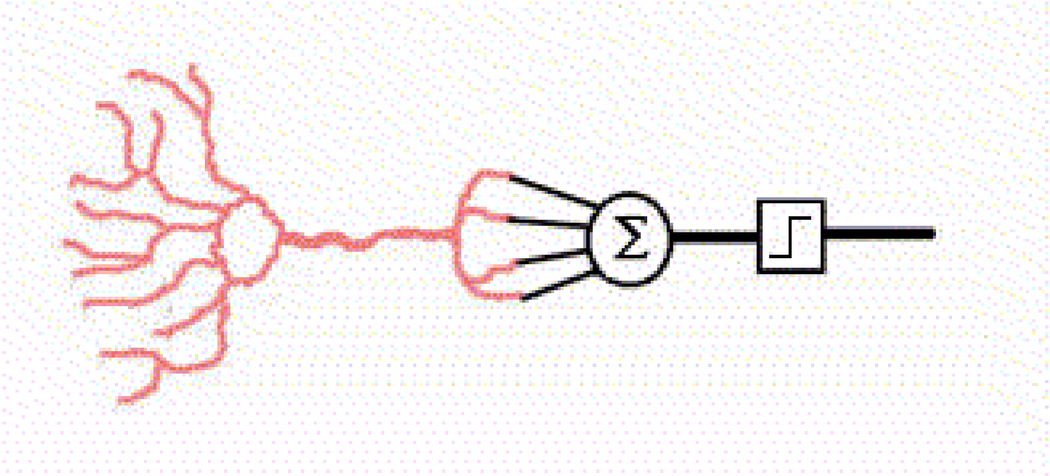


# NEURAL NETWORKS (CS010 805G02)

Shiney Thomas
AP,CSE,AJCE

# Introduction

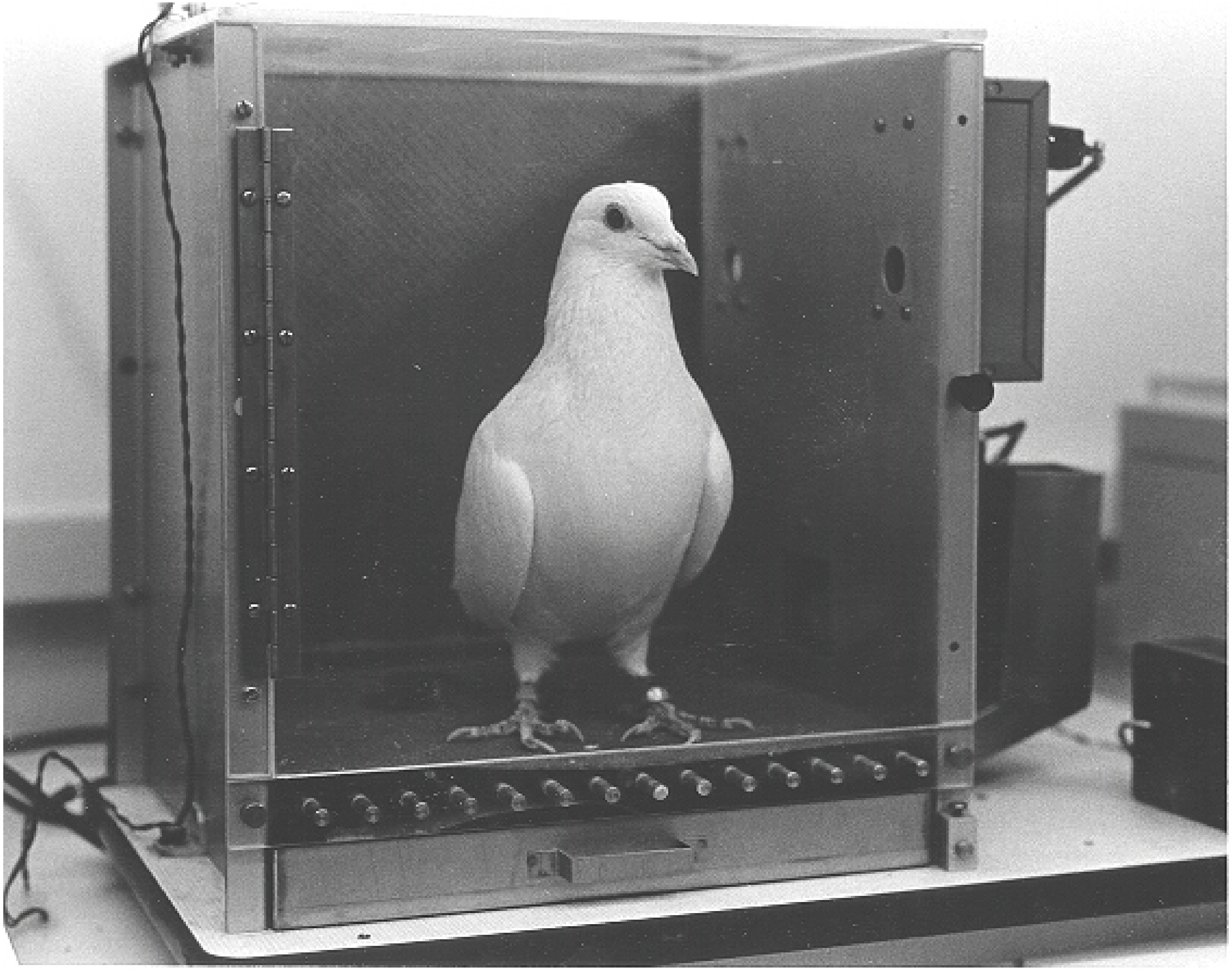
- Trends in Computing
  - Technology
  - Software
  - Architecture
  - AI Concepts
- AI systems
  - Build machines that can demonstrate intelligence similar to human beings
- Pattern Vs Data

# The brain metaphor

- The brain makes us who we are
- It is responsible for every emotion of ours, for every intellectual act of ours
- What is the nature of computation in human brain?
  - Numerical Problem
- Recognition with Partial and Noisy Information
  - Who are these?
- Abstraction
- Optical Illusions –
  - Kanizsa Square Illusion - One more

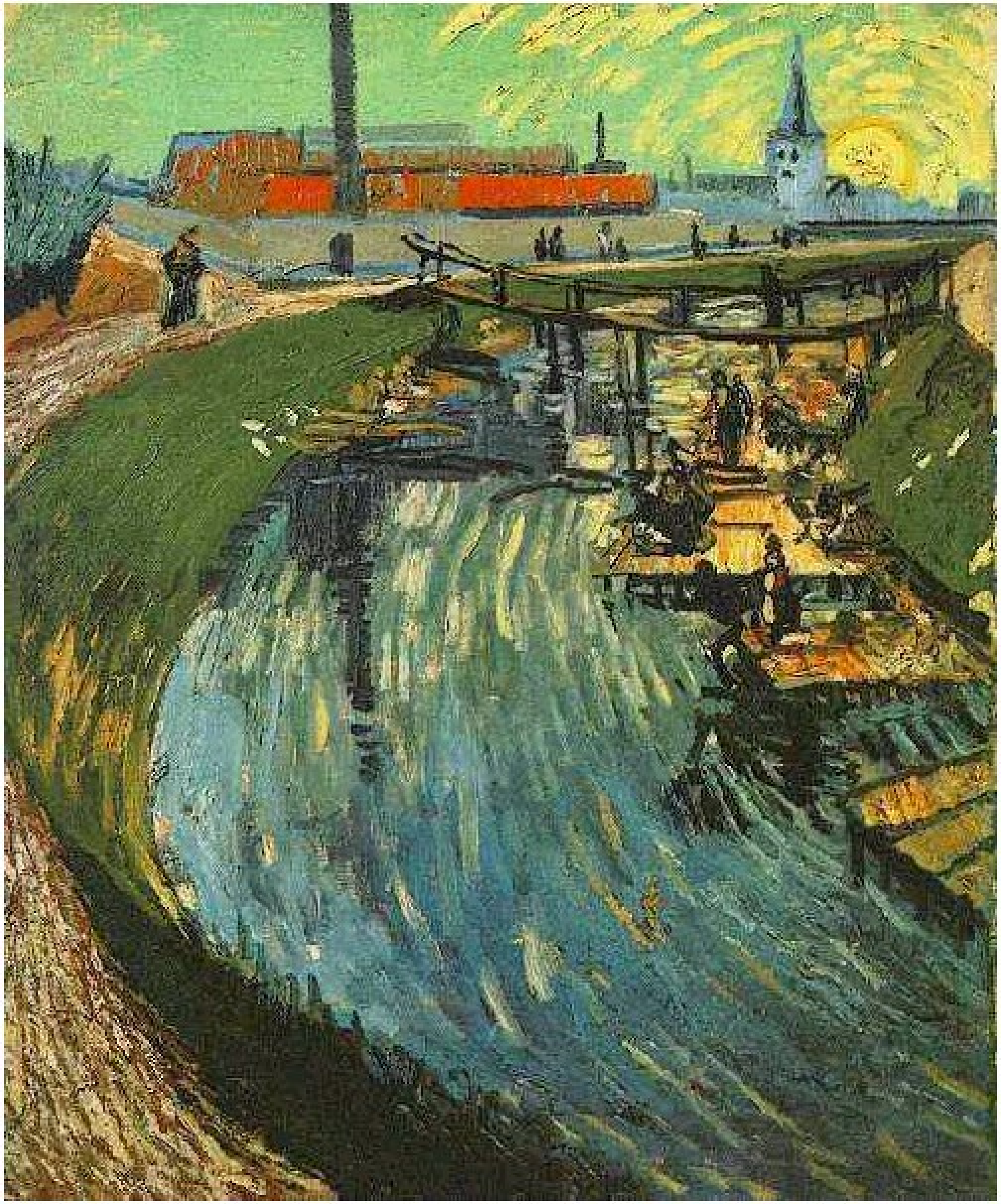
# Inspiration from pigeons- generalization

