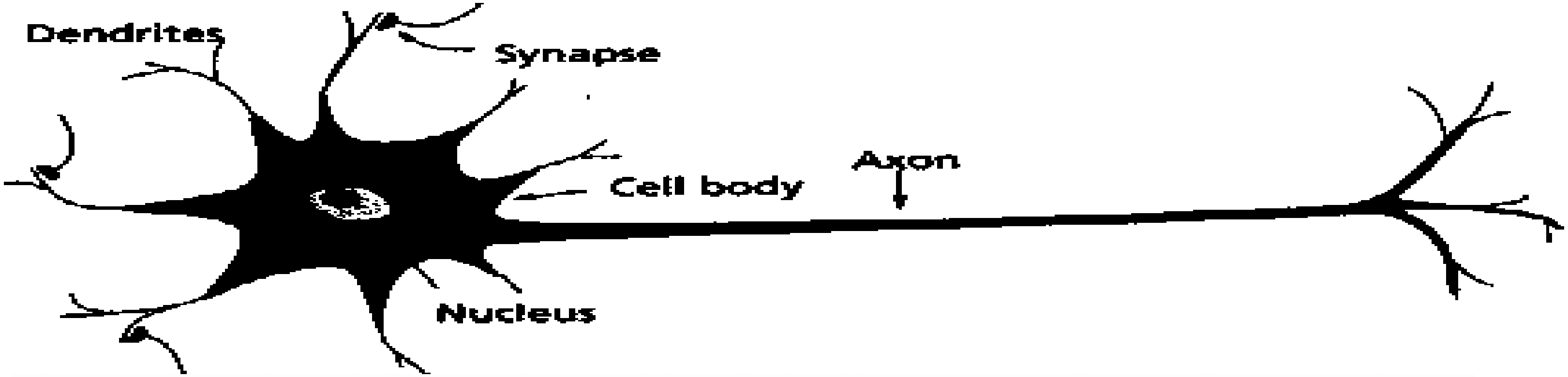
An Artificial Neuron

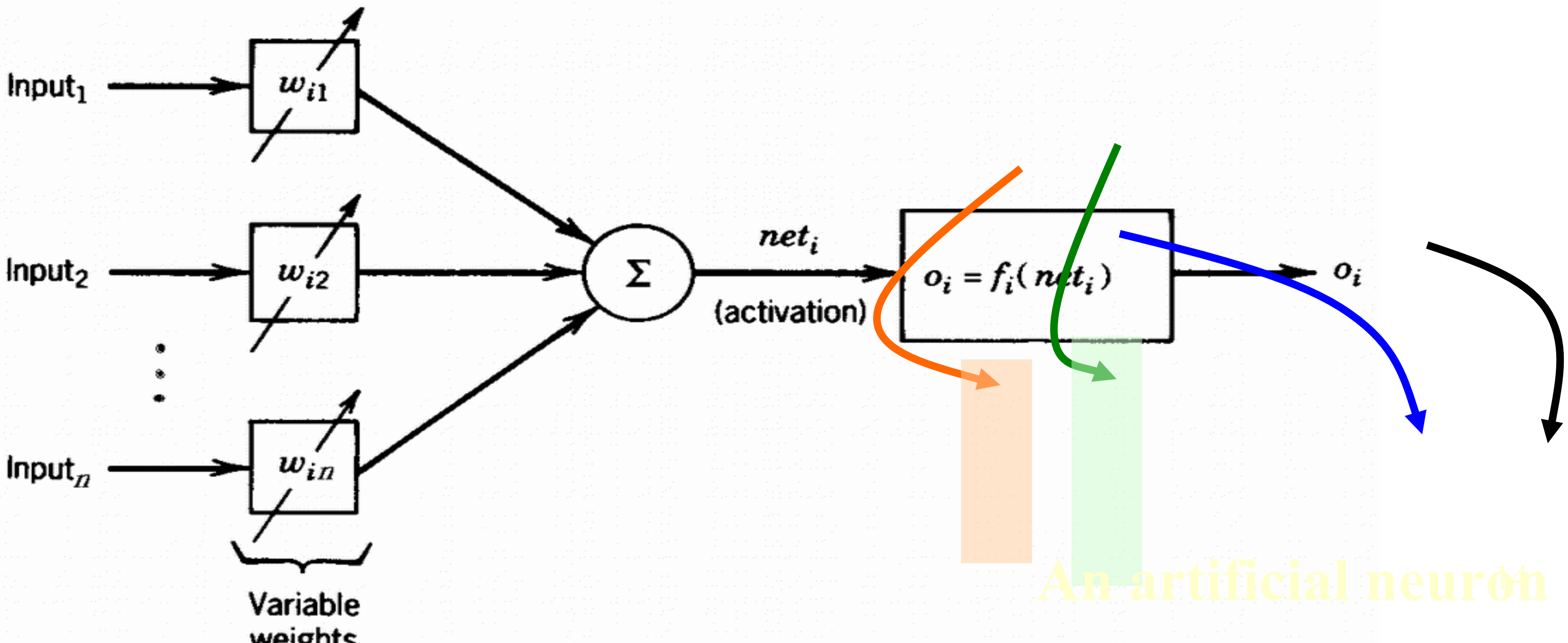
- Each hidden or output neuron has weighted input connections from each of the units in the preceding layer.
- The unit performs a weighted sum of its inputs, and subtracts its threshold value, to give its activation level.
- Activation level is passed through a sigmoid activation function to determine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

