Department of Computer Engineering
University of Kurdistan

Neural Networks (Graduate level)

Introduction

By: Dr. Alireza Abdollahpouri

Course Info

- **Instructor**

  Dr. Alireza Abdollahpouri

  Email: abdollahpouri@gmail.com

  Office: Room 219

- **Course Web Page**

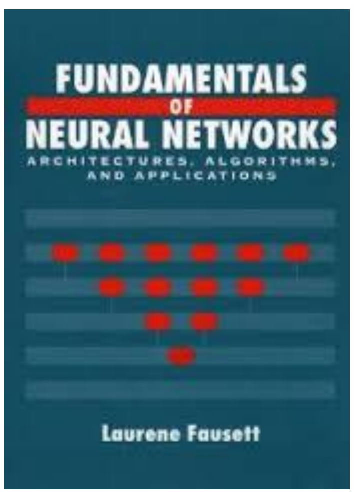
  http://eng.uok.ac.ir/abdollahpouri/NN.html

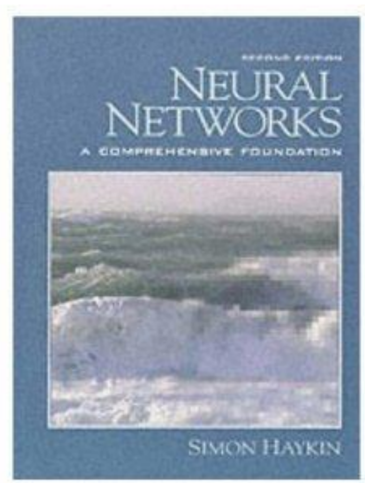
- **Grading Policy**

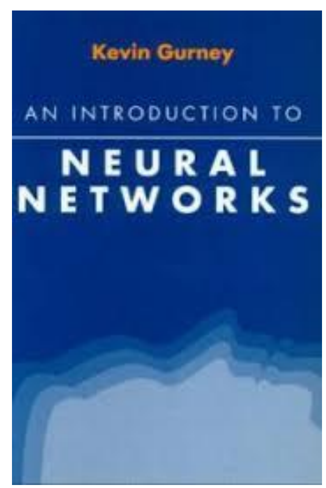
Projects	25%
Presentation	15%
Final exam	55%
Class participation	5%

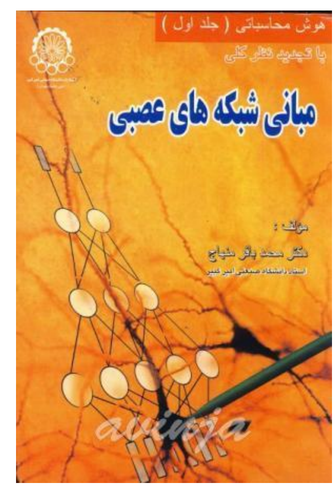
Course Info

- ## References
  - Neural Networks and Learning Machines, Simon O. Haykin
  - Fundamentals of Neural Networks: Architectures, Algorithms And Applications, Laurene V. Fausett
  - An Introduction to Neural Networks, Kevin Gurney
  - M.B. Menhaj: Neural Networks (in Persian)









Some Ground Rules

- Let's make this educational and enjoyable.
- It's a big size class, I enjoy questions and ideas from the class.
  - Ask questions and raise points.
  - Listen to other people's questions.
  - Be here.
  - Be here on time.

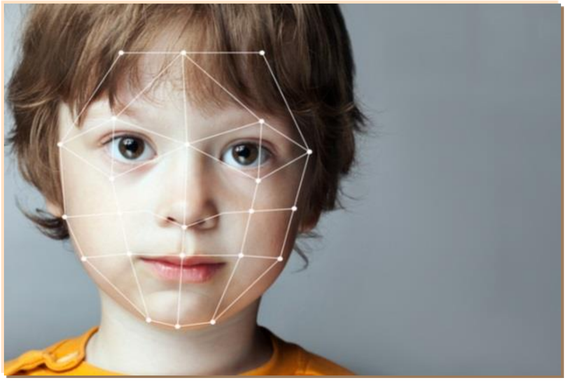
Humans vs. Computers

Computers are excellent in computing and searching, but they miserably fail when programmed to reproduce typical human activities:

- Sensory perception
- Sensory-motor coordination
- Image recognition
- Adaptive behaviors

Babies beat Computers

Although a PC can beat the world's chess champion, the most powerful computer in the planet is not capable to compete with a 3-year baby in



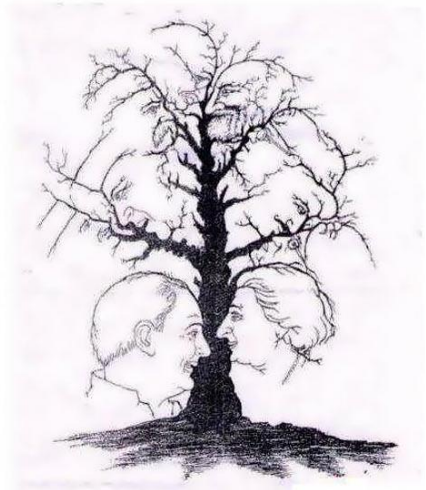
Face recognition

Voice recognition

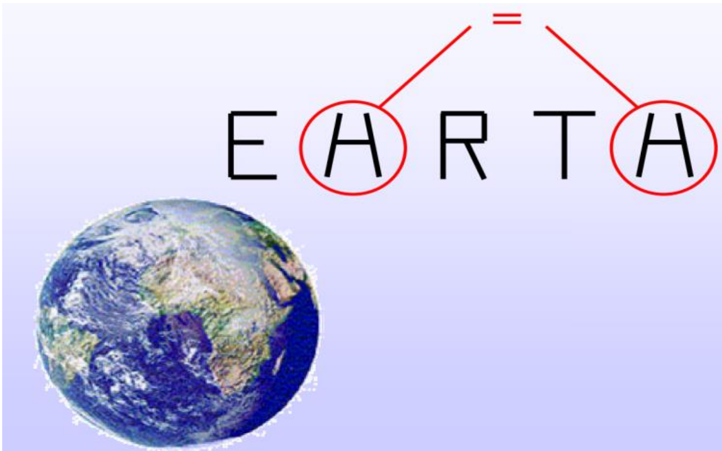
Lego construction

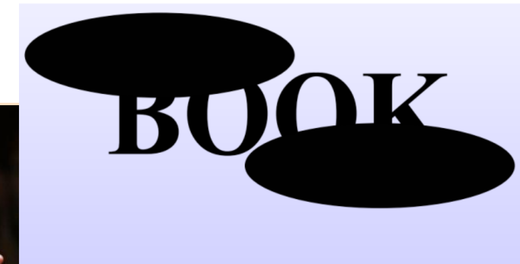
Capabilities of Human brain

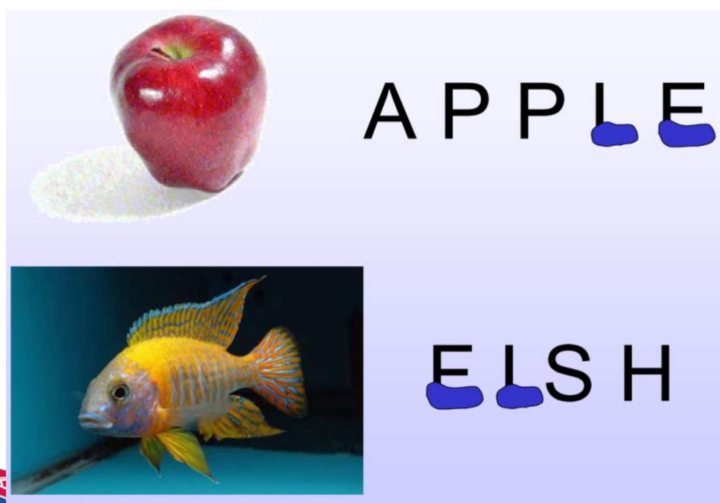
- Face recognition
- Learning
- Decision making
- Calculation
- Information storage
- …



Capabilities of Human brain







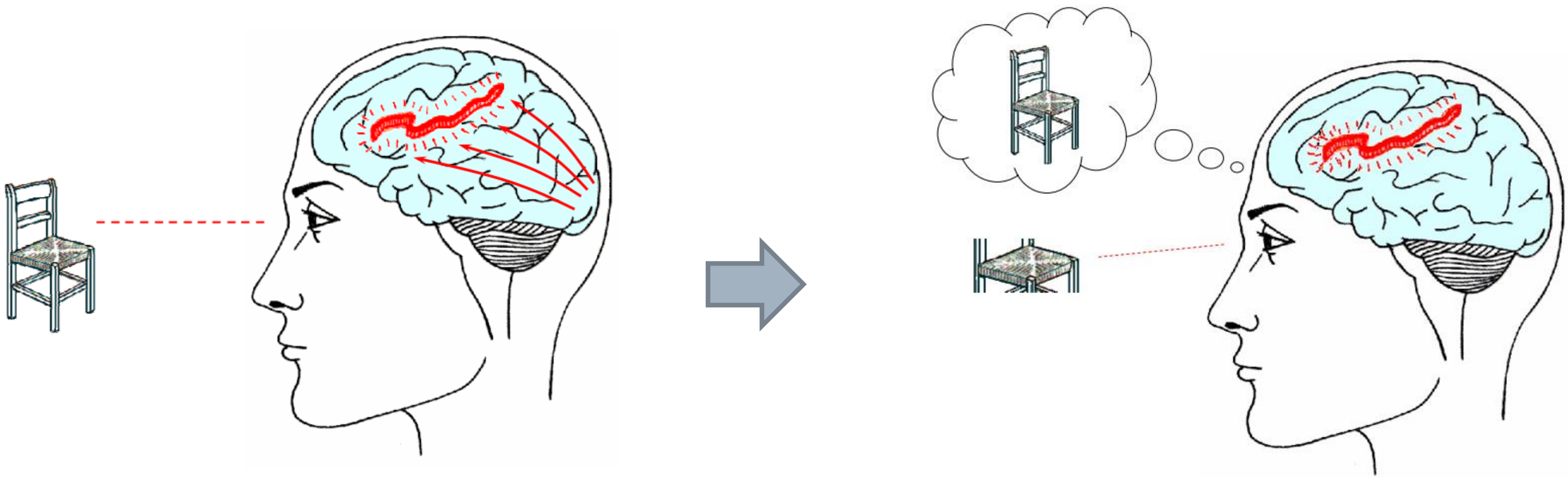
Capabilities of Human brain

The mouse on the desk is broken

The mouse on the desk is eating cheese

How does the brain work?