- Pigeons as art experts (Watanabe *et al.* 1995)
    - Experiment:
        - Pigeon in Skinner box
        - Present paintings of two different artists (e.g. Mark Chagall / Van Gogh)
        - Reward for pecking when presented a particular artist (e.g. Van Gogh)



The Starry Night by Van Gogh

The Poet and The Birds By Chagall



The Roubine du Roi Canal with Washerwomen - Vincent Van Gogh

Russian Wedding, 1909 - Marc Chagall

- Pigeons were able to discriminate between Van Gogh and Chagall with 95% accuracy (when presented with pictures they had been trained on)
- Discrimination still 85% successful for previously unseen paintings of the artists
- Video – Pigeon Learning

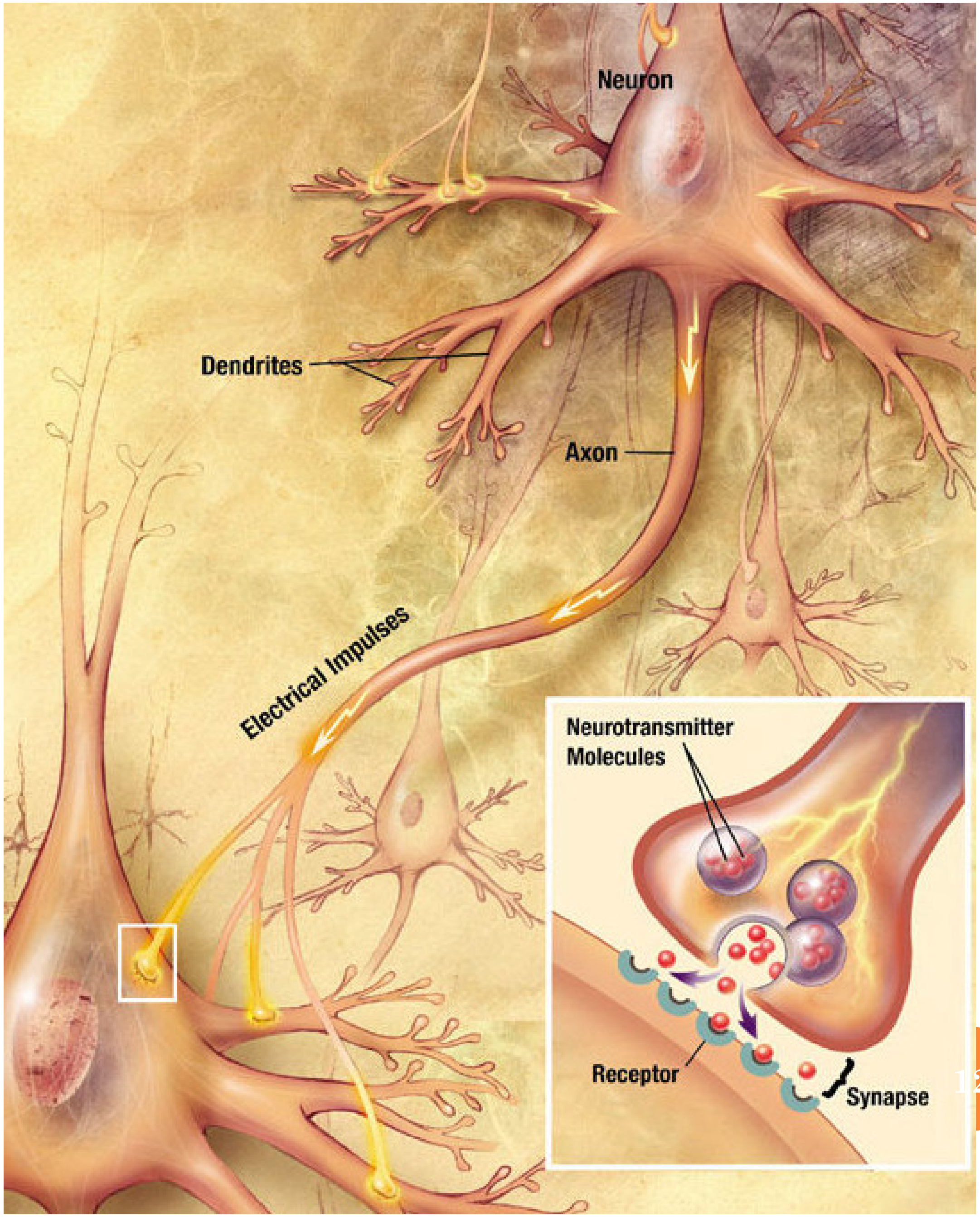
- Pigeons do not simply memorise the pictures
- They can extract and recognise patterns (the ‘style’)
- They generalise from the already seen to make predictions

- This is what neural networks (biological and artificial) are good at (unlike conventional computer)

# Neural Networks

- **Origin**: Algorithms that try to mimic the brain.
- Very widely used in 80s and early 90s; popularity diminished in late 90s.
- **Recent resurgence**: State-of-the-art technique for many applications
- Artificial neural networks are not nearly as complex or intricate as the actual brain structure

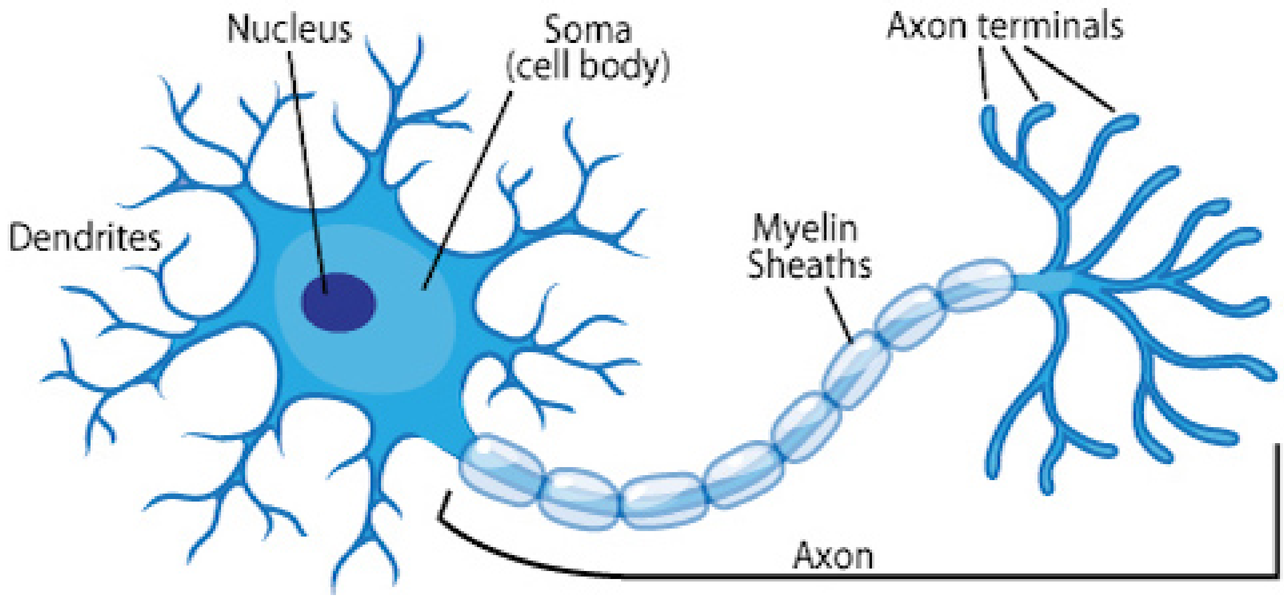
# Biological Neuron



# Features of Biological NN

1. Robustness and Fault Tolerance
2. Flexibility
3. Ability to deal with a variety of Data situations
4. Collective Computation