Biological Neural Networks

- 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.

- From experience: examples / training data
- Strength of connection between the neurons is stored as a weight-value for the specific connection.
- Learning the solution to a problem = changing the connection weights



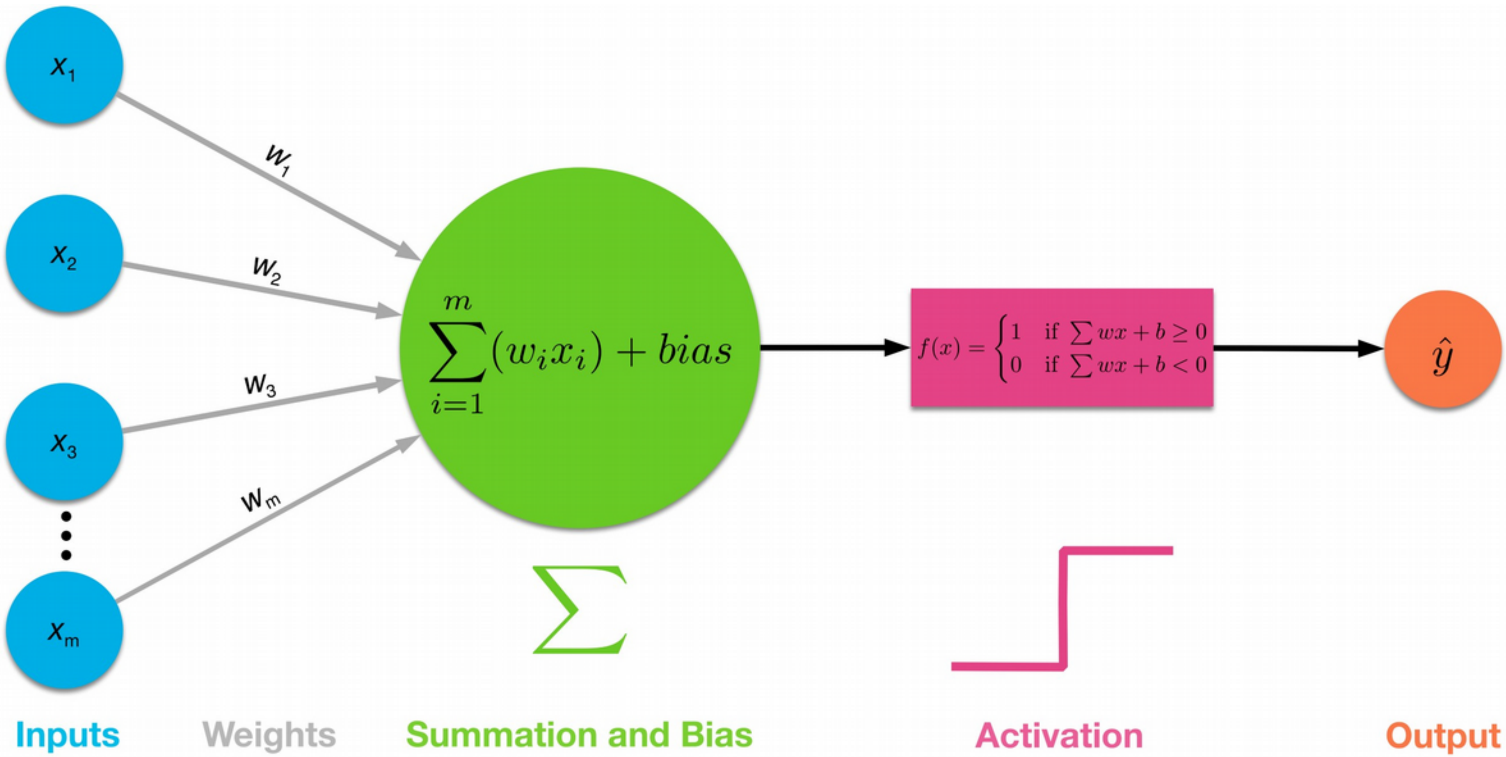


- A neural net consists of a large number of simple processing elements called neurons, units, cells or nodes.