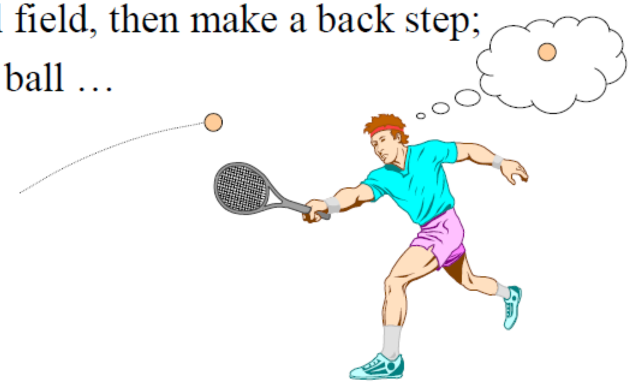
- When we recognize a face or we grasp an object we do not solve equations.
- The brain works using an **associative** process.



Hitting a tennis ball

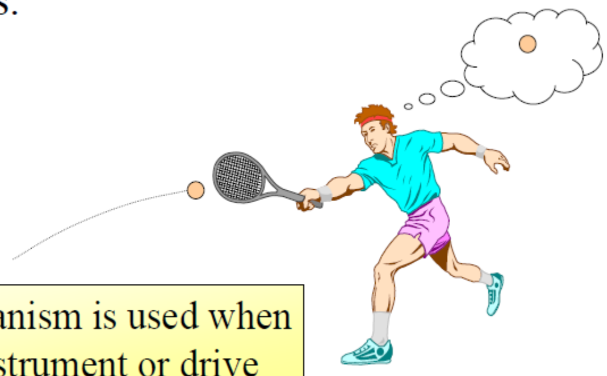
Learning phase

- In a learing phase, we try several actions and store the good ones:
  - If the ball is seen in the upper-left area of the visual field, then make a back step;
  - If the ball …



Operating Phase

- Once trained, the brain executes the actions *without thinking*, based on the learned associations.



A similar mechanism is used when we play an instrument or drive

Quotation

"If the human brain were simple enough for us to understand, we would still be so stupid that we couldn't understand it"

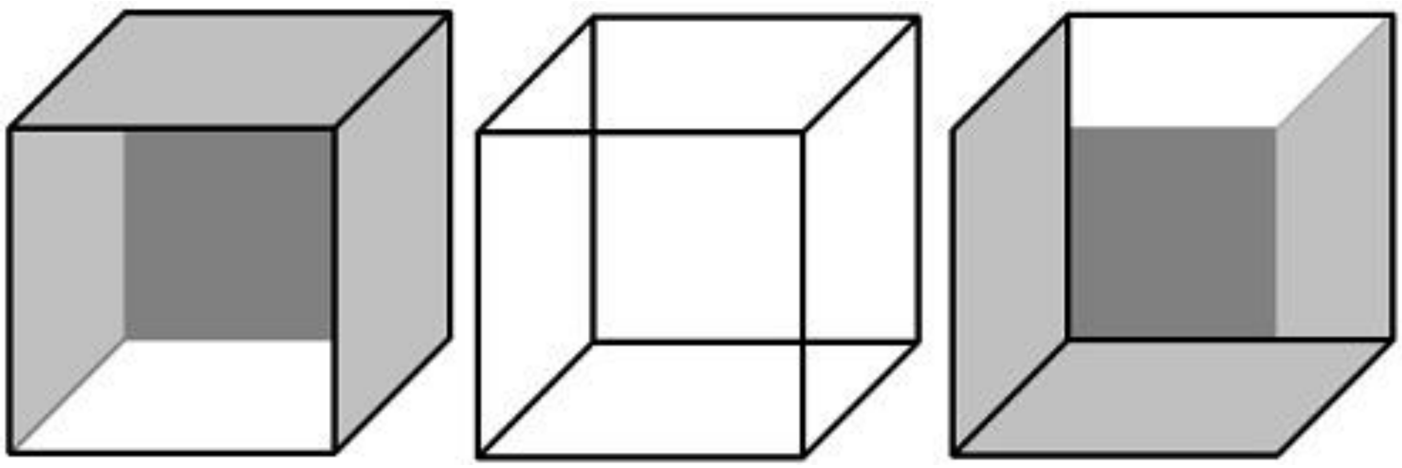
Jostein Gaarder

Digital versus biological computers

5 distinguishing properties

- Speed
- Robustness
- Flexibility
- Adaptivity
- Context-sensitivity

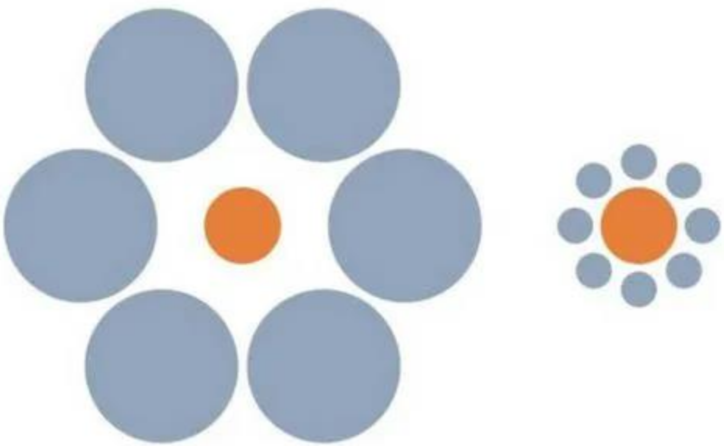
Flexibility: the Necker cube



Adaptivitiy

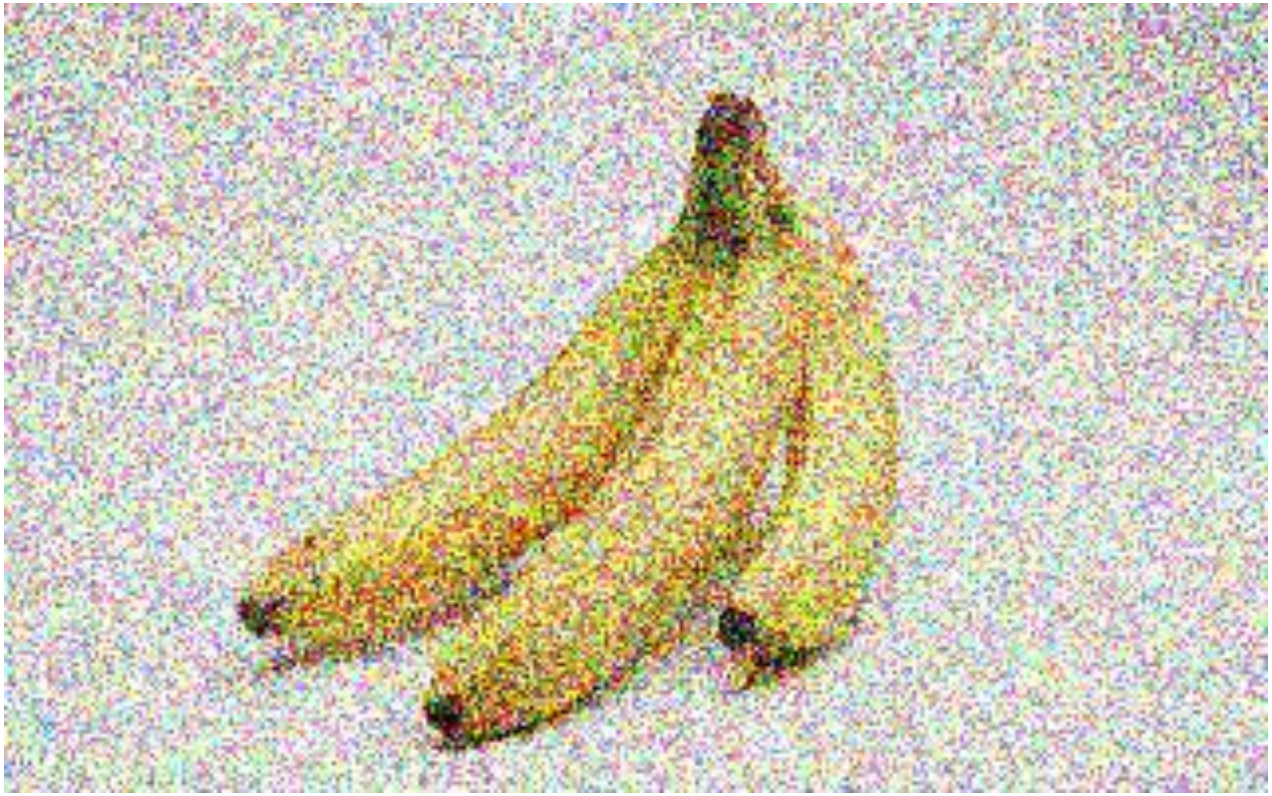
Processing implies learning
in biological computers
versus
processing does not imply learning
in digital computers

