innovate achieve lead

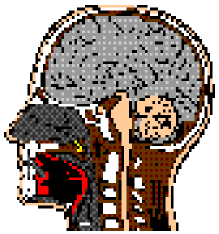
BITS Pilani

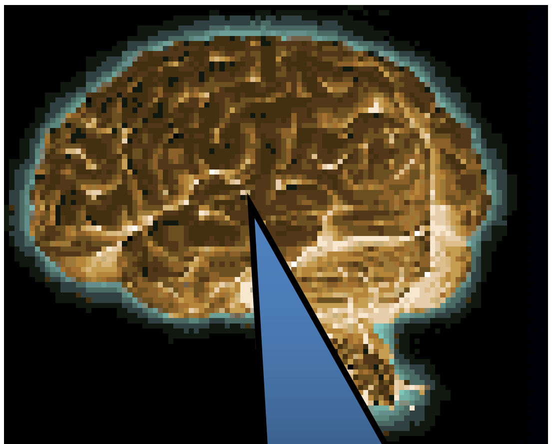
Machine Learning (IS ZC464) Session 9: Artificial Neural Networks

Artificial Neural Networks

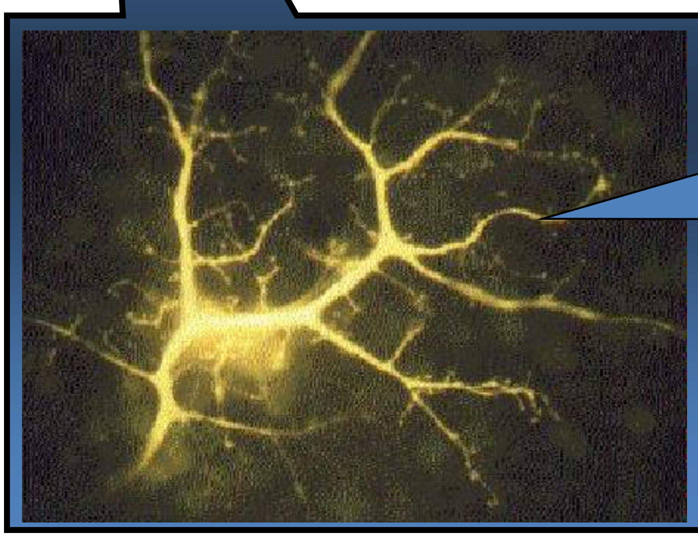
- Mathematical Models representing the massively parallel machines.
- Model inspired by the working of human nervous system.
- Has a number of neurons performing the task similar to human neuron.
- Each neuron triggers the received input according to the weight.
- A neural network captures the environment it has to learn in terms of the weights.

Brain Computer: What is it?

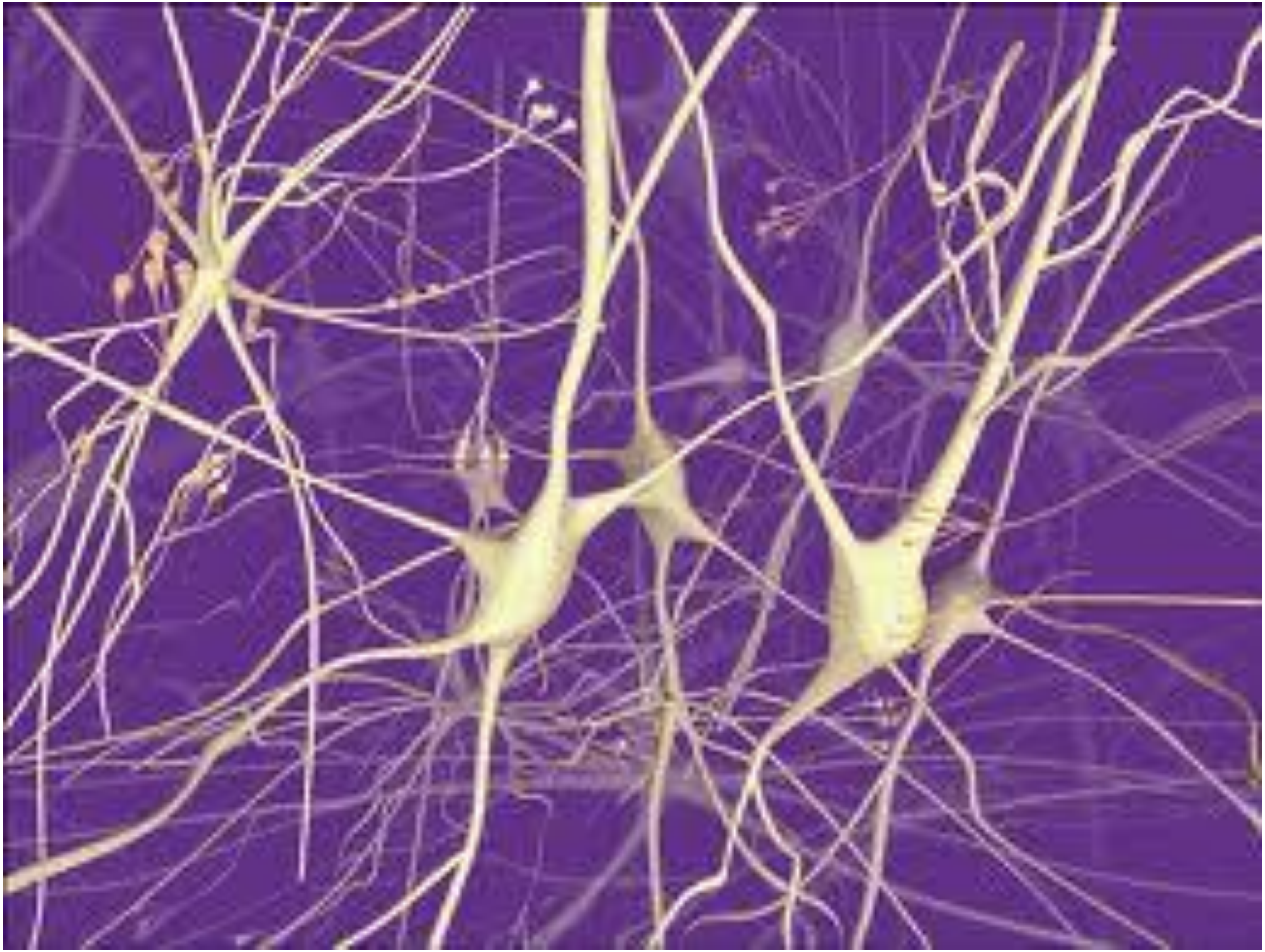




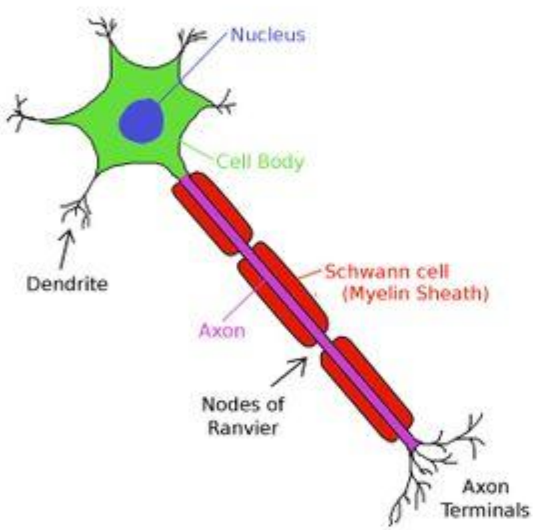
Human brain contains a massively interconnected net of 10^{10}-10^{11} (10 billion) neurons (cortical cells)