# Biological Neuron



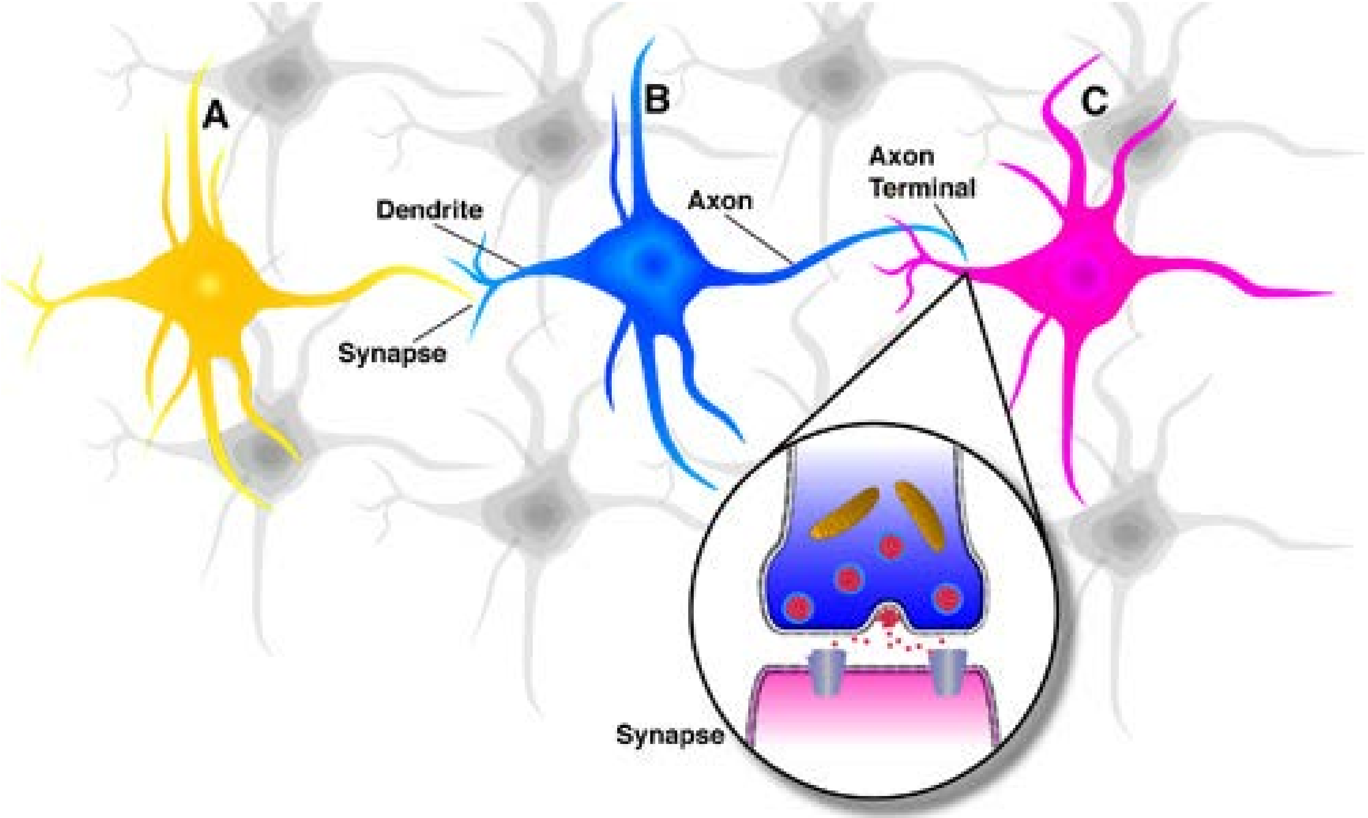

[A Quick Look]

# Biological Neuron

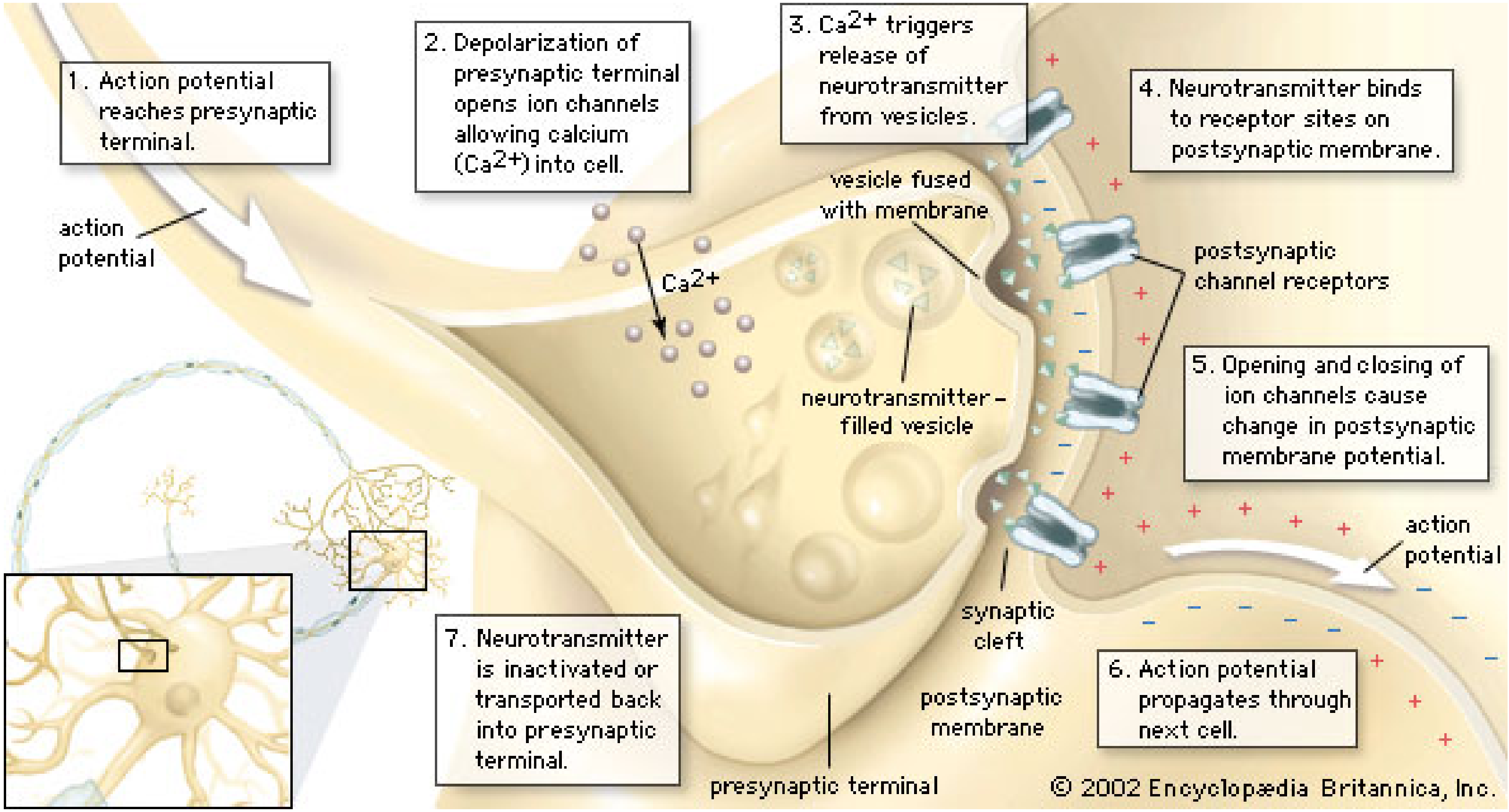
- 3 main parts

1. Soma or cell body
   - Where the cell nucleus is located
2. Dendrites
   - Where the nerve is connected to the cell body
3. Axon
   - Which carries the impulses of the neuron