Context-sensitivity: patterns



emergent properties

Robustness and context-sensitivity coping with noise



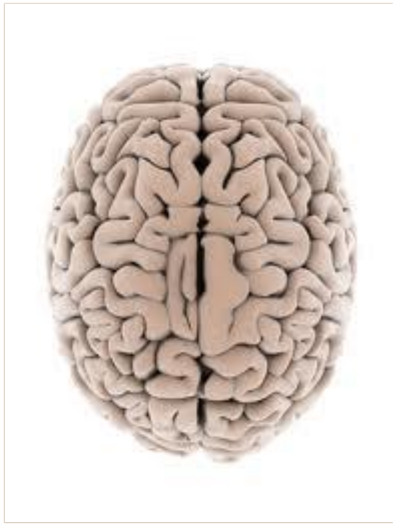
Properties of human brain

- Has self-organizing or self-learning capability. It
- can easily adjust to new environment by “learning” usually from experience or from a knowledge of its surrounding.
- Robust and fault tolerant.
- Parallel distributed architecture.
- Can deal with information that is fuzzy, probabilistic, noisy, or even inconsistent (nonlinear).
- Small, compact and dissipates very little power.

Biological Inspiration

Idea : To make the computer more robust, intelligent, and learn, …
Let’s model our computer software (and/or hardware) after the brain

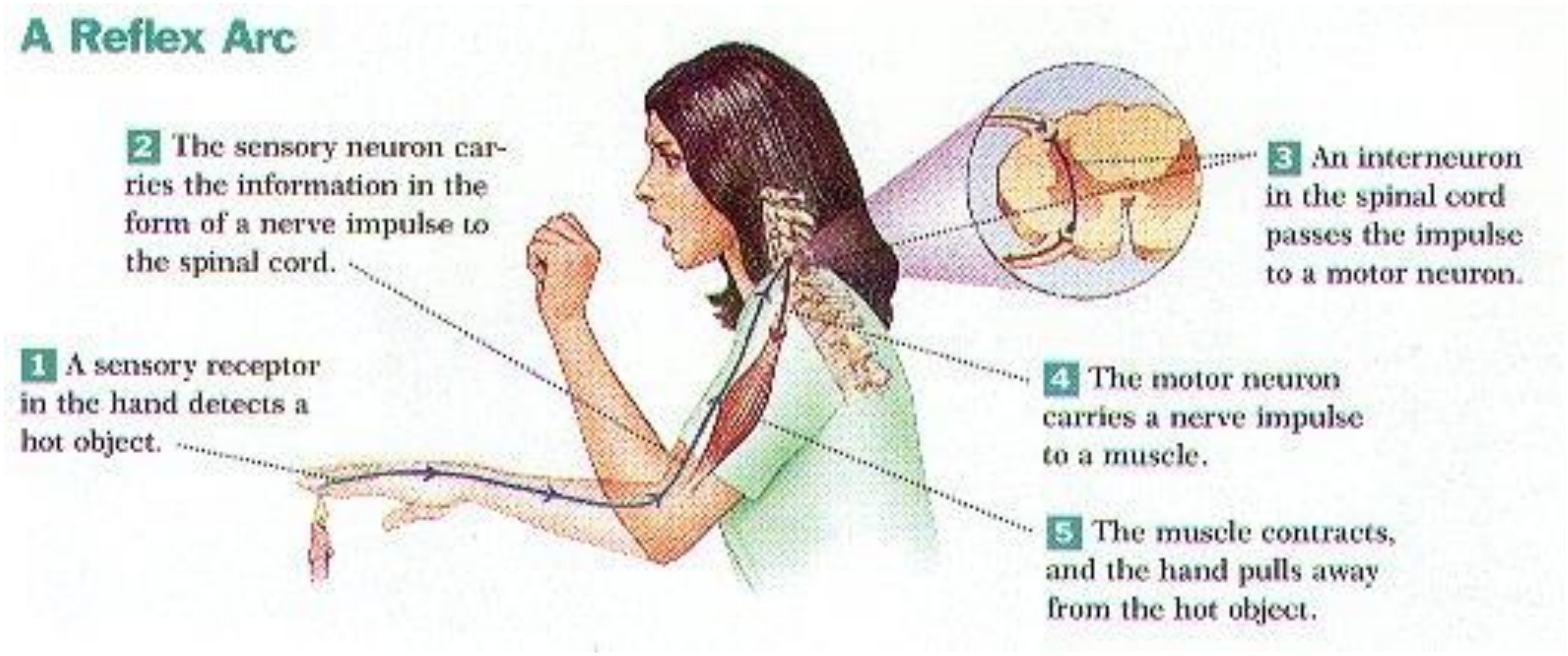
- **about 86 billion neurons** in the human cortex each connected to , on average, **10000** others.
- In total **80-1000 trillion synapses** of connections.
- The brain is a highly complex, nonlinear and parallel computer (information-processing system)

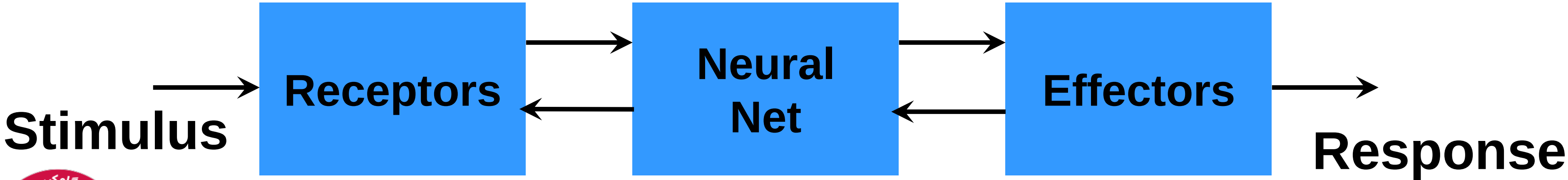


Biological inspiration

- Animals are able to react adaptively to changes in their external and internal environment, and they use their nervous system to perform these behaviours.
- An appropriate model/simulation of the nervous system should be able to produce similar responses and behaviours in artificial systems.
- The nervous system is build by relatively simple units, the neurons, so copying their behavior and functionality should be the solution.

Human nervous structure

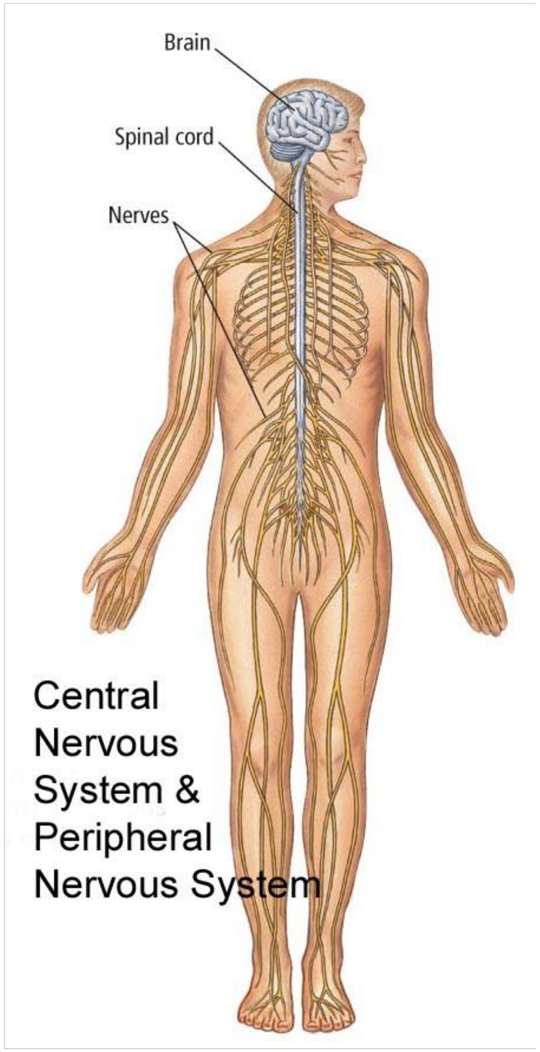




Human nervous structure

The real structure of the human nervous corresponding to last block-diagram.