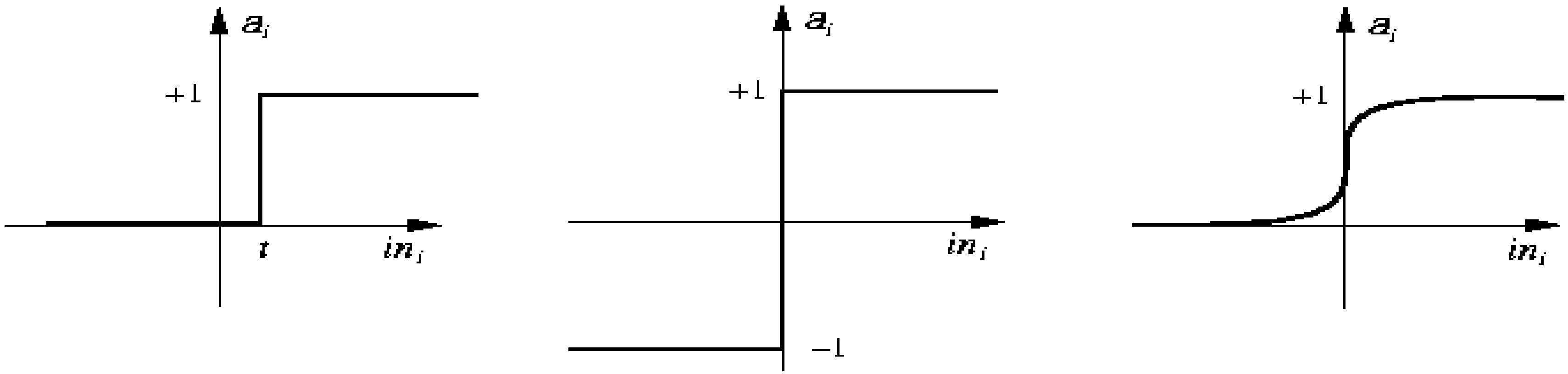
- Each neuron is connected to other neurons by means of directed communication links, each with associated weight.

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

Neuron Model



Activation Functions



(a) Step function (b) Sign function (c) Sigmoid function

- $\text{Step}_t(x)$ = 1 if x >= t, else 0
- Sign(x) = +1 if x >= 0, else –1
- Sigmoid(x) = $1/(1+e^{-x})$