# Synapse

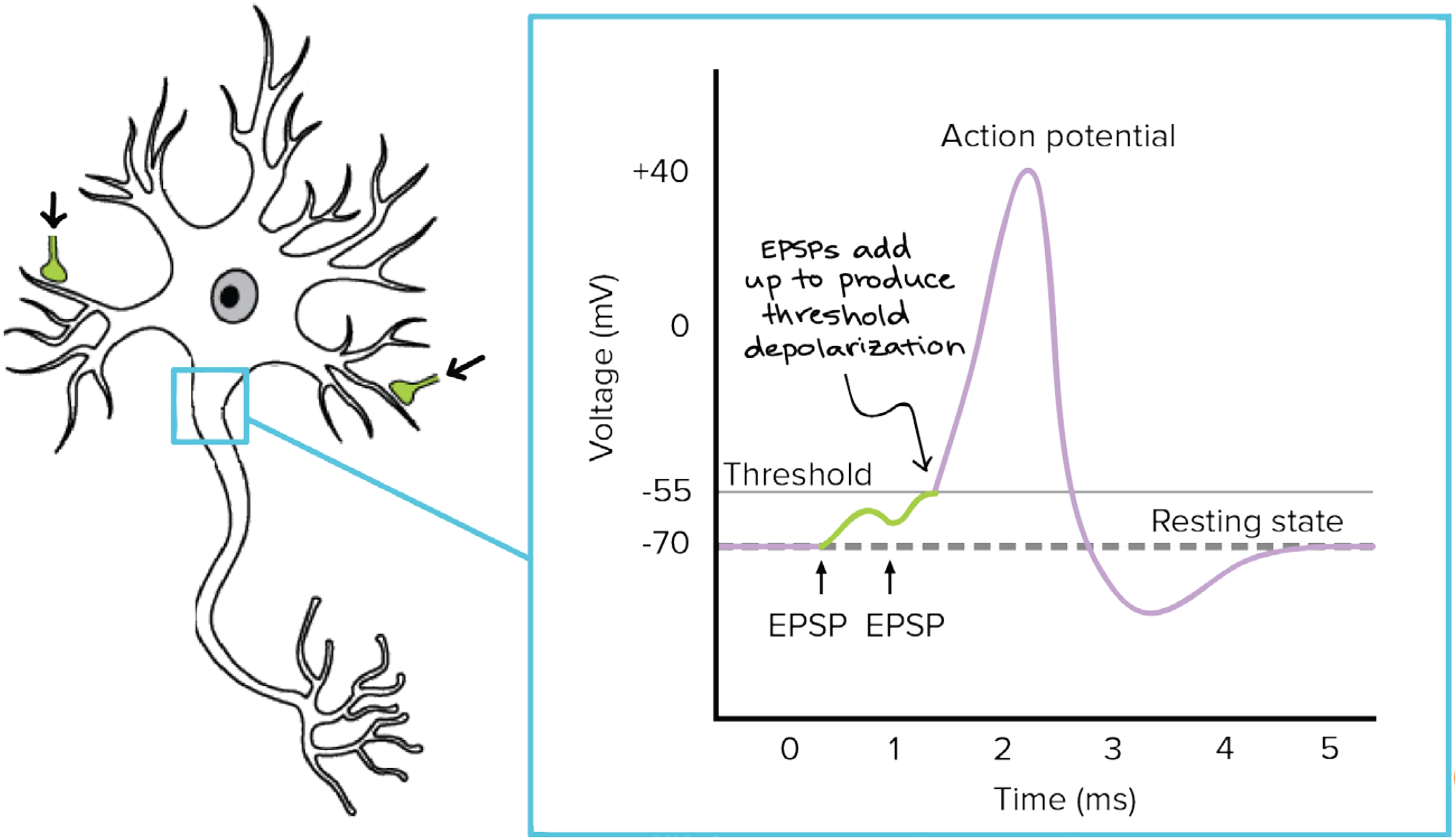


# Synapse



1. Action potential reaches presynaptic terminal.
action potential
2. Depolarization of presynaptic terminal opens ion channels allowing calcium ($Ca^{2+}$) into cell.
$Ca^{2+}$
3. $Ca^{2+}$ triggers release of neurotransmitter from vesicles.
vesicle fused with membrane
neurotransmitter-filled vesicle
4. Neurotransmitter binds to receptor sites on postsynaptic membrane.
postsynaptic channel receptors
5. Opening and closing of ion channels cause change in postsynaptic membrane potential.
action potential
synaptic cleft
postsynaptic membrane
6. Action potential propagates through next cell.
7. Neurotransmitter is inactivated or transported back into presynaptic terminal.
presynaptic terminal
© 2002 Encyclopædia Britannica, Inc.

# Synapse


Action potential
+40
0
-55
-70
Voltage (mV)
EPSPs add up to produce threshold depolarization
Threshold
Resting state
EPSP
EPSP
0
1
2
3
4
5
Time (ms)

# Artificial Neural Networks

- Neural networks are a new method of programming computers.
- They are exceptionally good at performing pattern recognition and other tasks that are very difficult to program using conventional techniques.
- Programs that employ neural nets are also capable of learning on their own and adapting to changing conditions.

# Definition of NN

**According to the DARPA Neural Network Study (1988, AFCEA International Press, p. 60):**

- • ... a neural network is a system composed of many simple processing elements operating in parallel whose function is determined by network structure, connection strengths, and the processing performed at computing elements or nodes.

**According to Haykin (1994), p. 2:**

- A neural network is a massively parallel distributed processor that has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two respects:
  - Knowledge is acquired by the network through a learning process.
  - Interneuron connection strengths known as synaptic weights are used to store the knowledge.

# Performance Comparison of Computer and Biological NN

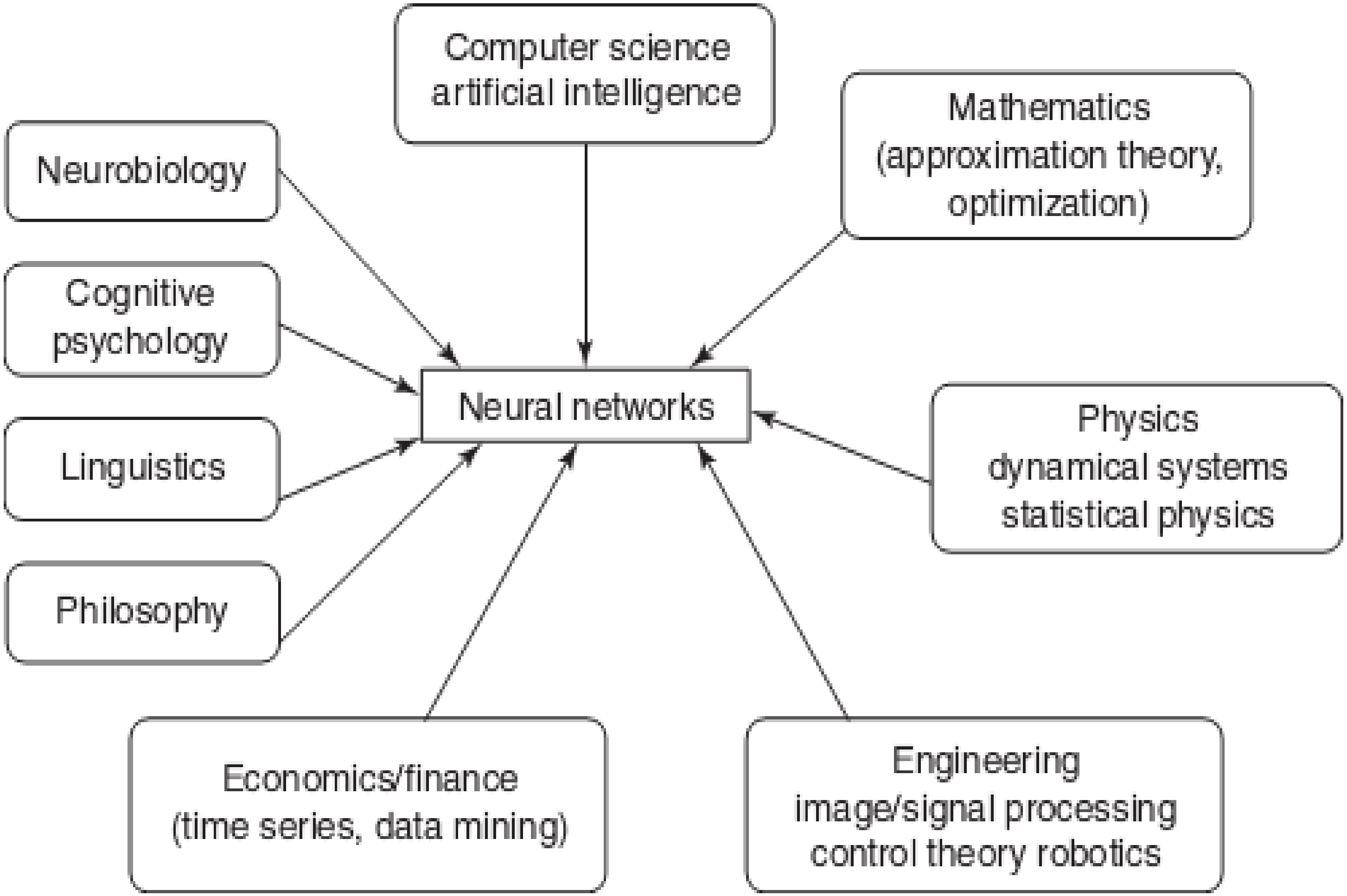
- Speed
- Processing
- Size and Complexity
- Storage
- Fault Tolerance
- Control Mechanism