- **It contains the neurons to transfer the signal form the receptors to brain and vice-versa to the effectors.**

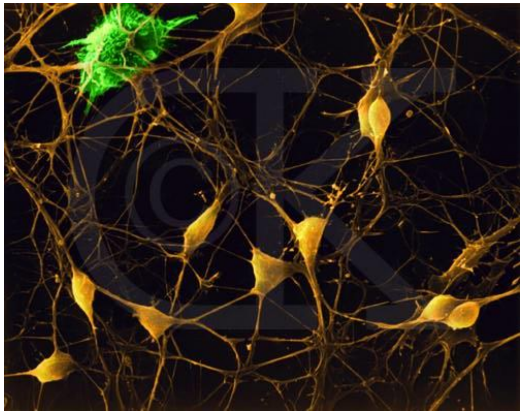


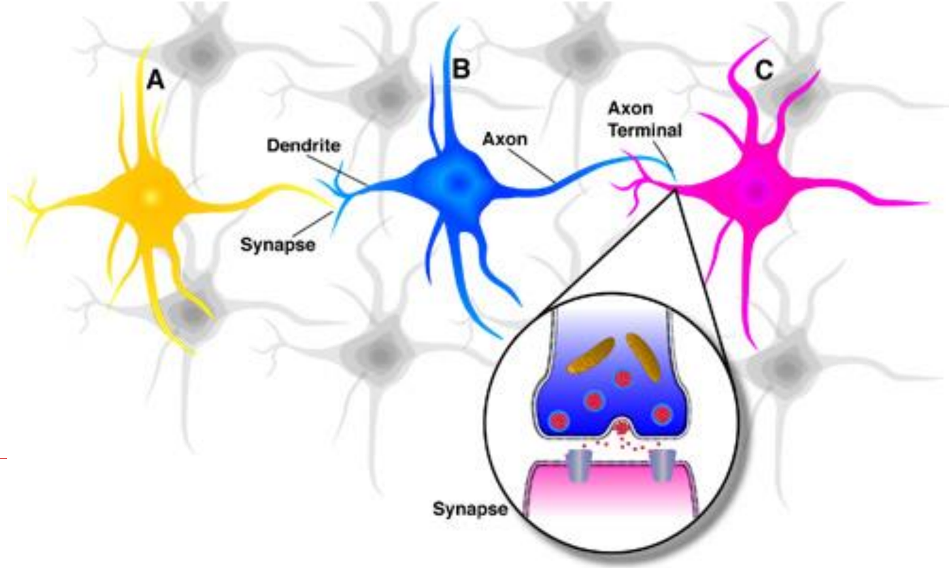
Learning in the Brain

- Brains learn
  - Altering strength between neurons
  - Creating/deleting connections
- Hebb's Postulate (Hebbian Learning)
  - When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased.
- Long Term Potentiation (LTP)
  - Cellular basis for learning and memory
  - LTP is the long-lasting strengthening of the connection between two nerve cells in response to stimulation
  - Discovered in many regions of the cortex

Neurons in the Brain

- Although heterogeneous, at a low level the brain is composed of neurons
  - A neuron receives input from other neurons (generally thousands) from its synapses
  - Inputs are approximately summed
  - When the input exceeds a threshold the neuron sends an electrical spike that travels that travels from the body, down the axon, to the next neuron(s)





Principles of Brain Processing

A process of pattern recognition and pattern manipulation is based on:

Massive parallelism

Brain computer as an information or signal processing system, is composed of a large number of a simple processing elements, called neurons. These neurons are interconnected by numerous direct links, which are called connection, and cooperate which other to perform a parallel distributed processing (PDP) in order to soft a desired computation tasks.

Connectionism

Brain computer is a highly interconnected neurons system in such a way that the state of one neuron affects the potential of the large number of other neurons which are connected according to weights or strength. The key idea of such principle is the functional capacity of biological neural nets deters mostly not so of a single neuron but of its connections

Associative distributed memory

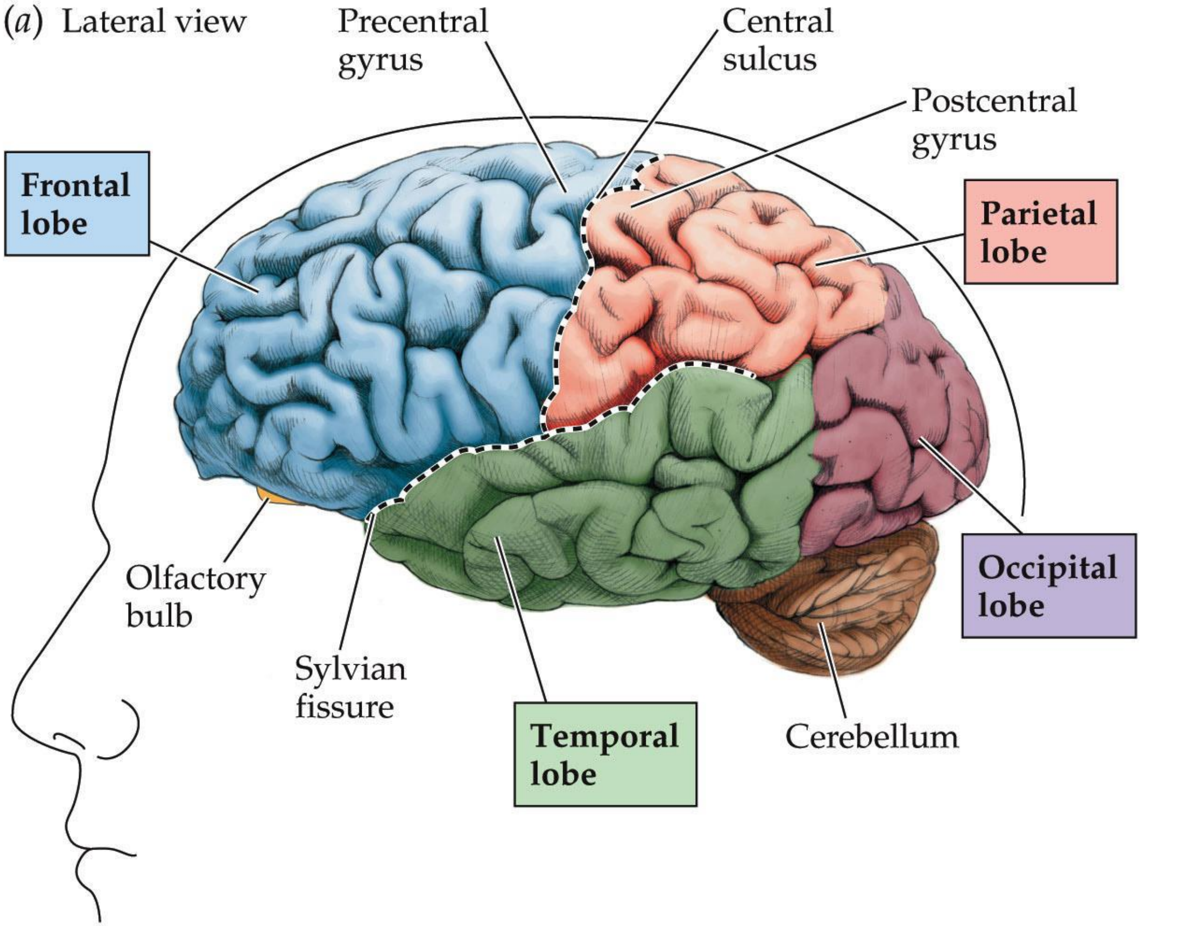
Storage of information in a brain is supposed to be concentrated in synaptic connections of brain neural network, or more precisely, in the pattern of these connections and strengths (weights) of the synaptic connections.

The Brain vs. Computer

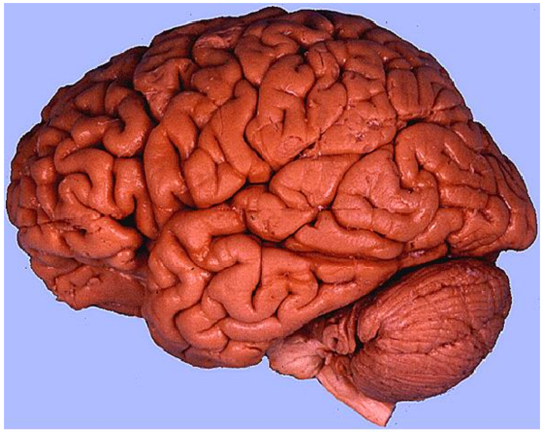
- 100 billion neurons, 32 trillion synapses
- Element size: 10^{-6} m
- Energy use: 25W
- Processing speed: 100 Hz
- Parallel, Distributed
- Fault Tolerant
- Learns: Yes
- Intelligent/Conscious: Usually