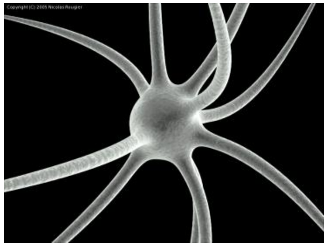


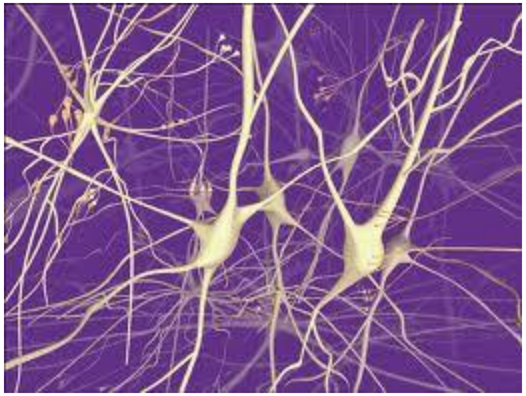
Biological Neuron
- The simple "arithmetic computing" element



Human nervous system

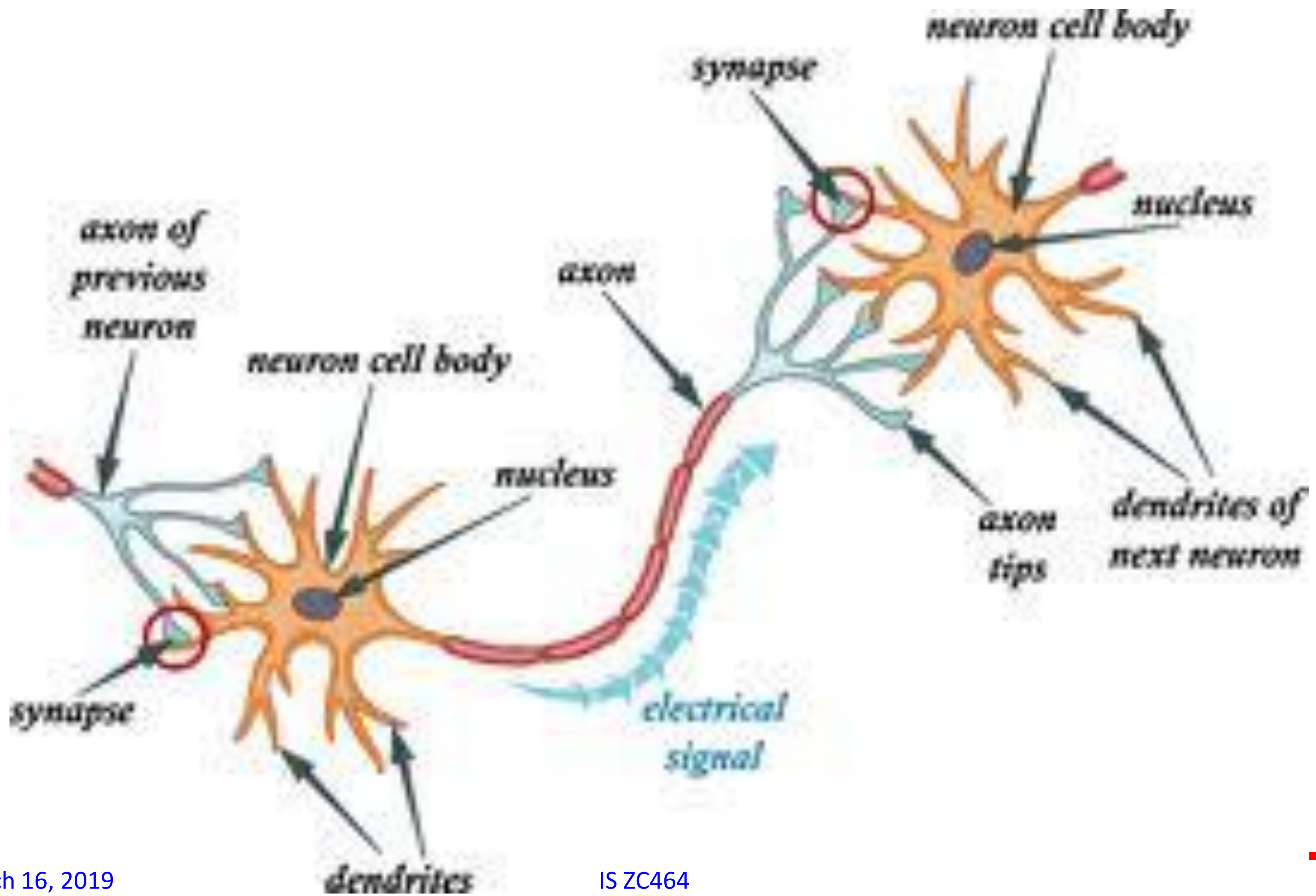






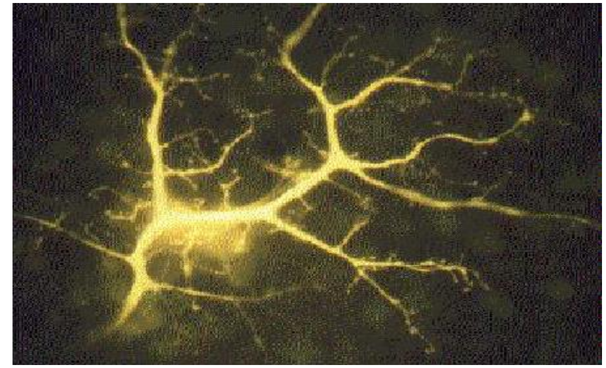
Images: google search

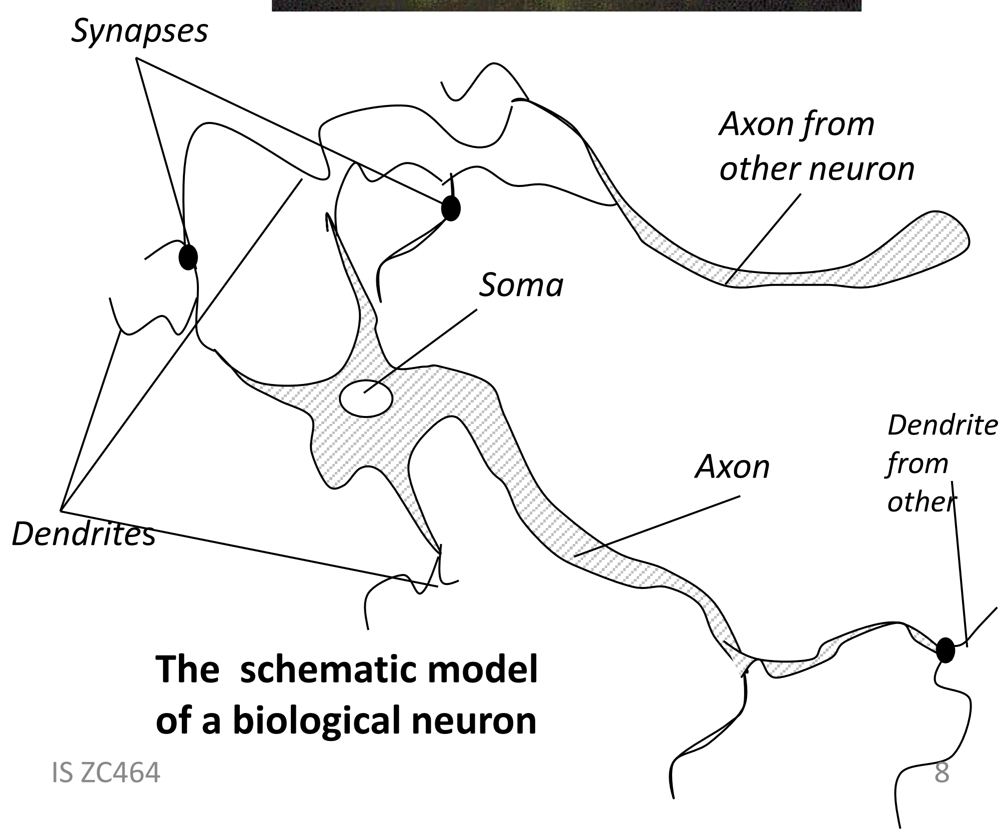
Human Neuron



Biological Neurons

1. ***Soma* or *body cell* -** is a large, round central body in which almost all the logical functions of the neuron are realized.
2. ***The axon* (output)**, is a nerve fibre attached to the soma which can serve as a final output channel of the neuron. An axon is usually highly branched.
3. ***The dendrites* (inputs)-** represent a highly branching tree of fibres. These long irregularly shaped nerve fibres (processes) are attached to the soma.
4. ***Synapses*** are specialized contacts on a neuron which are the termination points for the axons from other neurons.





The schematic model of a biological neuron

A human neuron

- A neuron is a nerve cell that is the basic building block of the human nervous system.
- Neurons transmit information throughout the body.
- These neurons transmit the information in their chemical and electrical form.
- Each neuron contains a nucleus that has genetic structure.
- The connection between cells are known as synapses.

Communication between synapses

- Electrical signals communicate between the neurons.
- Also there are neurotransmitters, which are the chemical messengers to transmit information from one to another neuron.
- Some major neurotransmitters are as follows
- **Acetylcholine:**
  - Associated with memory, muscle contractions, and learning. A lack of acetylcholine in the brain is associated with Alzheimer's disease.
- **Endorphins:**
  - Associated with emotions and pain perception. The body releases endorphins in response to fear or trauma..
- **Dopamine:**