# Comparison of Brains and Traditional Computers

- 200 billion neurons, 32 trillion synapses
- Element size: $10^{-6}$ m
- Energy use: 25W
- Processing speed: 100 Hz
- Parallel, Distributed
- Fault Tolerant
- Learns: Yes
- Intelligent/Conscious: Usually

- 1 billion bytes RAM but trillions of bytes on disk
- Element size: $10^{-9}$ m
- Energy watt: 30-90W (CPU)
- Processing speed: $10^9$ Hz
- Serial, Centralized
- Generally not Fault Tolerant
- Learns: Some
- Intelligent/Conscious: Generally No

# Multidisciplinary View Of Neural Networks



# History of NN

- 1943: McCulloch–Pitts “neuron”
  - Started the field
- 1962: Rosenblatt’s perceptron
  - Learned its own weight values; convergence proof
- 1969: Minsky & Papert book on perceptrons
  - Proved limitations of single-layer perceptron networks
- 1982: Hopfield and convergence in symmetric networks
  - Introduced energy-function concept
- 1986: Backpropagation of errors
  - Method for training multilayer networks
- Present: Probabilistic interpretations, Bayesian and spiking networks

# History of NN

**Table 1.1** Historical Development of Neural Network Principles

| Key developments | Other significant contributions |
|---|---|
| McCulloch and Pitts (1943) | von Neumann (1946)—General purpose electronic computer |
| • Model of neuron | Norbert Weiner (1948)—Cybernetics |
| • Logic operations | Shannon (1948)—Information theory |
| • Lack of learning | Ashby (1952)—Design for a Brain |
| Hebb (1949) | Gabor (1954)—Nonlinear adaptive filter |
| | Uttley (1956)—Theoretical machine |
| • Synaptic modifications | Caianiello (1961)—Statistical theory and learning |
| • Hebb's learning law | |

# History of NN

Minsky (1954)

- Learning machines

Rosenblatt (1958)

- Perceptron learning and convergence
- Pattern classification
- Linear separability constraint

Widrow and Hoff (1960)

- Adaline—LMS learning
- Adaptive signal processing

Minsky (1961)—Artificial intelligence

Steinbuch (1961)—Learnmatrix

Minsky and Selfridge (1961)—Credit assignment problem

Nilsson (1965)—Learning machine

Amari (1967)—Mathematical solution to credit assignment

Kohonen (1971)—Associative memories

Willshaw (1971)—Self-organization and generalization

Malsburg (1973)—Self-organization

Tikhonov (1973)—Regularization theory

Little (1974)—Ising model and neural network

Grossberg (1976)—Adaptive resonance theory

# History of NN

Minsky and Papert (1969)

- Perceptron—Multilayer perceptron (MLP)
- Hard problems
- No learning for MLP

Werbos (1974)

- Error backpropagation

Hopfield (1982)

- Energy analysis

Anderson (1977)—Brain state-in-box model

Little and Shaw (1978)—Stochastic law for NN, spin glasses

Fukushima (1980)—Neocognitron

Kohonen (1982)—Feature mapping

Barto, Sutton and Anderson (1983)—Reinforcement learning

Kirkpatrick (1983)—Simulated annealing

Peretto (1984)—Stochastic units

Mead (1985)—Analog VLSI

Amit (1985)—Statistical machines and stochastic networks

# History of NN

Ackley, Hinton and Sejnowski (1985)

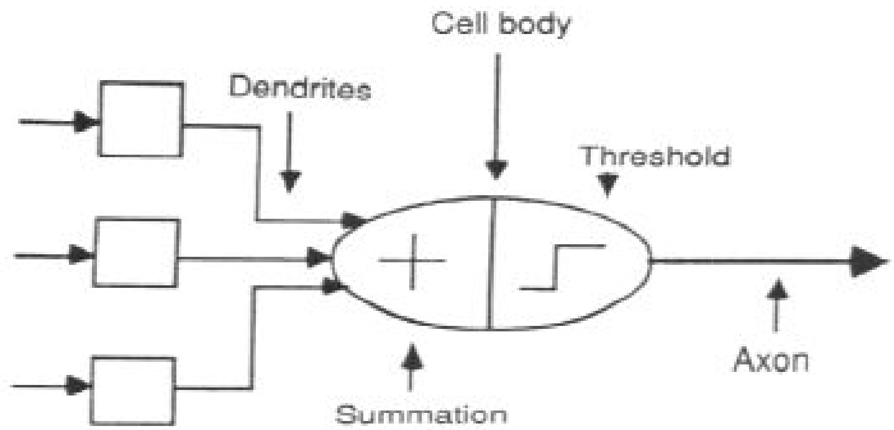
- Boltzmann machine

Rumelhart, Hinton and Williams (1986)

- Generalised delta rule

Klopf (1986)—Drive-reinforcement learning
Hecht-Nielsen (1987)—Counterpropagation
Linsker (1988)—Self-organization based on information preservation
Kosko (1988)—BAM, Fuzzy logic in ANN
Broomhead (1988)—Radial basis functions (RBF)
Poggio and Girosi (1990)—RBF and regularization theory

# ARCHITECTURE OF ANN



# The Neuron

- The neuron is the basic information processing unit of a NN. It consists of:
  1. A set of synapses or connecting links, each link characterized by a weight:

     $$W_1, W_2, \ldots, W_m$$

  2. An adder function (linear combiner) which computes the weighted sum of the inputs:

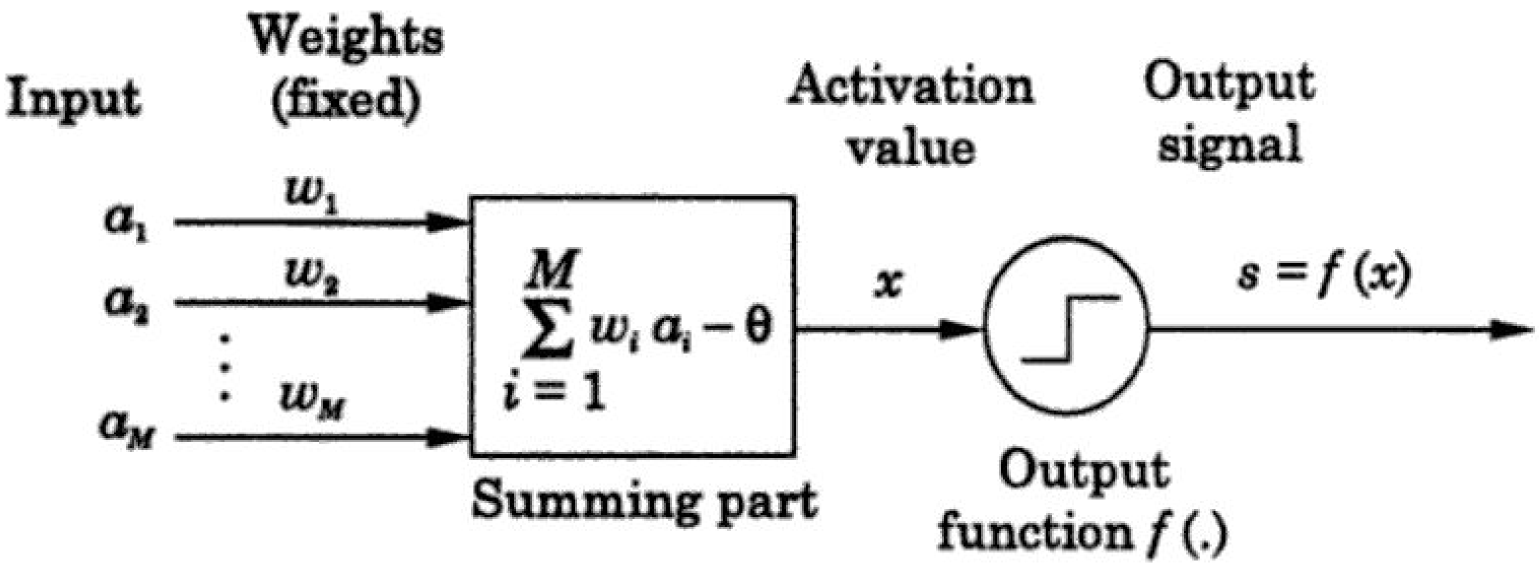
     $$u = \sum_{j=1}^{m} w_j x_j$$

  3. Activation function (squashing function) $\varphi$ for limiting the amplitude of the output of the neuron.

     $$y = \varphi(u + b)$$

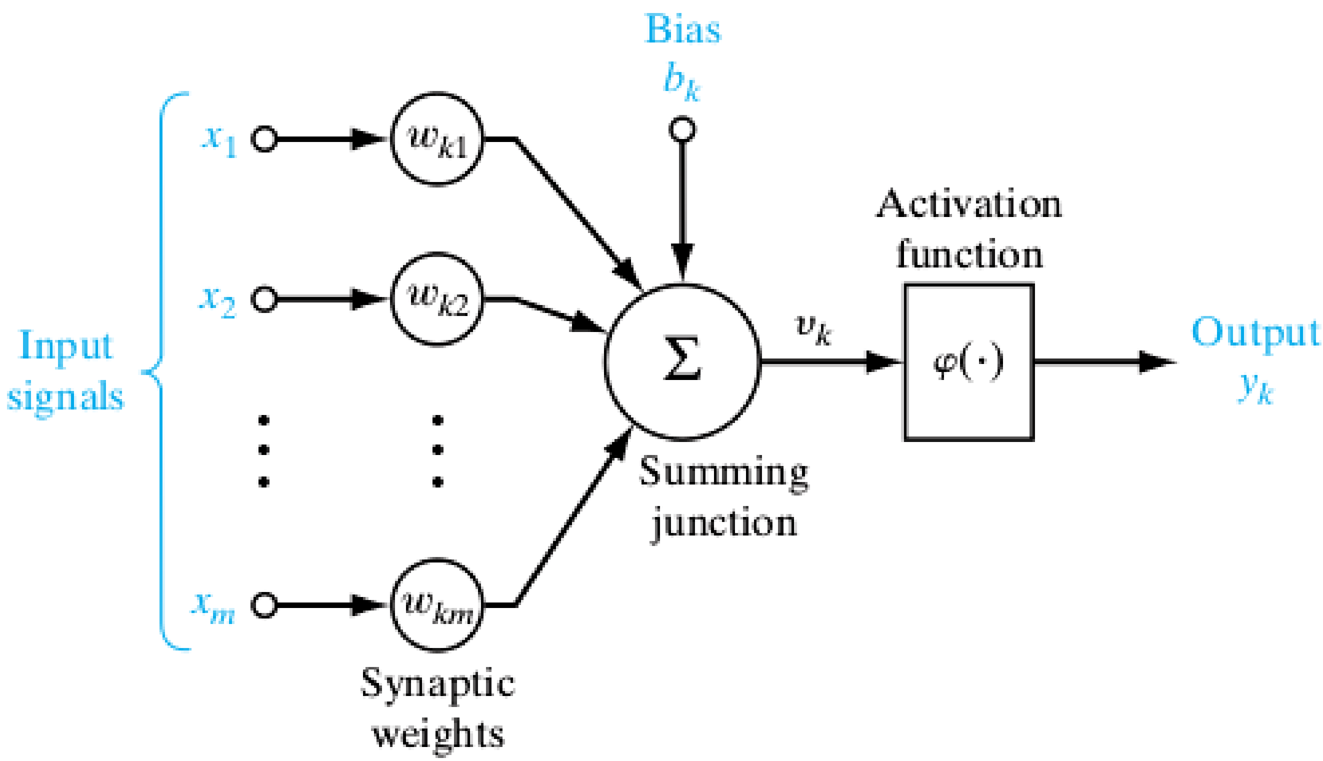
# Models Of Neurons

- McCulloch-Pits Model (MP)



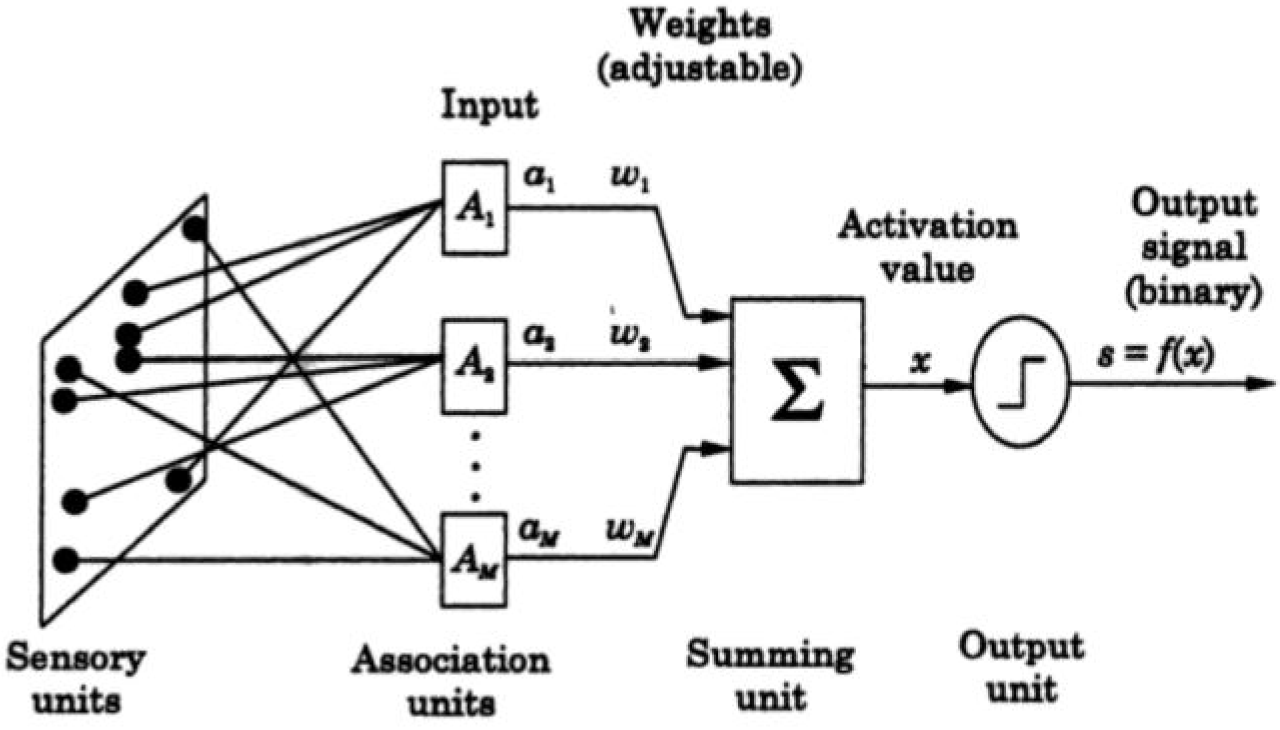
# Models Of Neurons

- McCulloch-Pits Model (MP)