- 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

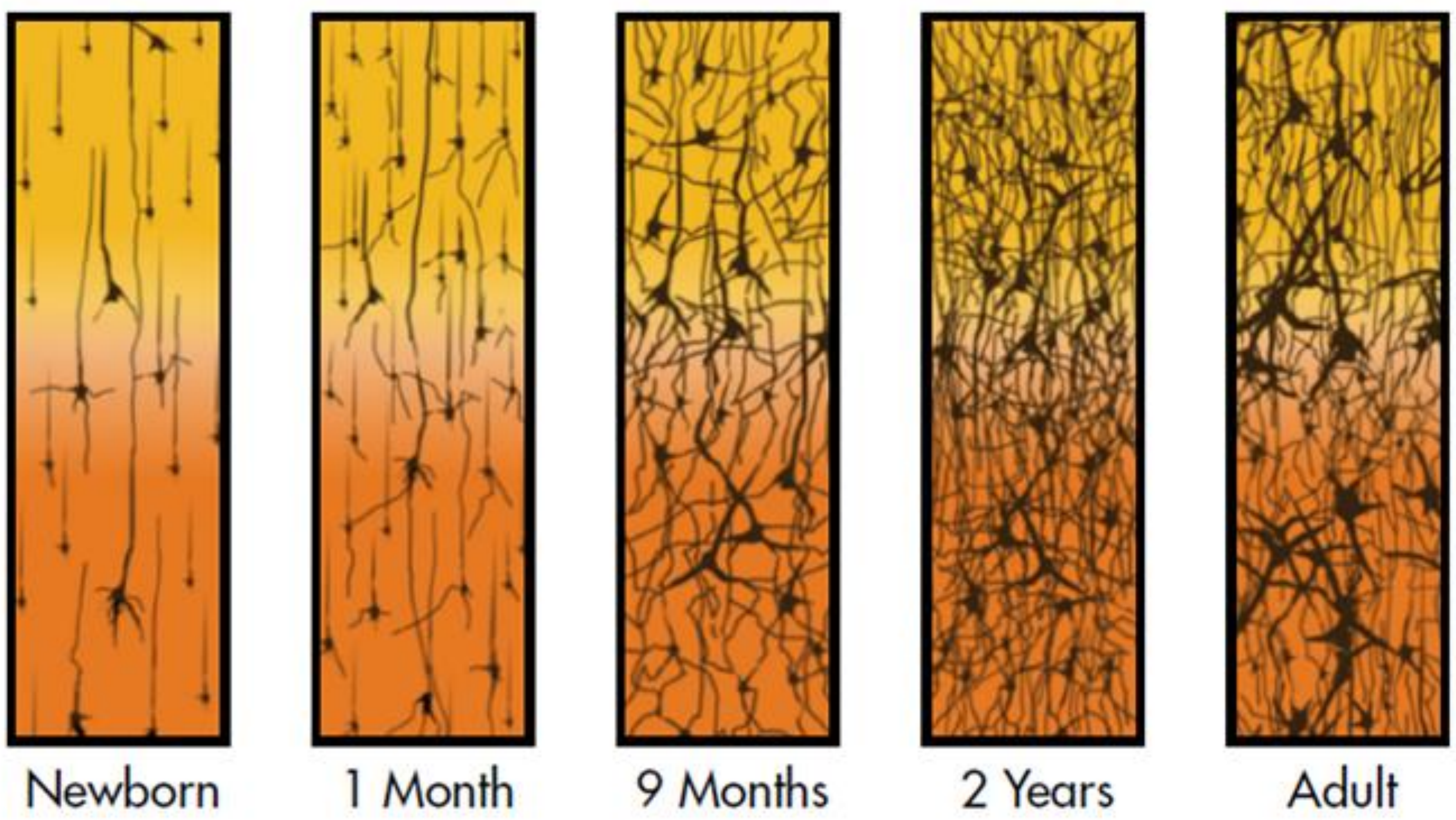
Human brain

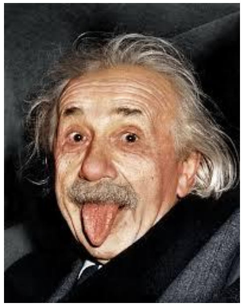


***Biological Psychology 6e*, Figure 2.12 (Part 1)**



Growth of Neural Density in a Human Brain





http://www.urbanchildinstitute.org/why-0-3/baby-and-brain

Biological Neuron

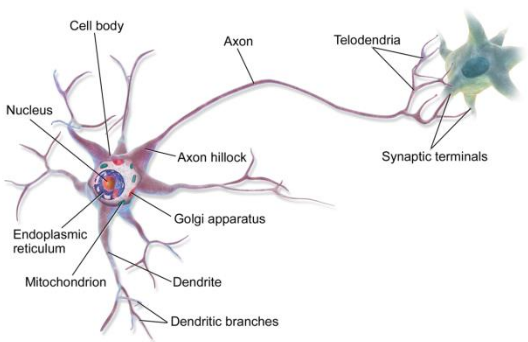
Cell structures

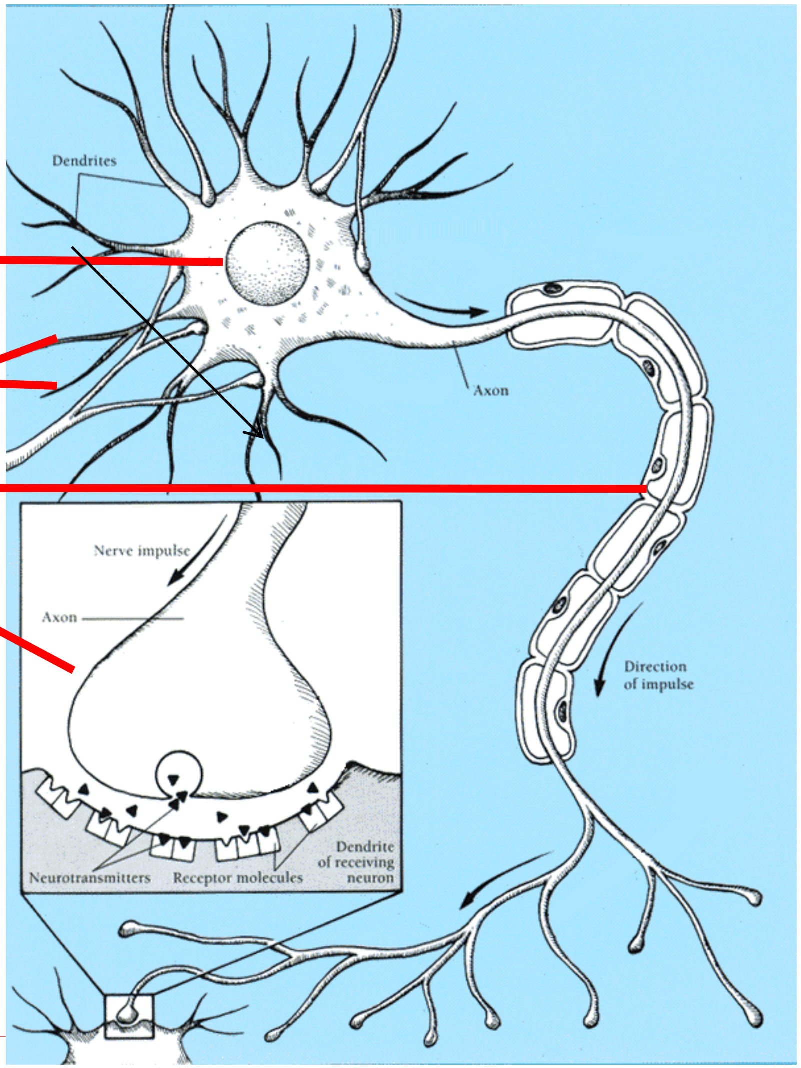
Cell body

Dendrites

Axon

Synaptic terminals



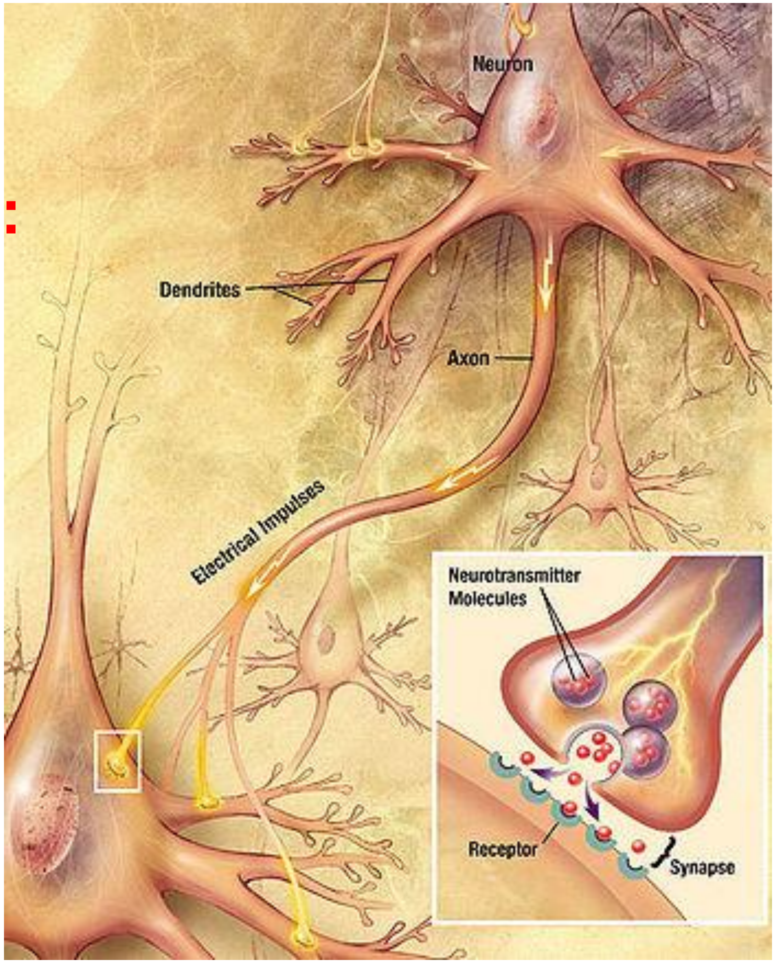


Inter-Neural Communication

Communication is Electrochemical:

Electrical (via ions) along axons

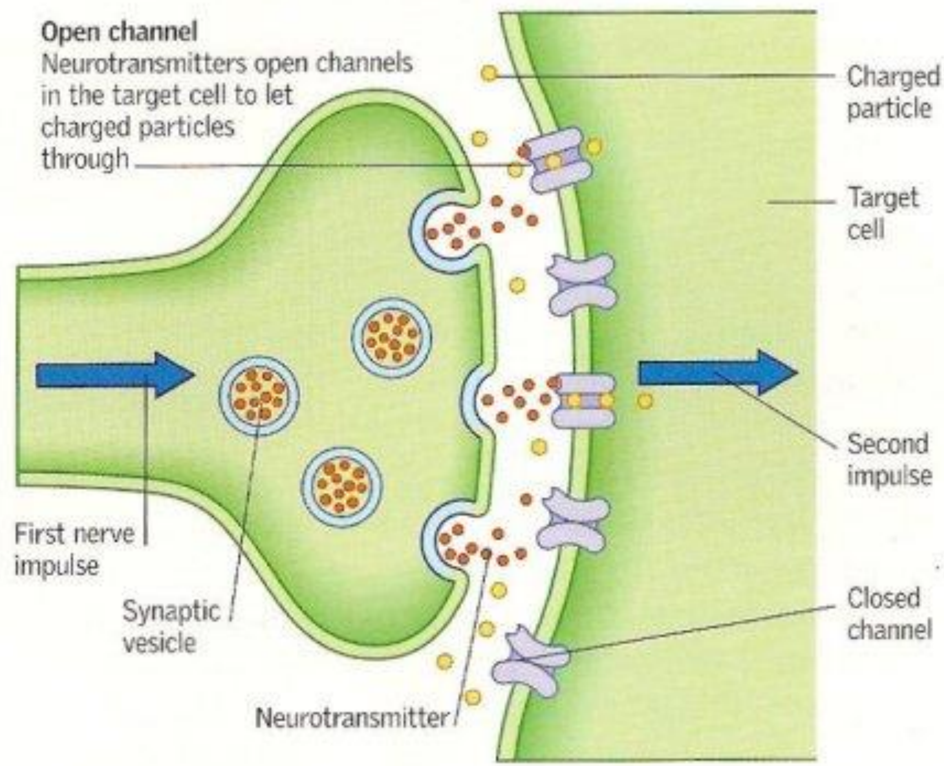
Chemical (via molecules) across synapse



http://en.wikipedia.org/wiki/Neurons

Inter-Neural Communication

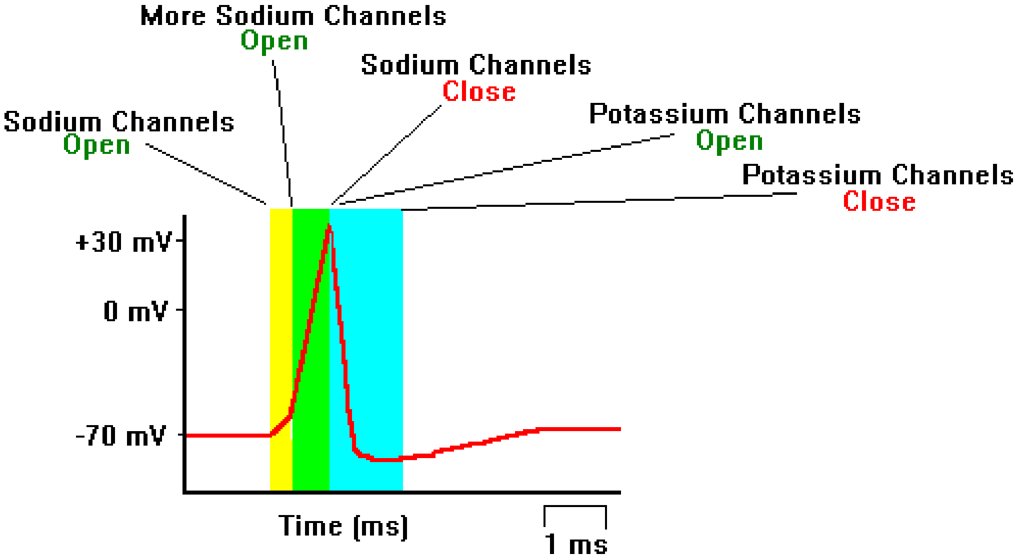
The arrival of a nerve impulse stimulates the release of neurotransmitters from vesicles. They pass across the synapse and open channels in the target cell.
Charged particles can then enter and trigger a second impulse.



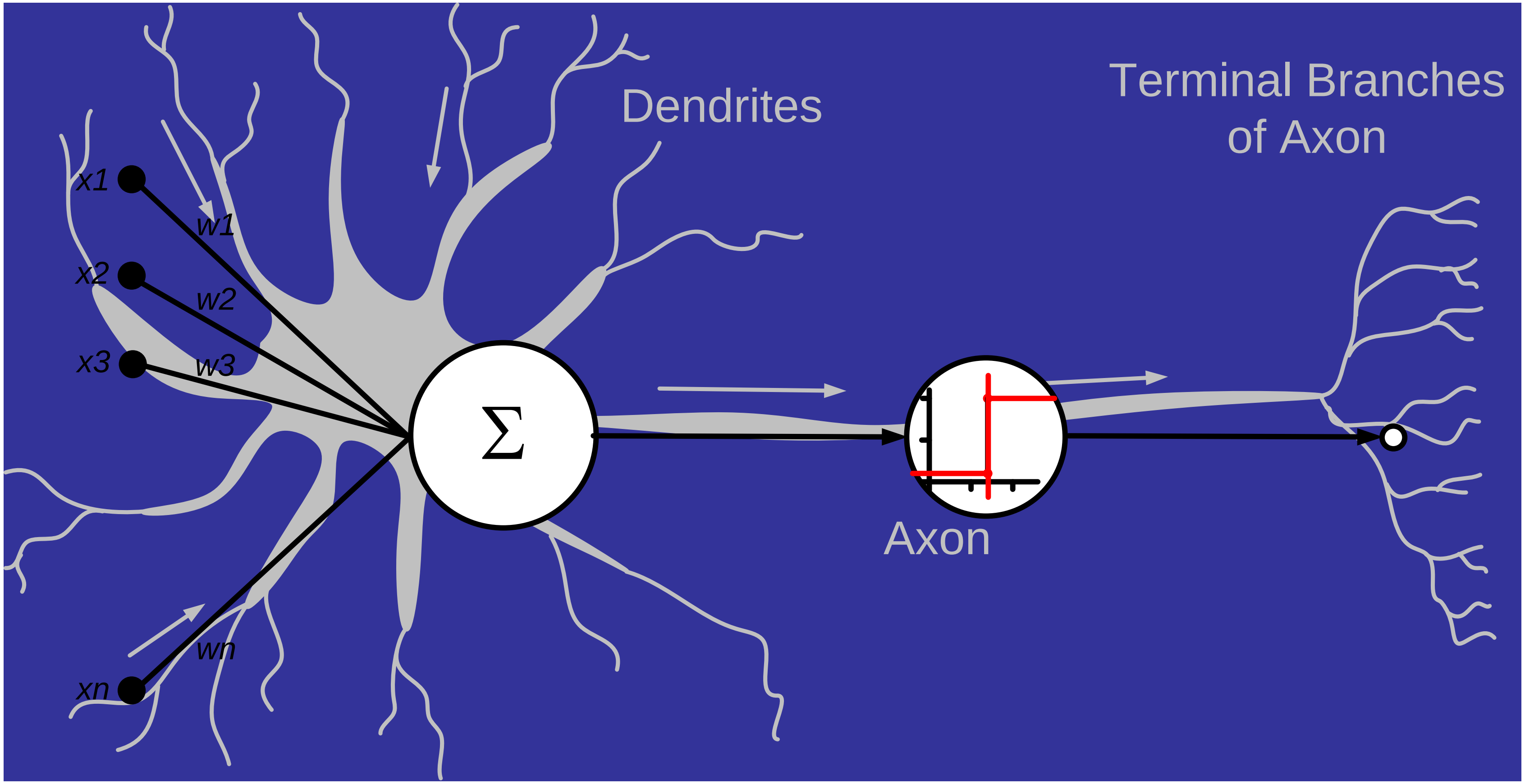
http://www.daviddarling.info/encyclopedia/N/neurotransmitter.html

Action potential

A neuron receives input from other neurons (typically many thousands). Inputs sum (approximately). Once input exceeds a critical level, the neuron discharges a spike - an electrical pulse that travels from the body, down the axon, to the next neuron(s) (or other receptors)