  - Associated with thought and pleasurable feelings. Parkinson's disease is one illness associated with deficits in dopamine, while schizophrenia is strongly linked to excessive amounts of this chemical messenger.

Summation process in human neuron

- Each neuron is connected with numerous other neurons, receiving numerous impulses from them.
- **Summation** is the adding together of these impulses at the axon hillock.
- If the neuron only gets excitatory impulses, it will also generate an action potential;
- but if the neuron gets as many inhibitory as excitatory impulses, the inhibition cancels out the excitation and the nerve impulse will stop there.
- Summation takes place at the axon hillock.
- **Spatial summation** means several firings on different places of the neuron, that in themselves are not strong enough to cause a neuron to fire. However, if they fire simultaneously, their combined effects will cause an action potential.

Text adapted from wikipedia

How do humans recognize a face?

1. Eye captures the information and passes it to the brain via synapses.
2. There are about 1000 billion neurons in the human body.
3. The information is passed through these neurons to the brain.
4. The brain trains itself with the prominent and discriminatory features of an individual
5. Whenever the person sees the same face again, a query is automatically generated in the brain.
6. The brain answers the query “Who that person is?’ on the basis of training already done.

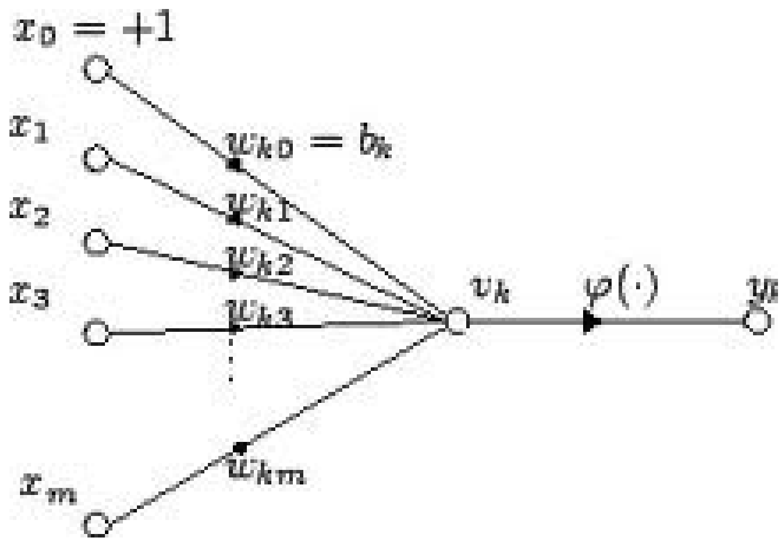
Neural Networks

- These are mathematical models to solve problems whose algorithmic solutions are difficult to be designed or algorithmic solutions do not exist.
- The technology of neural networks is more useful when the learning has to take place on the basis of observations.
- These models are based on the human neural system.
- It collects input through various interconnection links called synapses and passes the weighted information to the neurons.

