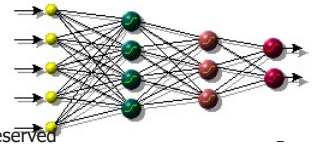
Artificial Neural Networks



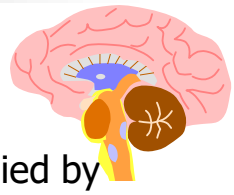
Outline

- The Brain
- Perceptrons
- Gradient descent
- Multi-layer networks
- Backpropagation

Artificial Neural Networks

- Other terms/names
  - connectionist
  - parallel distributed processing
  - neural computation
  - adaptive networks..
- History
  - 1943-McCulloch & Pitts are generally recognised as the designers of the first neural network
  - 1949-First learning rule
  - 1969-Minsky & Papert - perceptron limitation - Death of ANN
  - 1980's - Re-emergence of ANN - multi-layer networks

The biological inspiration



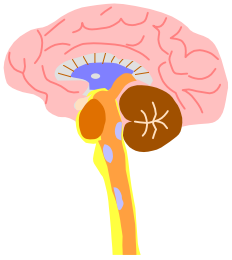
- The brain has been extensively studied by scientists.
- Vast complexity prevents all but rudimentary understanding.
- Even the behaviour of an individual neuron is extremely complex

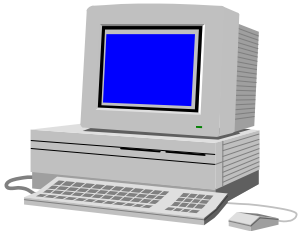
Features of the Brain

- Ten billion (10^{10}) neurons
- Neuron switching time $>10^{-3}$secs
- Face Recognition ~0.1secs
- On average, each neuron has several thousand connections
- Hundreds of operations per second
- High degree of parallel computation
- Distributed representations
- Die off frequently (never replaced)
- Compensated for problems by massive parallelism

Brain and Machine

- The Brain
  - Pattern Recognition
  - Association
  - Complexity
  - Noise Tolerance



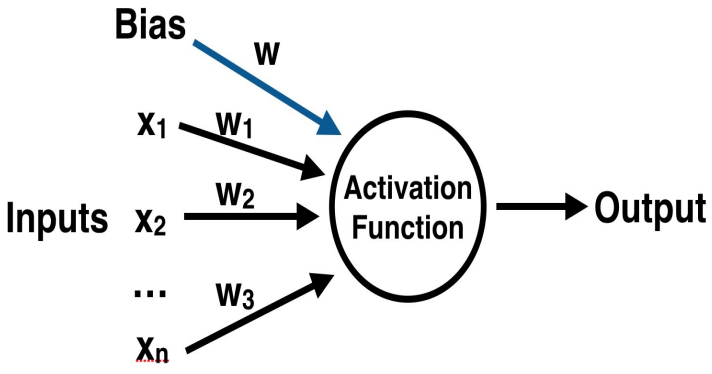


- The Machine
  - Calculation
  - Precision
  - Logic

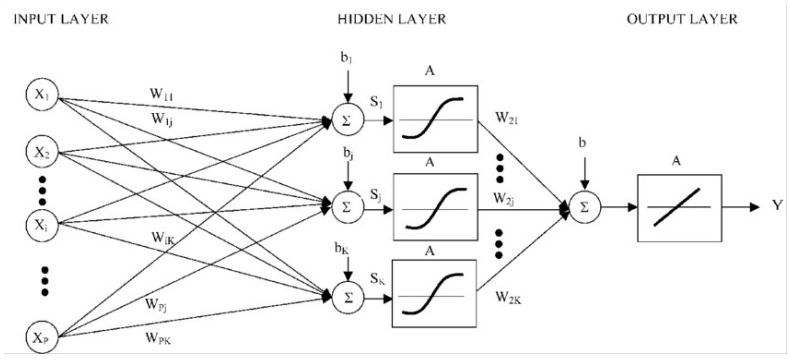
Neural Networks

- Neural Networks are a machine learning framework that attempts to mimic the learning pattern of natural biological neural networks:

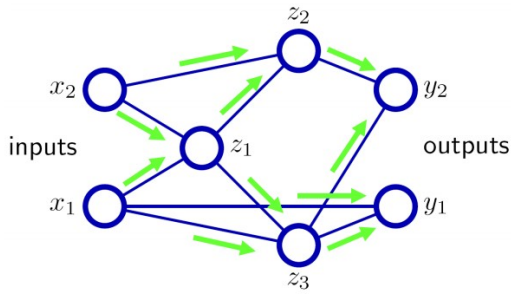
The Perceptron



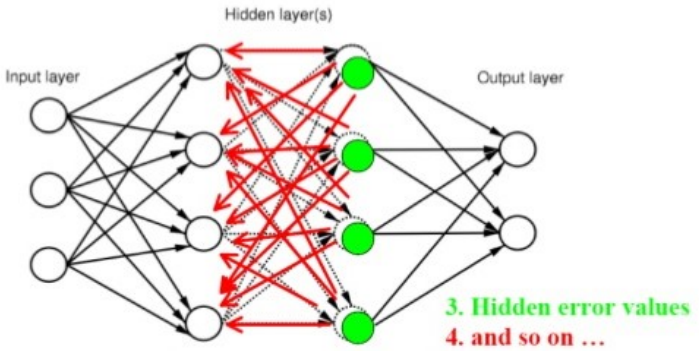
Activation function	Equation	Example	1D Graph
Unit step (Heaviside)	$\phi(z) = \begin{cases} 0, & z < 0, \\ 0.5, & z = 0, \\ 1, & z > 0, \end{cases}$	Perceptron variant	
Sign (Signum)	$\phi(z) = \begin{cases} -1, & z < 0, \\ 0, & z = 0, \\ 1, & z > 0, \end{cases}$	Perceptron variant	
Linear	$\phi(z) = z$	Adaline, linear regression	
Piece-wise linear	$\phi(z) = \begin{cases} 1, & z \geq \frac{1}{2}, \\ z + \frac{1}{2}, & -\frac{1}{2} < z < \frac{1}{2}, \\ 0, & z \leq -\frac{1}{2}, \end{cases}$	Support vector machine	
Logistic (sigmoid)	$\phi(z) = \frac{1}{1 + e^{-z}}$	Logistic regression, Multi-layer NN	
Hyperbolic tangent	$\phi(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$	Multi-layer NN	

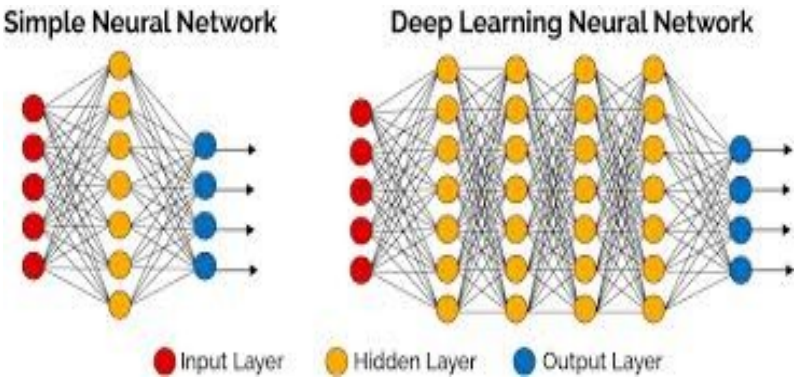


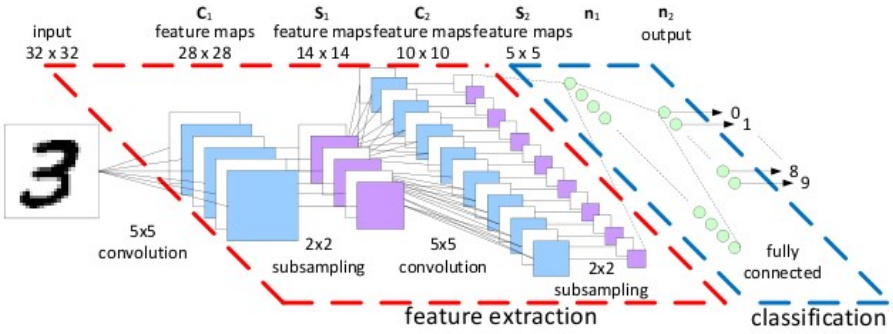
- Forward Propagation :
  - Sum inputs, produce activation, feed-forward

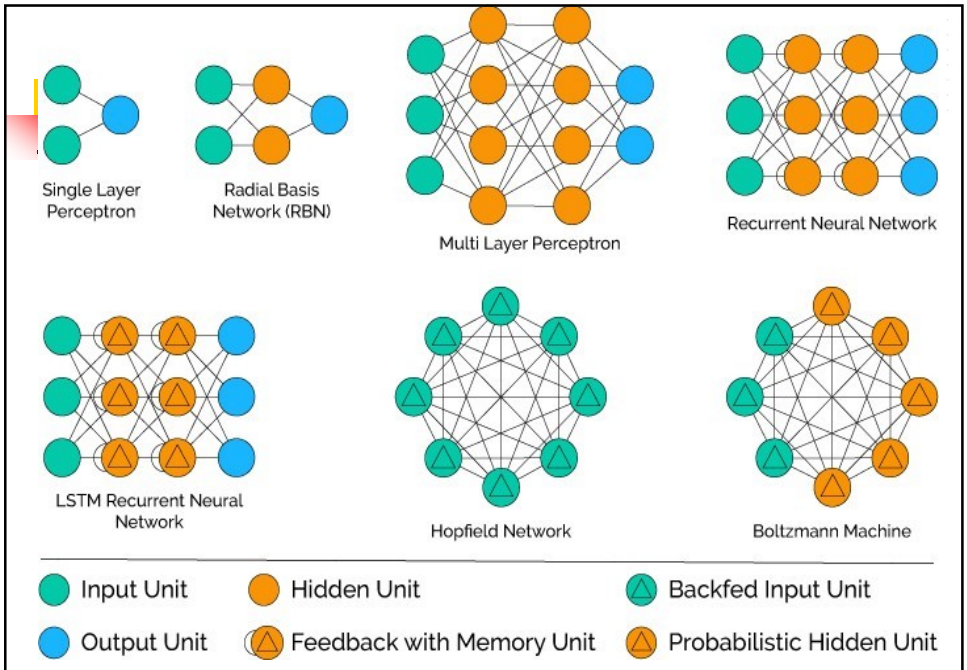


Backpropagation Learning









Convolutional Neural Network

- The process of building a Convolutional Neural Network always involves four major steps.
- **Step - 1 : Convolution**
- **Step - 2 : Pooling**
- **Step - 3 : Flattening**
- **Step - 4 : Full connection**