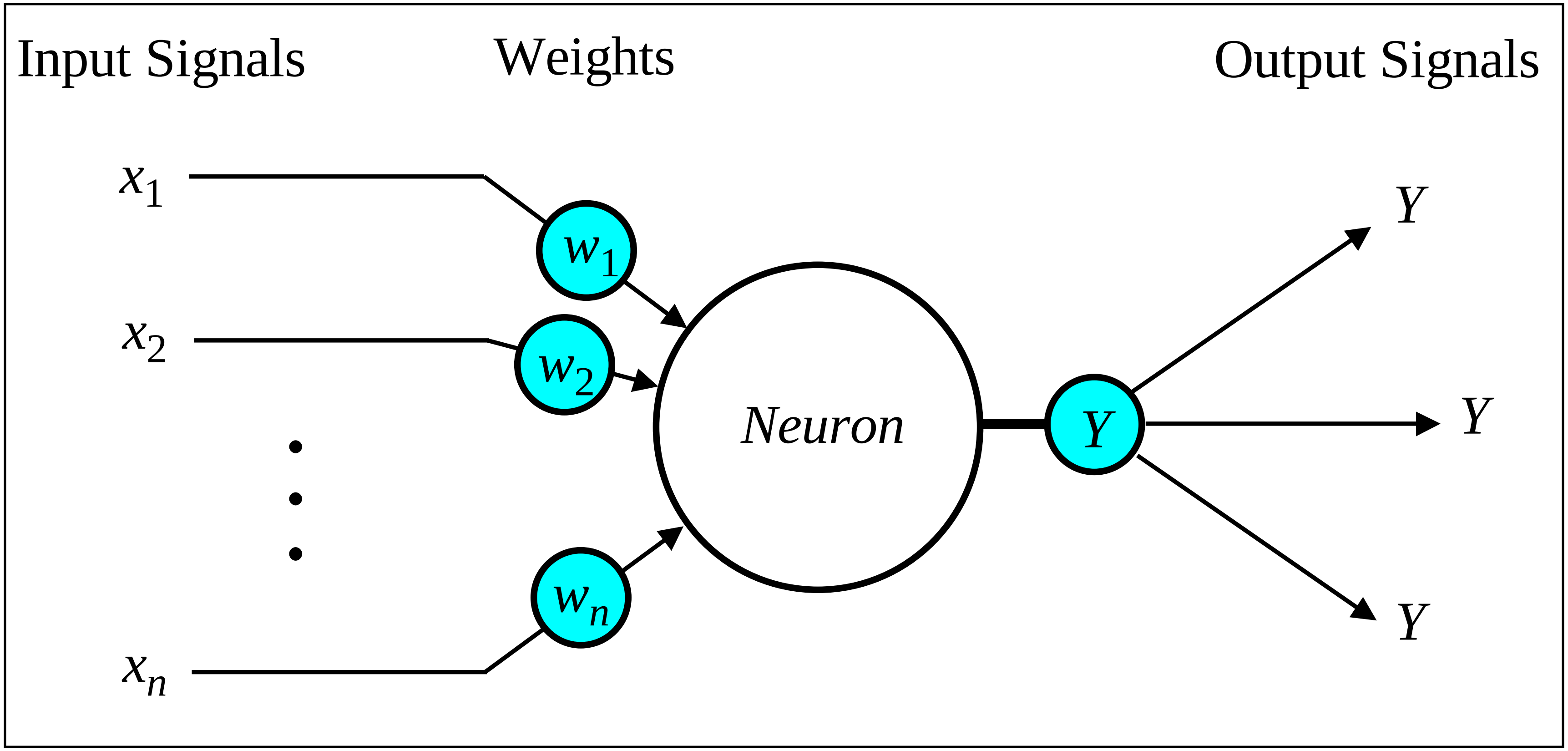


Models of A Neuron



The neuron as a simple computing element

Diagram of a neuron

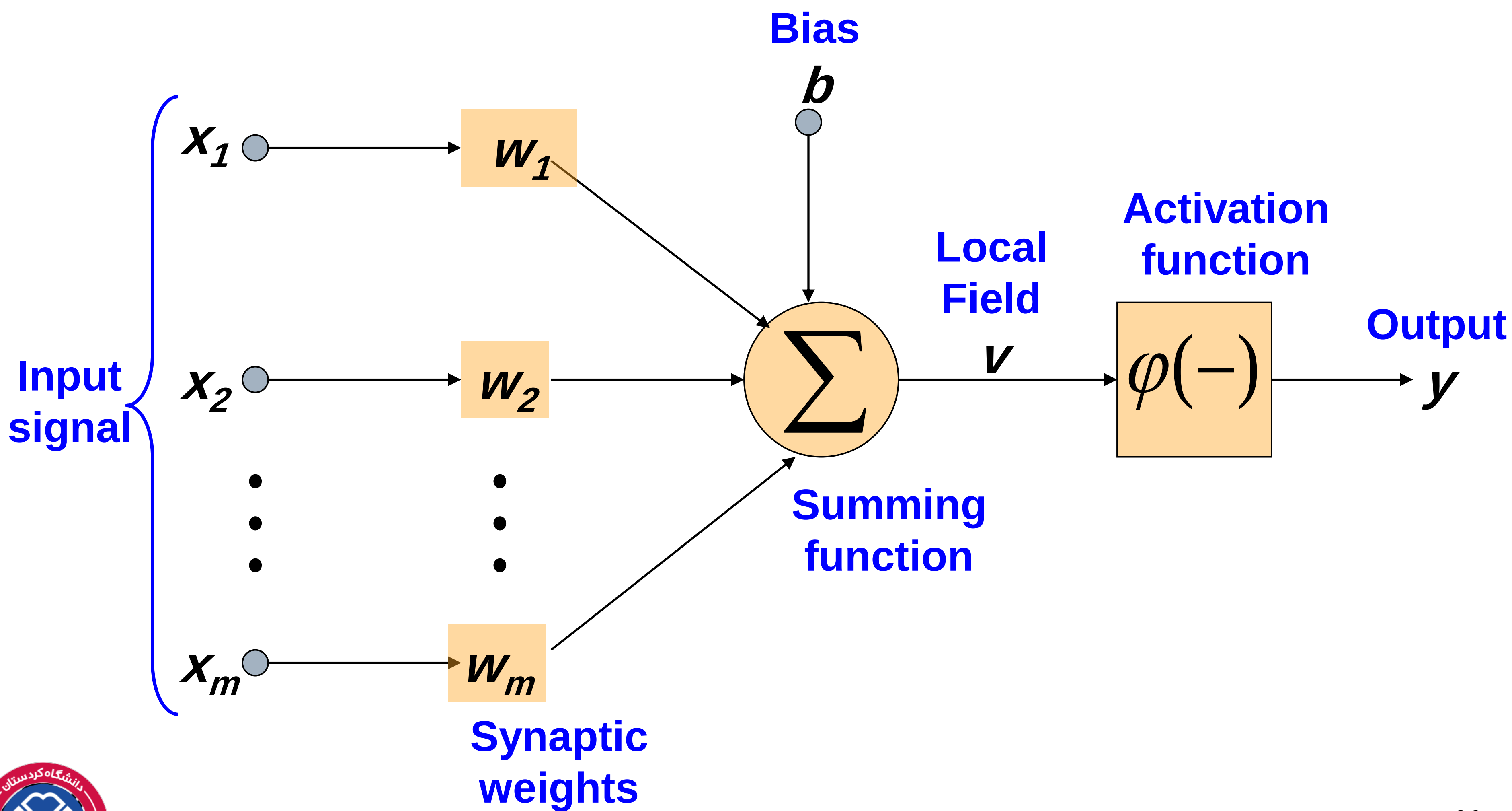


Models of A Neuron

Three elements:

1. **A set of synapses, or connection link:** each of which is characterized by a ***weight*** or strength of its own w_{kj}. Specifically, a signal x_j at the input synapse 'j' connected to neuron 'k' is multiplied by the synaptic w_{kj}
2. **An adder:** For summing the input signals, weighted by respective synaptic strengths of the neuron in a linear operation.
3. **Activation function:** For limiting of the amplitude of the output of the neuron to limited range. The activation function is referred to as a Squashing (i.e. limiting) function {interval [0,1], or, alternatively [-1,1]}

The Neuron

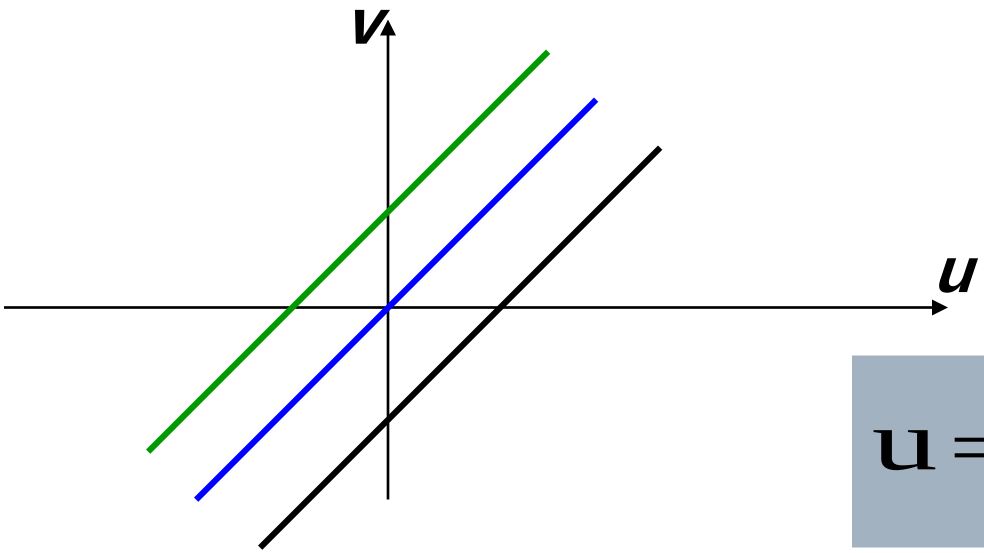

Bias
b
x1
x2
xm
w1
w2
wm
Input
signal
Local
Field
v
Activation
function
φ(−)
Output
y
Summing
function
Synaptic
weights