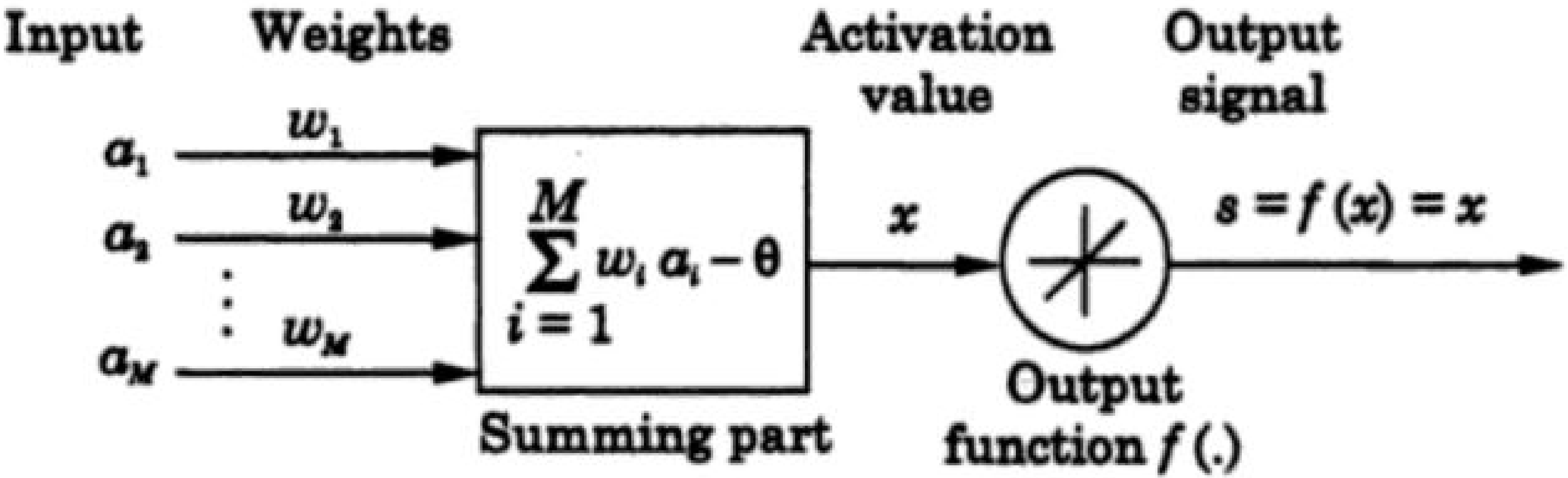
# Models Of Neurons

- Perceptron (Rosenblatt's)



# Models Of Neurons

- Adaline(ADAptive LINear Element)- Widrow's



# Dimensions of a Neural Network

- Various types of neurons
- Various network architectures
- Various learning algorithms
- Various applications

# Network Architectures

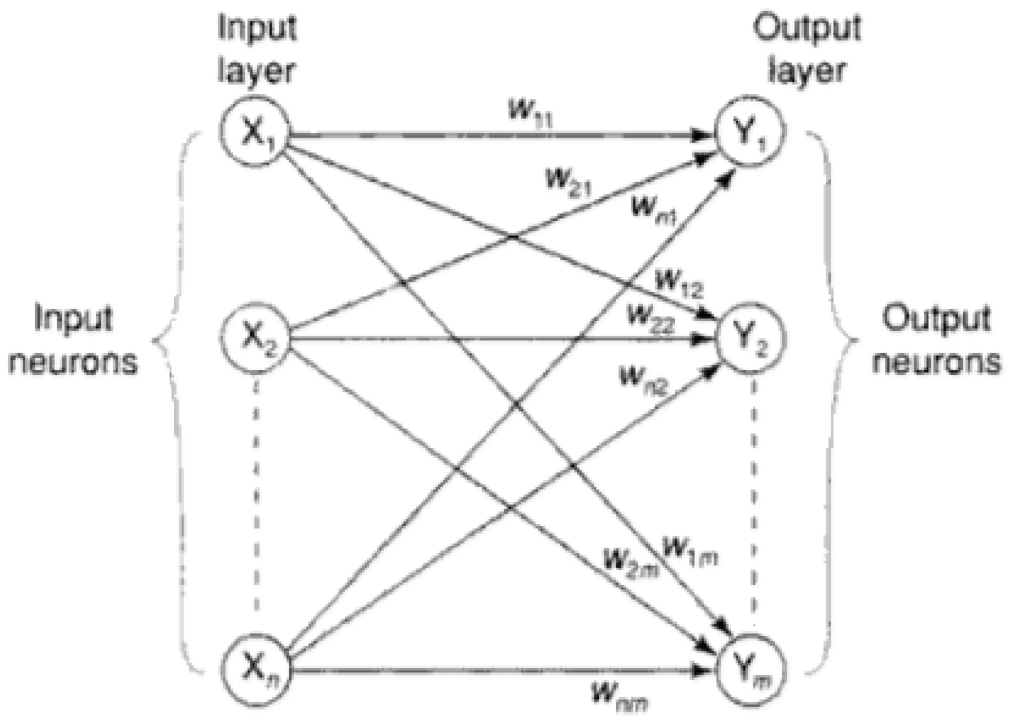
- Three different classes of network architectures

  - single-layer feed-forward
  - multi-layer feed-forward
  - Recurrent networks

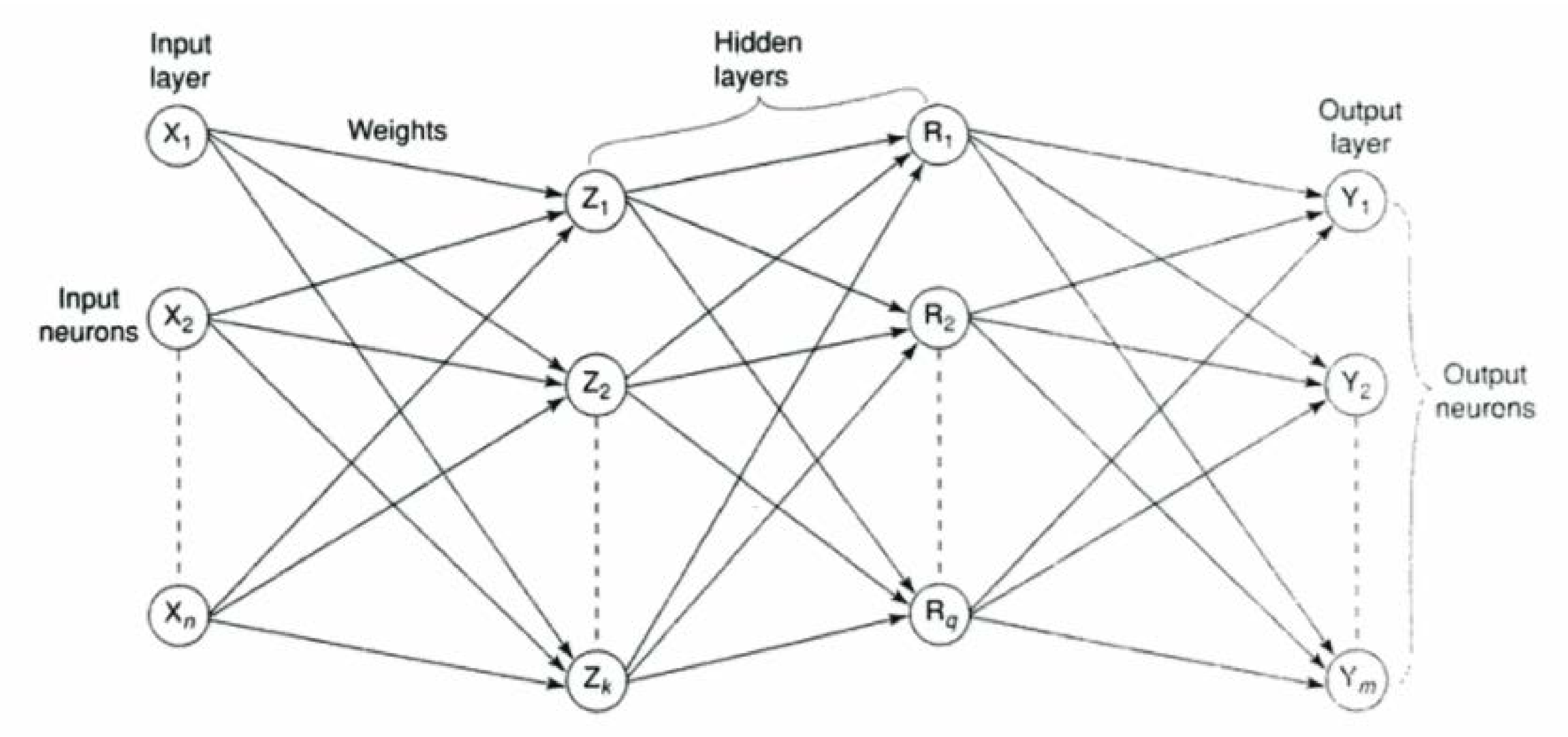
  neurons are organized in acyclic layers