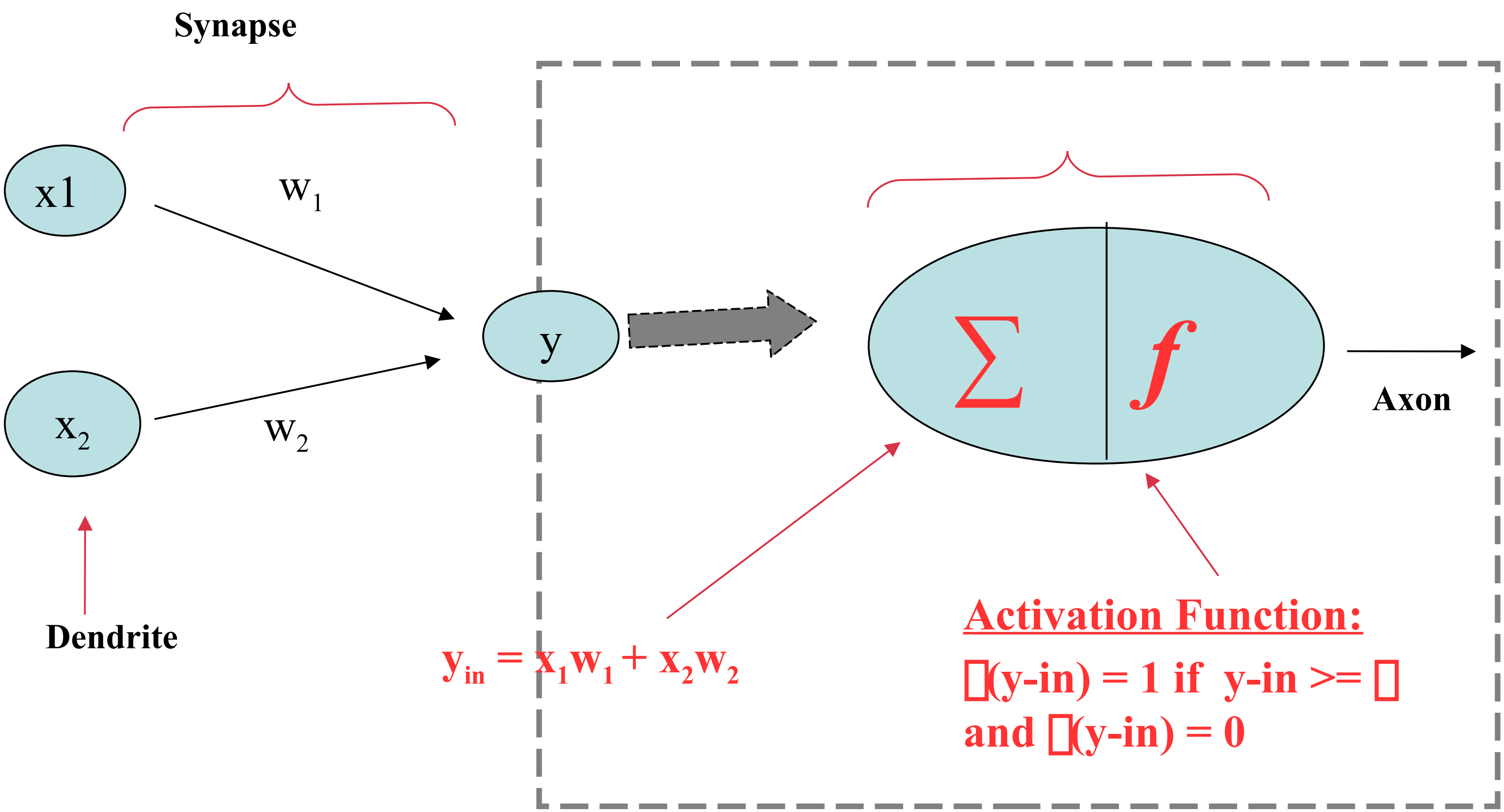
- Each neuron has an internal state, called its activation or activity level, which is a function of the inputs it has received. Typically, a neuron sends its activation as a signal to several other neurons.

- 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 (ANNs)