An Artificial Neuron

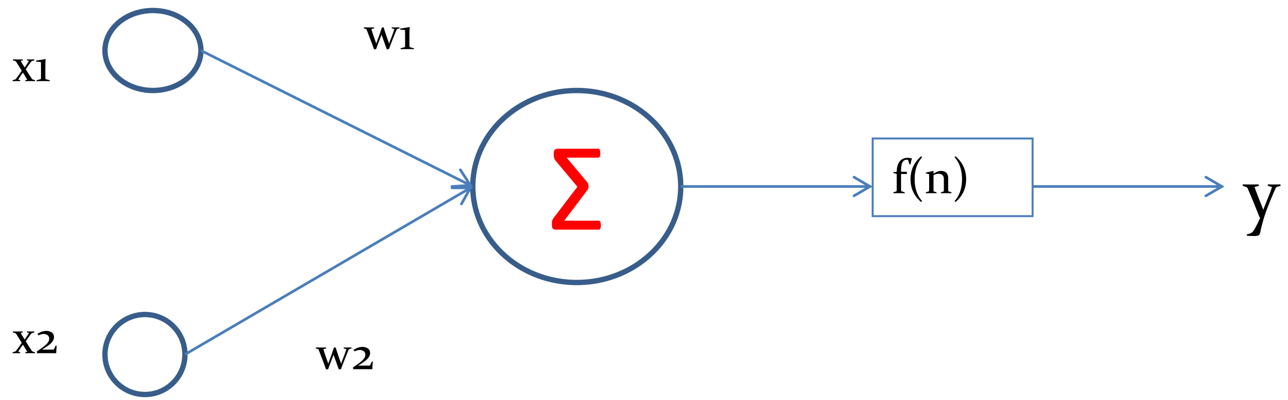
- A mathematical neuron is a processing unit capable of receiving inputs from single or multiple neurons and triggers a desired response.
- Each neuron has an associated activation function which takes as input the weighted sum of the inputs coming to the neuron and triggers a response depending on the associated threshold

Artificial Neuron



Artificial Neuron

- Weighted sum (n) = $\Sigma\, w_i x_i$
- Activation function f(n) = f($\Sigma\, w_i x_i$)
- Threshold T such that a response is triggered if f(n) > T



Example : Simulating a neuron to perform a logical AND

- Truth table for AND

x1	x2	x1 AND x2
0	0	0
0	1	0
1	0	0
1	1	1

- Let the activation function f(n) be defined as

$$f(n) = \begin{cases} 1 & \text{if } n>1 \\ 0 & \text{otherwise} \end{cases}$$

Estimate the weights

- $n = w_1x_1 + w_2x_2$
- Case $x_1 = 0$ and $x_2 = 0$

  $n = 0$ for any pair of weight parameters
- Case $x_1 = 0$ and $x_2 = 1$

  $n = w_2$
- Case $x_1 = 1$ and $x_2 = 0$

  $n = w_1$
- Case $x_1 = 1$ and $x_2 = 1$

  $n = w_1 + w_2$

Other Neuron parameters

- Threshold = 1
- Activation function f(n) = 1 if n>T

  = 0 if n<=T

x1	x2	w1	w2	n	f(n)
0	0				0
0	1				0
1	0				0
1	1				1

View f(n) as the desired output.

Other Neuron parameters

- Threshold = 1
- Activation function f(n) = 1 if n>T

  = 0 if n<=T

x1	x2	w1	w2	n	f(n)
0	0				0
0	1				0
1	0				0
1	1	1	1	2	1

Other Neuron parameters

- Threshold = 1
- Activation function f(n) = 1 if n>T

  = 0 if n<=T

x1	x2	w1	w2	n	f(n)
0	0	1	1	0	0
0	1	1	1	1	0
1	0	1	1	1	0
1	1	1	1	2	1

Supervised learning

- Knowing the expected output and adjusting the weight to get that output is known as supervised learning.
- Learning is in terms of obtaining weights.
- The neuron is trained with the input patterns <0,0> <0,1> <1,0> and <1,1> so as to learn that the patterns should ne named as 0,0,0 and 1 respectively. (relate to face recognition)
- Once a neuron is trained, it is expected to use its learning to answer any input pattern.
- Example: what is <1,0>? Answer is 0. [No logical operation performed anywhere]

Testing

- Input a pattern <0.001, 0.987> [weak signal]
- n = 0.001 x 1 + 0.987 x 1 = 0.988
- f(n) = 0
- Notice that even a unseen pattern is also recognised correctly.

- Another input: <0.876, 0.991> [weak signal]
- n = 1.867 > T
- Hence f(n) = 1

Testing

- Input <1.009, 1.100> [strong signal/spikes]
- n = 1.009 + 1.100 = 2.109 > 1
- f(n) = 1 (expected)

Class Assignment

- Simulate a neuron to generate OR function

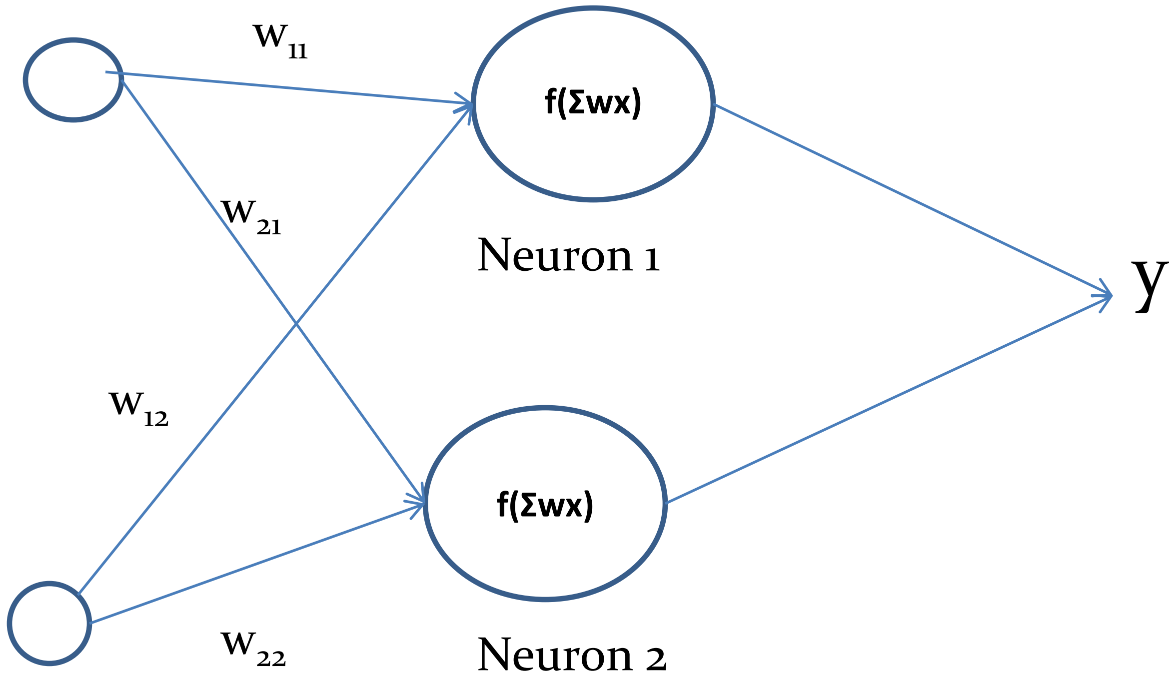
McCulloch-Pitts Neuron Model

- The first formal definition of an artificial neuron was formulated by Warren McCulloch and Walter Pitts in 1943.
- The McCulloch-Pitts Neuron allows binary 0 or 1 states only.
- The connection path can be excitatory or inhibitory.
- Excitatory connections have positive weights and inhibitory connections have negative weights.