- The architecture of a neural network is linked with the learning algorithm used to train

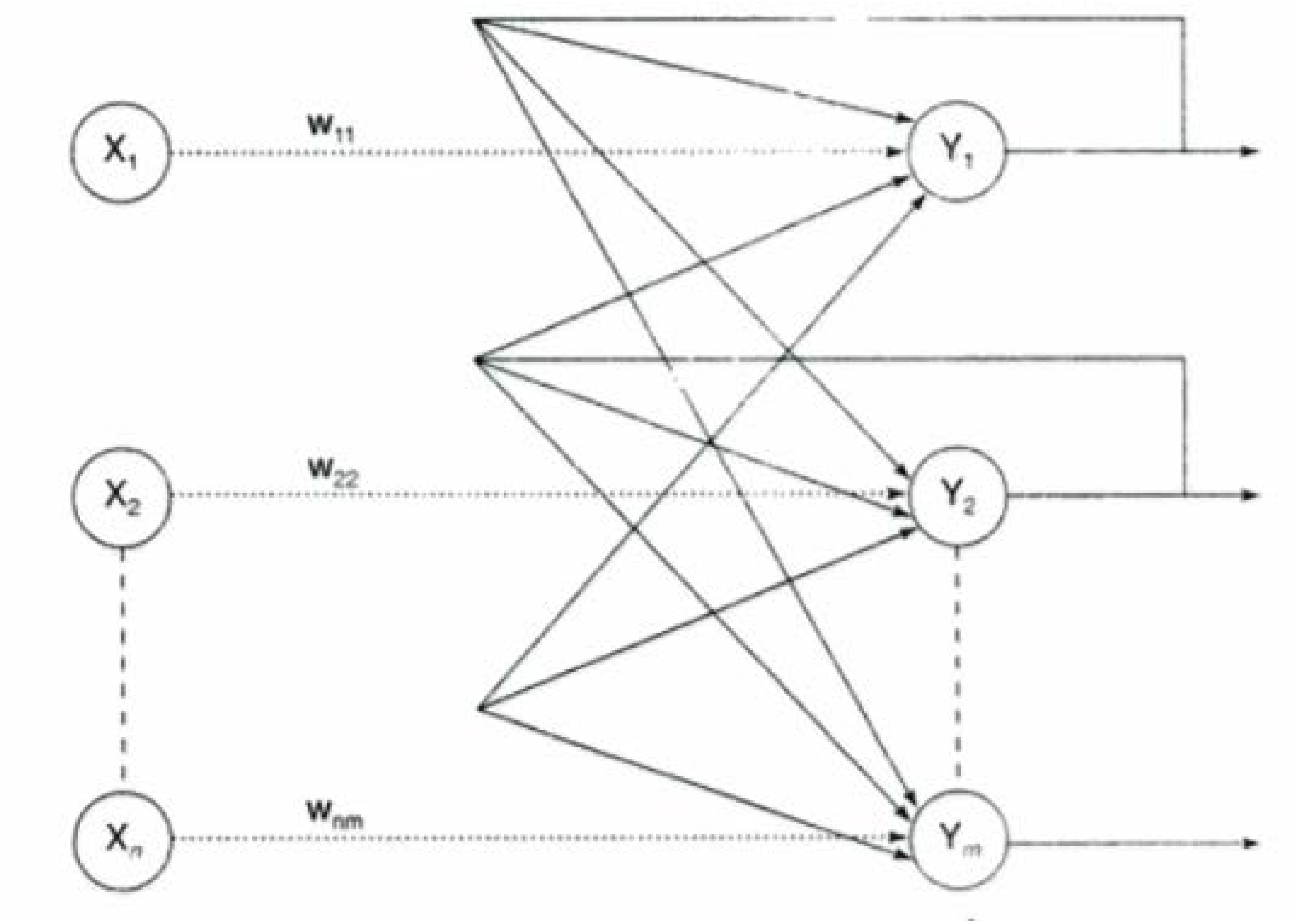
# SINGLE LAYER FEED-FORWARD NETWORK



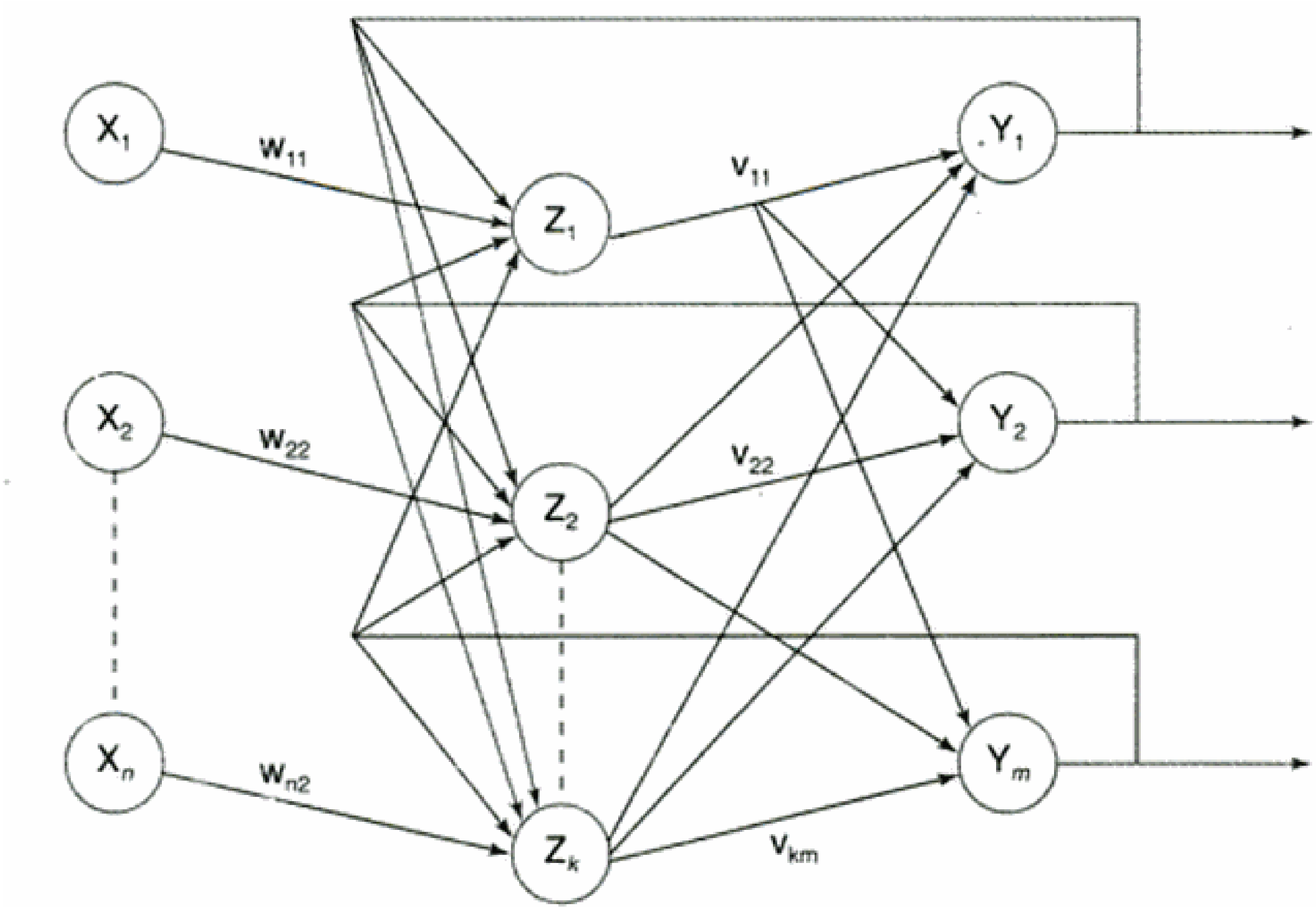
# Multi Layer Feed-Forward Network



# Single Layer Recurrent Network



# Multi Layer Recurrent Network



## Learning

- Supervised
- Unsupervised
- Reinforcement Learning

# Supervised Vs Unsupervised learning

- Supervised Learning
  - Recognizing hand-written digits, pattern recognition, regression.
  - Labeled examples
    (input , desired output)
  - Neural Network models: perceptron, feed-forward, radial basis function, support vector machine.
- Unsupervised Learning
  - Find similar groups of documents in the web, content addressable memory, clustering.
  - Unlabeled examples
    (different realizations of the input alone)
  - Neural Network models: self organizing maps, Hopfield networks.