Artificial Neural Network



- 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)

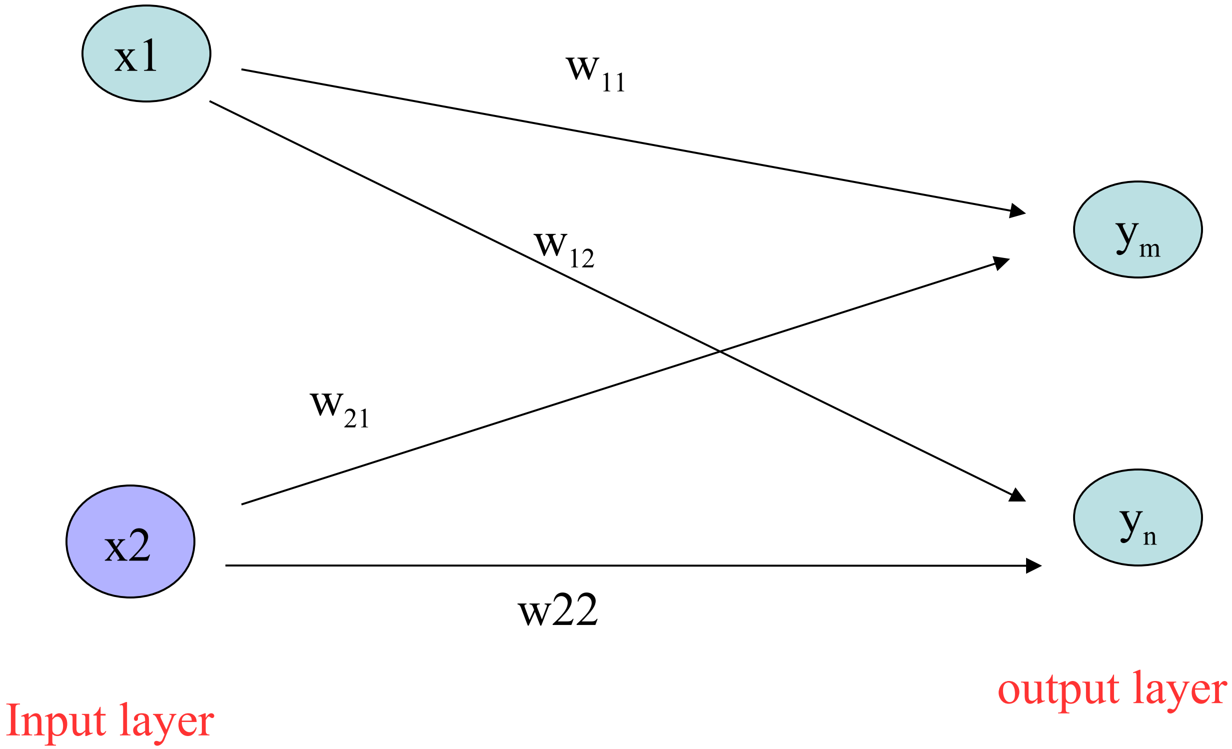
- The value is in the range of 0 and 1
- If the total weighted input greater than or equal the threshold value, the neuron will produce the