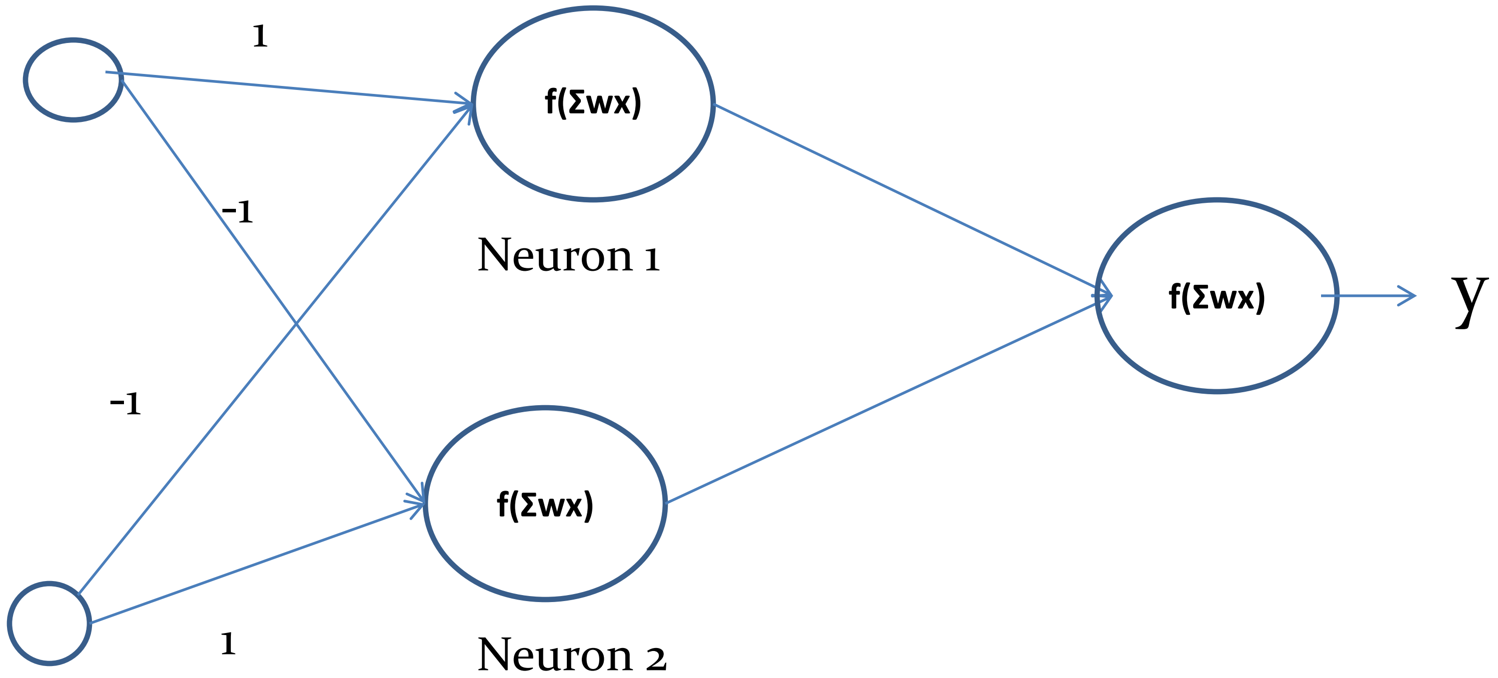
Exclusive OR simulation using McCulloch-Pitts neuron

- Draw the truth table of XOR.
- Try to plot the 2 dimensional pattern on a 2D Cartesian plane and plot a symbol appropriately for both 0 and 1 outputs.
- Observe that a single line cannot demarcate the boundary (which was possible in the AND and OR examples)
- This indicates that the neural simulation is not possible with one neuron. Use two neurons.

Neural Network for XOR



Neural Network for XOR



Perform Computations (T=1)

Activation Function

$$f(n) = 1 \text{ if } n >= 1$$
$$= 0 \text{ if } n < 1$$

x1	x2	w11	w12	w21	w22	n1	f(n1)	n2	f(n2)	f(m)
0	0	1	-1	-1	1	0	0	0	0	0
0	1	1	-1	-1	1	-1	0	1	1	1
1	0	1	-1	-1	1	1	1	-1	0	1
1	1	1	-1	-1	1	0	0	0	0	0

Activation Functions

- Activation functions trigger the received weighted sum of the input according to the expected output.
- Let g(n) be the activation function where n= $\Sigma\ w_i x_i$
- Based on the value of n, the g(n) is triggered.

- Different Activation Functions are
  - Step Function
  - Linear function
  - Sigmoid Function (Logistic Function)

Step Function

- $G(n) = 1$ if $n>T$

  $= 0$ if $n<=T$
- The output is always 1 or 0.
- Useful is pattern recognition where an output 1 can represent the class which the input test pattern belong to.

Example

- Let us have two discriminatory features <color, texture> for automatic fruit recognition.
- Let us code the attributes as 1, 2 and 3 for red, orange and yellow colors respectively. Also let the textures be defined as 1,2 and 3 representing the degrees of smoothness in increasing order.
- Let the training samples be

Training data

Color	Texture	Fruit (supervised learning)
1	2	Apple
3	2	Mango
2	1	Orange
1	3	Plum
1.4	2.8	plum
2	1.5	orange
2.5	2.2	mango
1	2.2	apple