output, and if the total weighted input less than the threshold value, no output will be produced

Types of networks

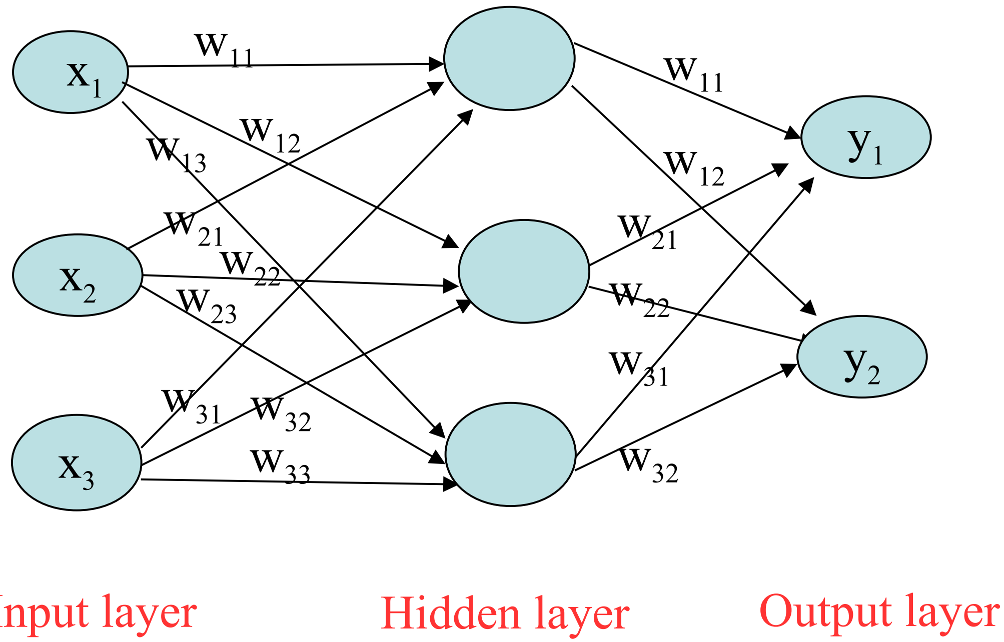
Single Layer Feed-forward NN

Input layer projecting into the output layer



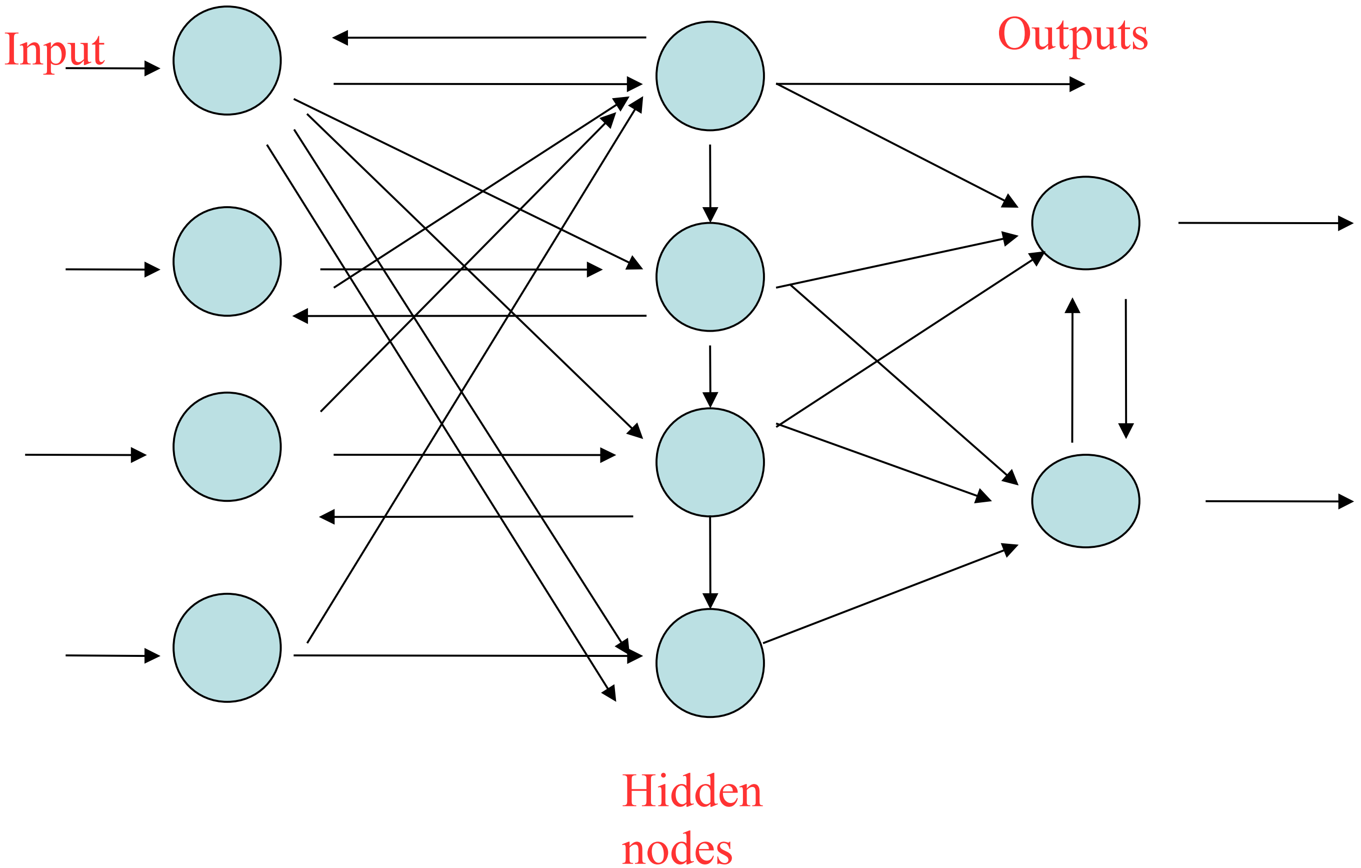
Multilayer feed-forward NN

- One or more hidden layers.
- – Input projects only from previous layers onto a layer.
- Typically, only from one layer to the next
- 2-layer or
- 1-hidden layer
- fully connected network



Recurrent NN

– A network with feedback, where some of its inputs are connected to some of its outputs (discrete time).



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.

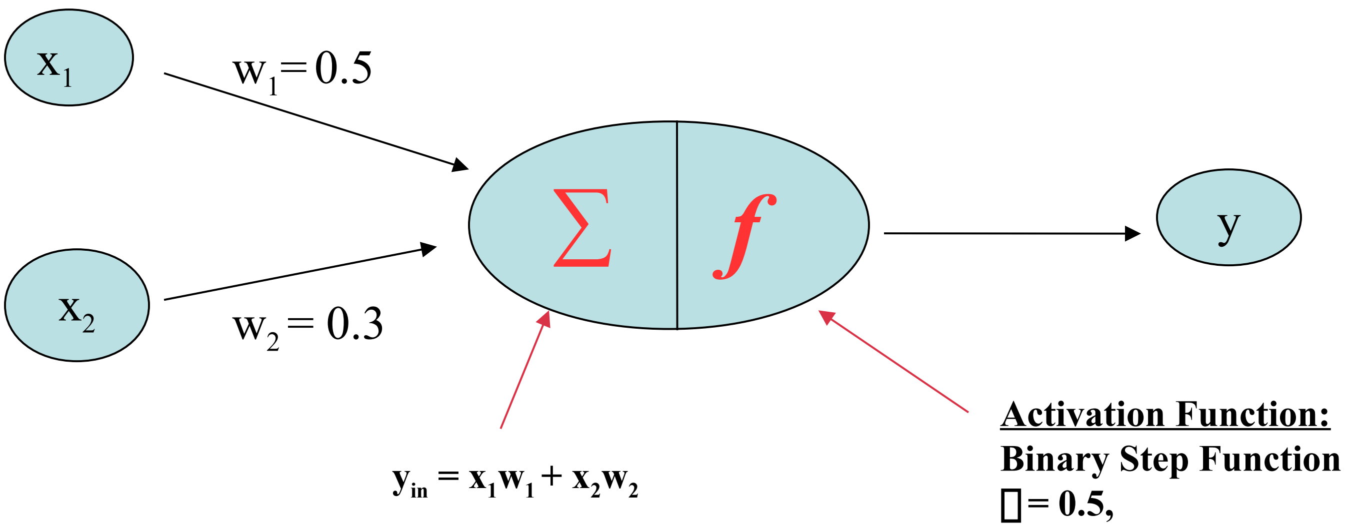
Early ANN Models:

- Perceptron, ADALINE, Hopfield Network

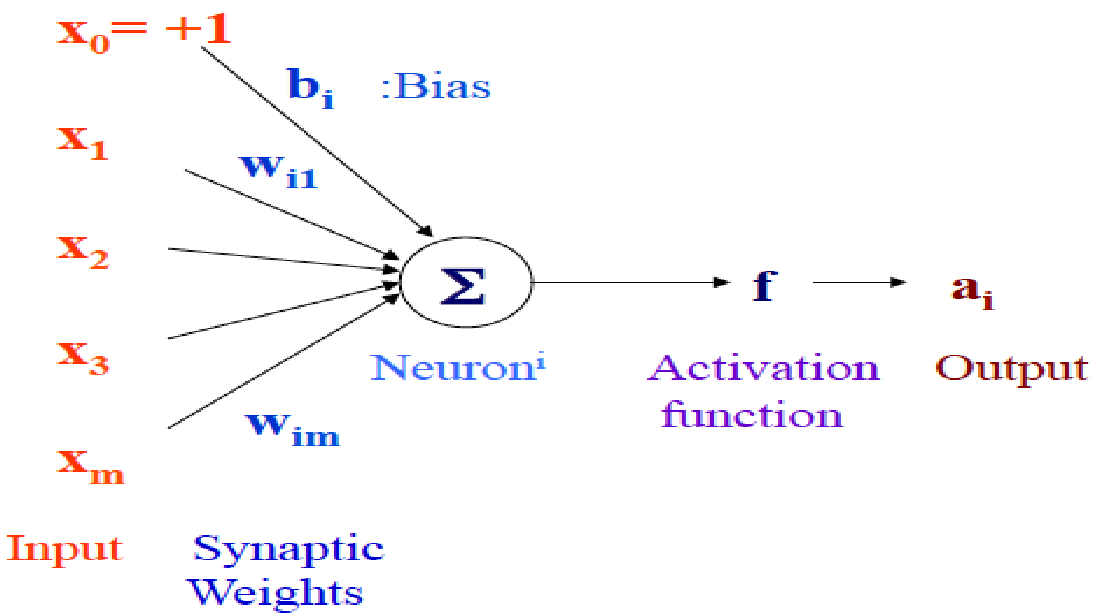
Current Models:

- Deep Learning Architectures
- Multilayer feedforward networks (Multilayer perceptrons)
- Radial Basis Function networks
- Self Organizing Networks

Example


x_1
$w_1 = 0.5$
x_2
$w_2 = 0.3$
Σ
f
y
$y_{in} = x_1w_1 + x_2w_2$
Activation Function:
Binary Step Function
□= 0.5,

Some Neurons have a Bias



Bias

$$a_i = f\,(n_i) = f\,\left(\sum_{j=1}^{n} w_{ij}x_j + b_i\right)$$

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

above zero.

Applications of ANNs

- 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

ANN 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.

Neural Network Architectures

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