Expected output

- Apple → 1
- Orange → 2
- Mango → 3
- Plum → 4

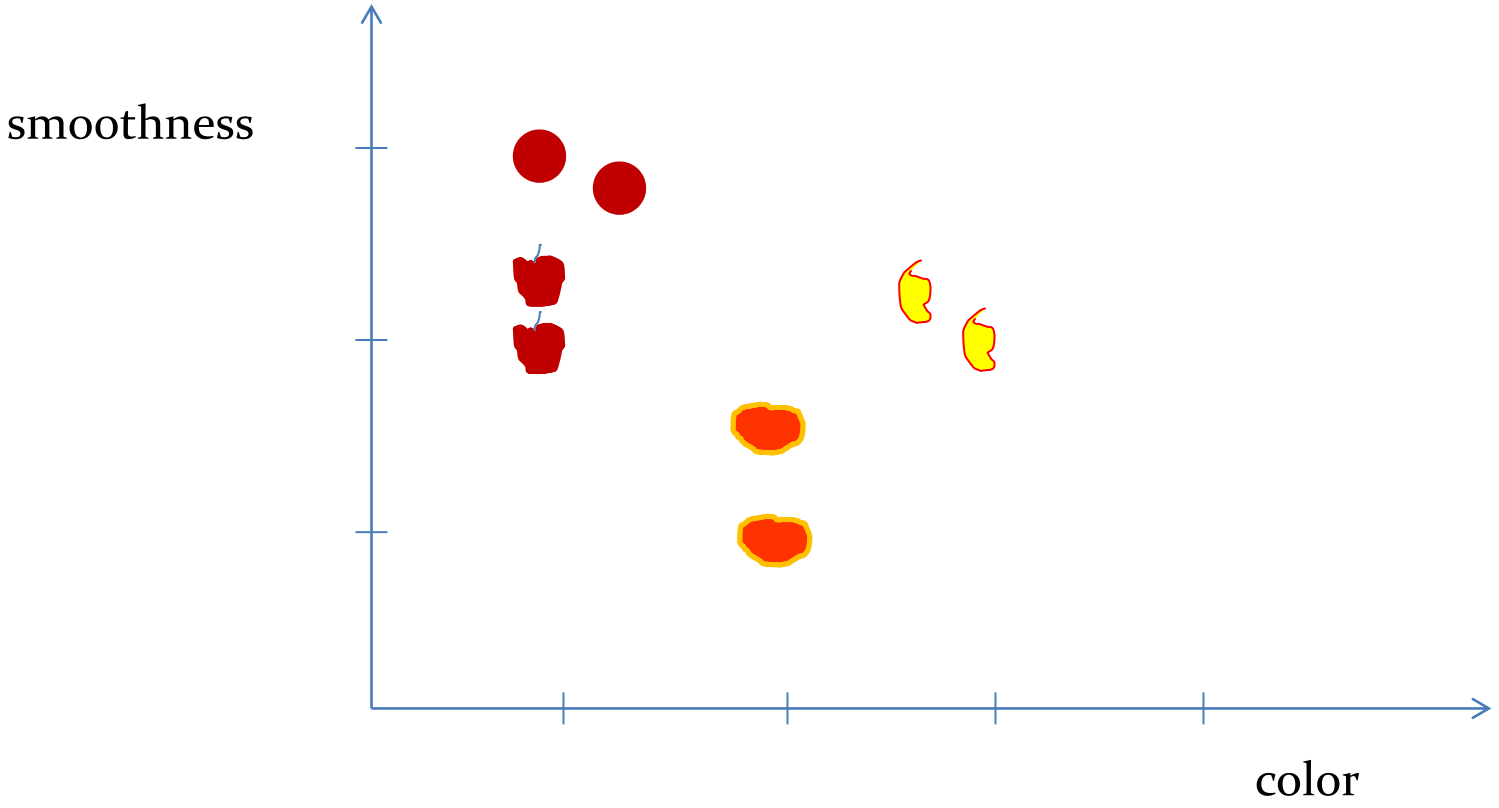
The expected output can also be modeled as a 4 tuple as <1,0,0,0> for apple

<0,1,0,0> for Orange

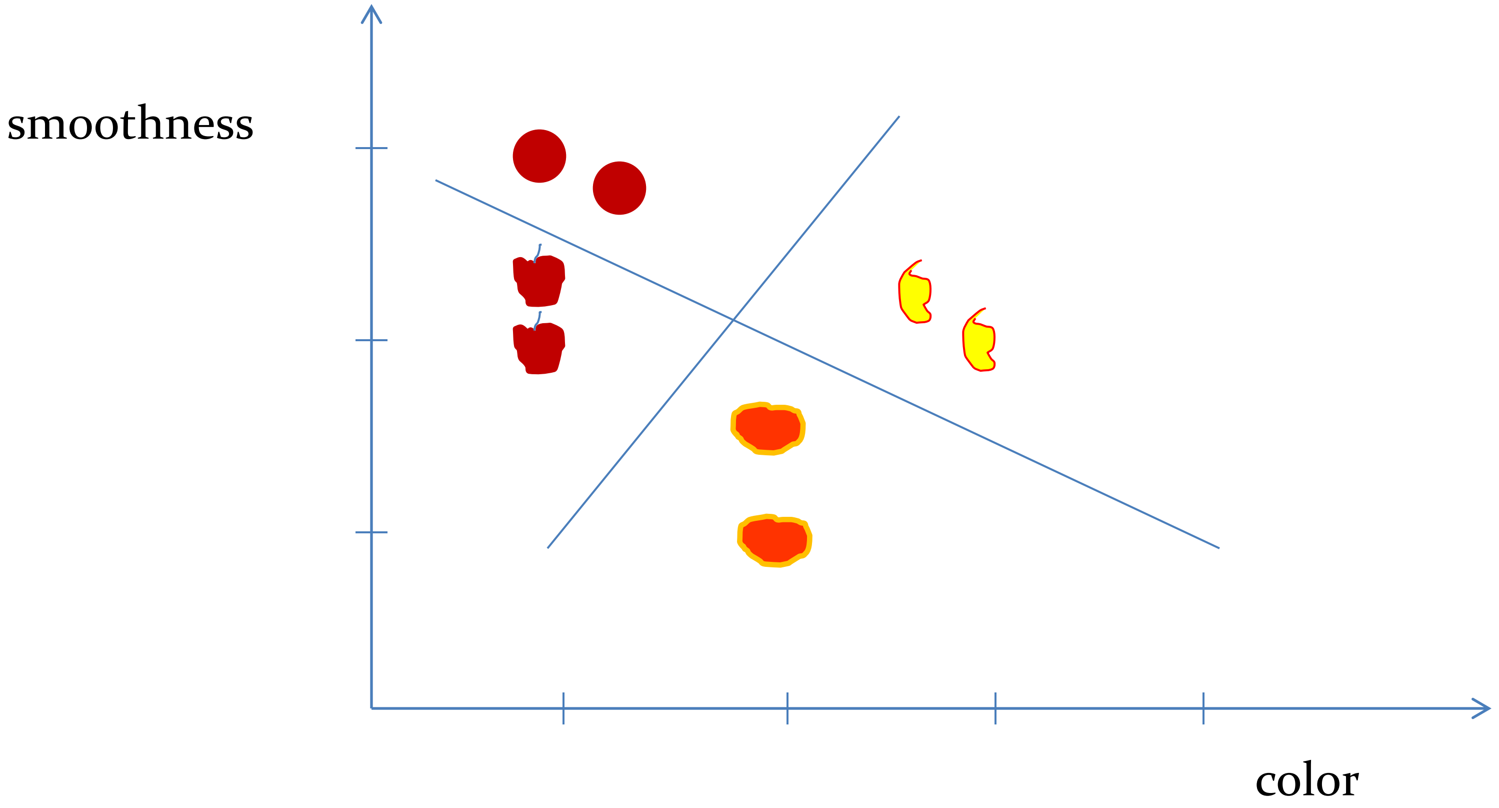
<0,0,1,0> for Mango

<0,0,0,1> for plum

Check the feasibility of using a single neuron for recognition of patterns

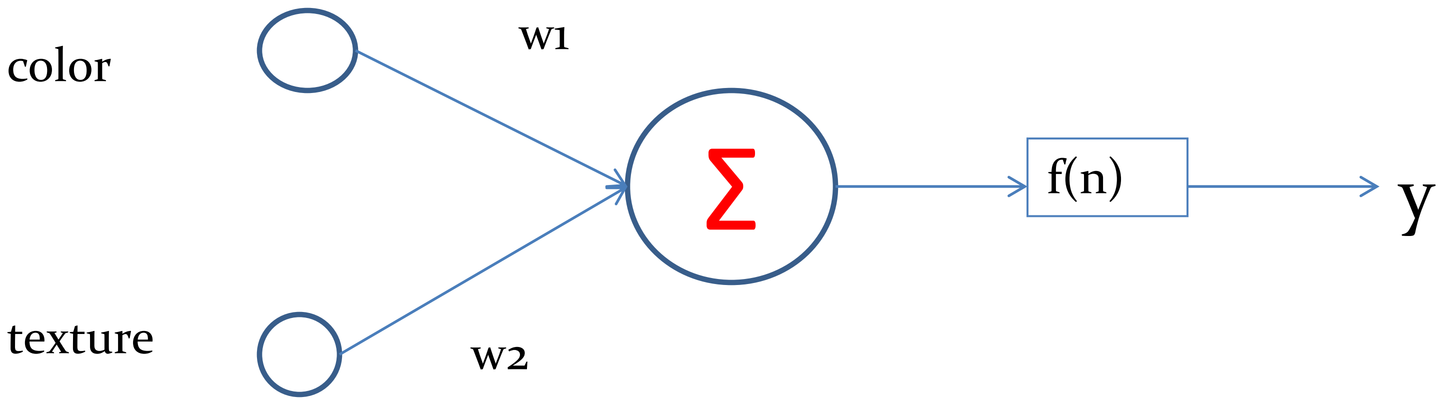


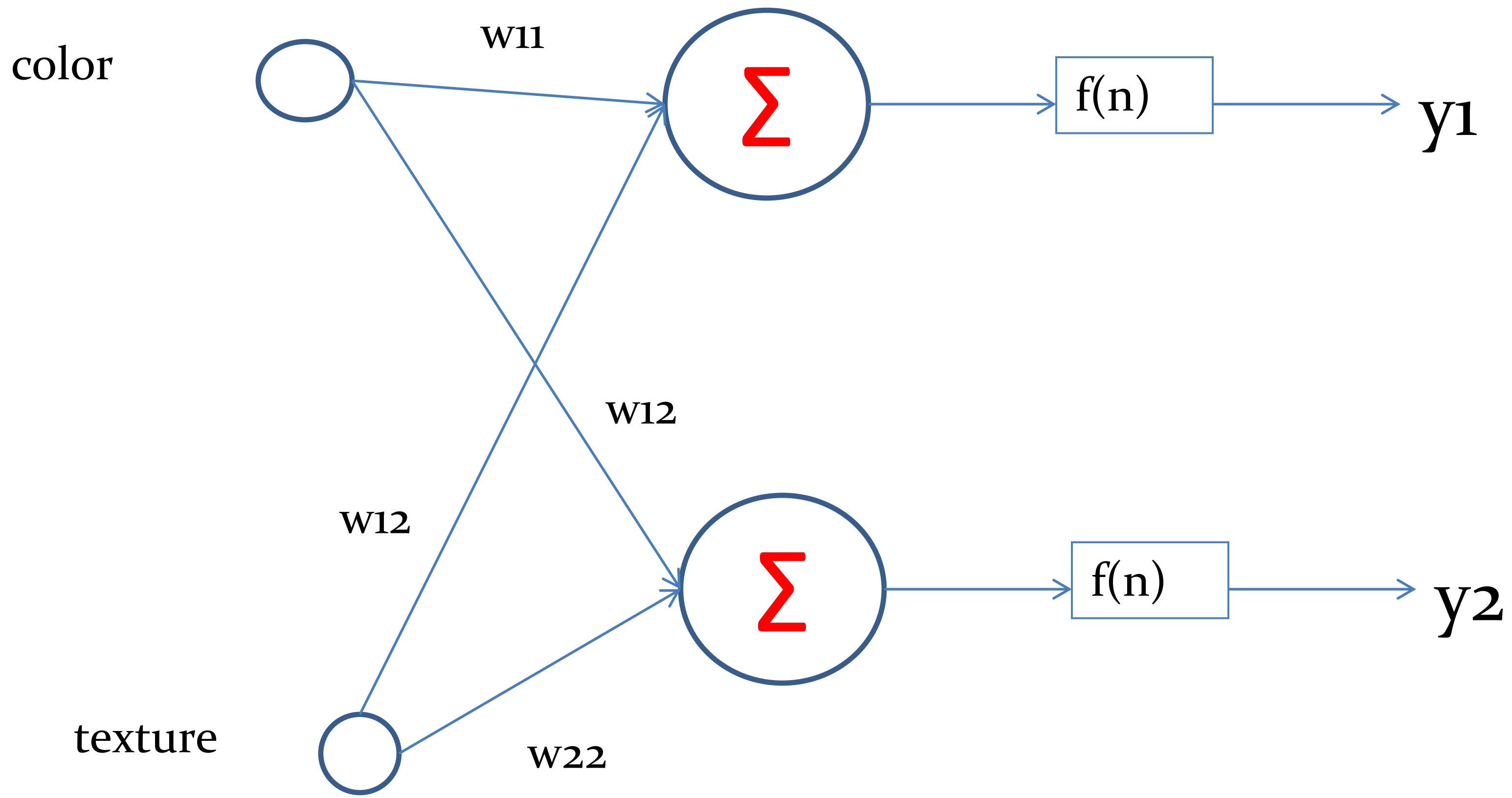
Try to draw lines that separate the classes: two decision boundaries hint the need for two neurons



Train the neuron: Select the activation function first

A simple step function simply gives 0 or 1 and is not sufficient to produce output for 4 classes





color
w11
y1
f(n)
Σ
w12
w12
texture
w22
Σ
f(n)
y2

Output modeling

- y1 = 0 , y2 = 0 represents that the test input is Apple
- y1 = 1 , y2 = 0 represents that the test input is Orange
- y1 = 1 , y2 = 0 represents that the test input is Mango
- y1 = 1 , y2 = 1 represents that the test input is Plum

Parameters for each neuron

- Neuron 1:
  - Activation Function: Step function
  - Threshold: 1
  - Weights: <1,-1>
- Neuron 2:
  - Activation Function: Step function
  - Threshold:1
  - Weights:<1,-1>

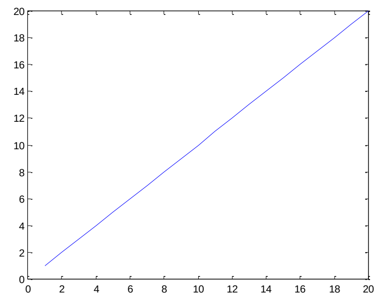
Training and Learning

- Training is supervised in nature.
- Neural network learns from training examples
- Knowledge is captured in terms of weights.
- Learning Algorithms require the initialization of weights which later adapt to other examples.
- Hit and trial methods do not work for larger data sets.
- Gradient Descent is the most popular learning algorithm.

Neural Network Architecture

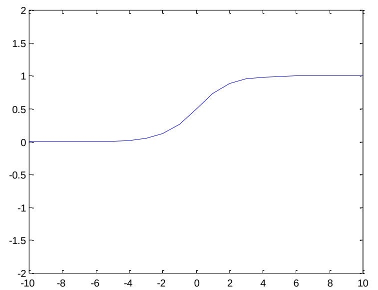
- The arrangement of neurons, along with the number of neurons, activation functions associated with the neurons and the input and output constitute the architecture of the Neural Network.
- Different Architectures include
  - Feed Forward Neural Networks-represents a function of its input
    - Single Layered
    - Multi Layered
  - Back Propagation Neural Networks- feeds its output back into its own inputs

Activation functions



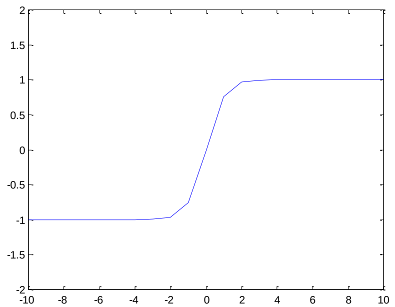
Linear

$$y = x$$



Logistic

$$y = \frac{1}{1+\exp(-x)}$$

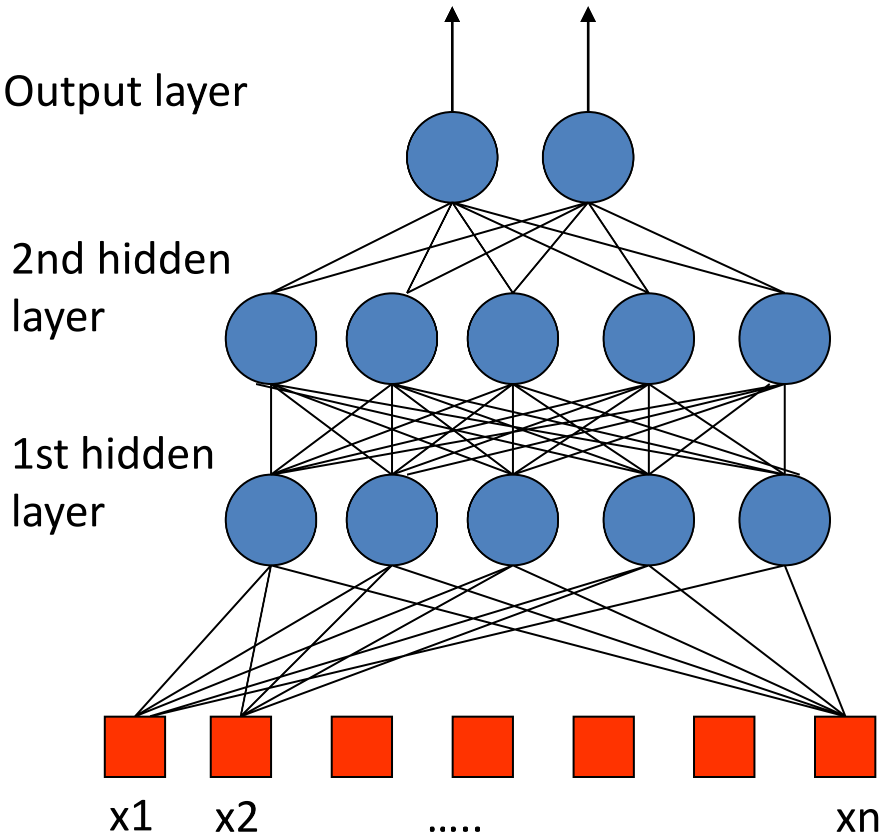


Hyperbolic tangent

$$y = \frac{\exp(x)-\exp(-x)}{\exp(x)+\exp(-x)}$$

*Slide adapted from : acat02.sinp.msu.ru/presentations/prevotet/tutorial.**ppt***

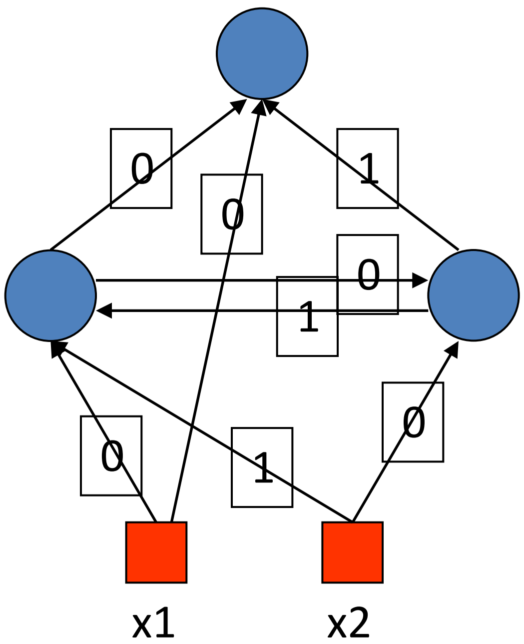
Feed Forward Neural Networks



- The information is propagated from the inputs to the outputs
- Time has no role (NO cycle between outputs and inputs)

*Slide adapted from : acat02.sinp.msu.ru/presentations/prevotet/tutorial.**ppt***

Recurrent Neural Networks



- Can have arbitrary topologies
- Can model systems with internal states (dynamic ones)
- Delays are associated to a specific weight
- Training is more difficult
- Performance may be problematic
  - Stable Outputs may be more difficult to evaluate
  - Unexpected behavior (oscillation, chaos, …)

*Slide adapted from : acat02.sinp.msu.ru/presentations/prevotet/tutorial.**ppt***