Bias of a Neuron

- Bias ***b*** has the effect of applying an affine transformation to ***u***

$$v = u + b$$

- ***v*** is the **induced field** of the neuron

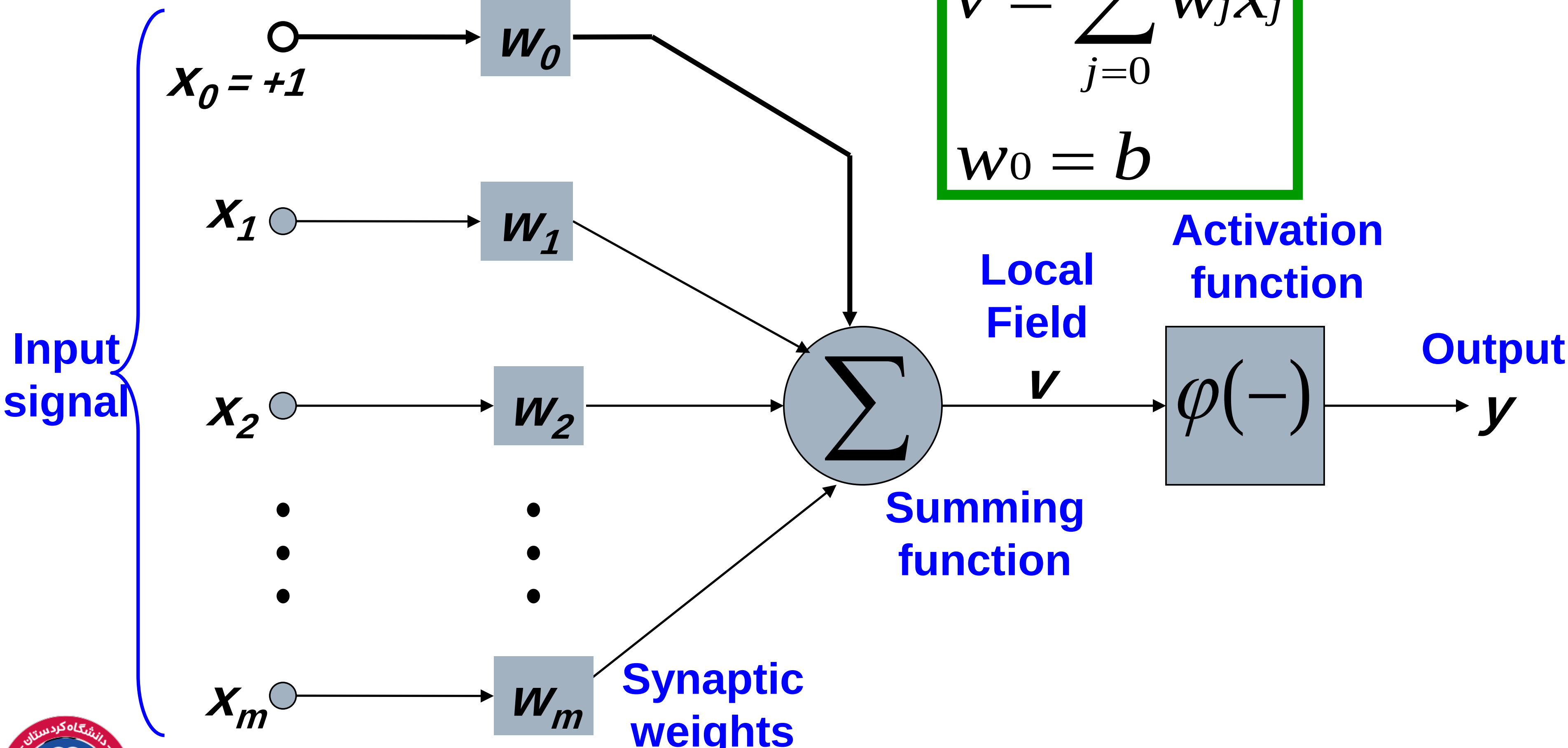


$$u = \sum_{j=1}^{m} w_j x_j$$

Bias as extra input

$$v = \sum_{j=0}^{m} w_j x_j$$

$$w_0 = b$$



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) φ for limiting the amplitude of the output of the neuron.

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

Dimensions of a Neural Network

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

Network architectures

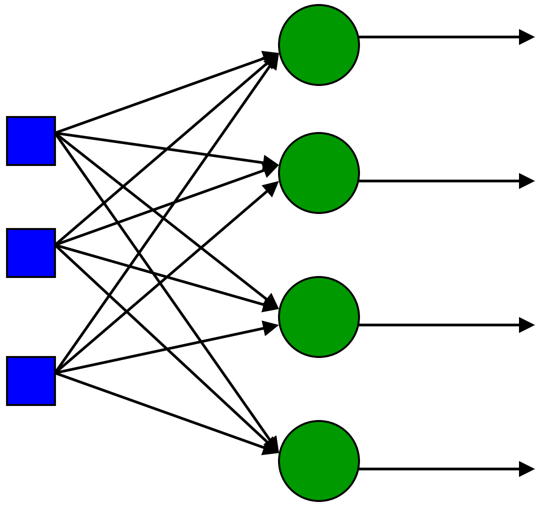
- Three different classes of network architectures

  - single-layer feed-forward } neurons are organized
  - multi-layer feed-forward } in acyclic layers
  - recurrent

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

Single Layer Feed-forward

Input layer of source nodes

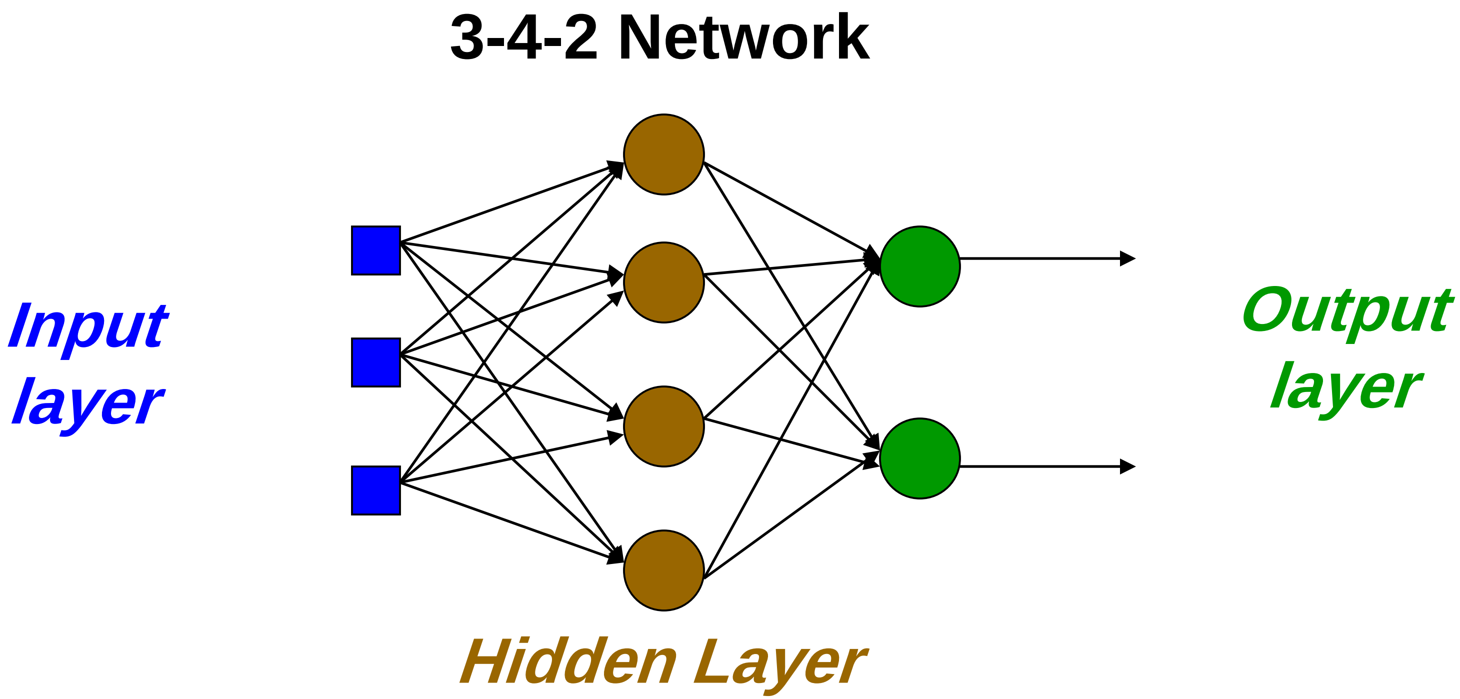


Output layer of neurons

- Input layer of source nodes that projects directly onto an output layer of neurons.

- “Single-layer” referring to the output layer of computation nodes (neuron).

Multi layer feed-forward

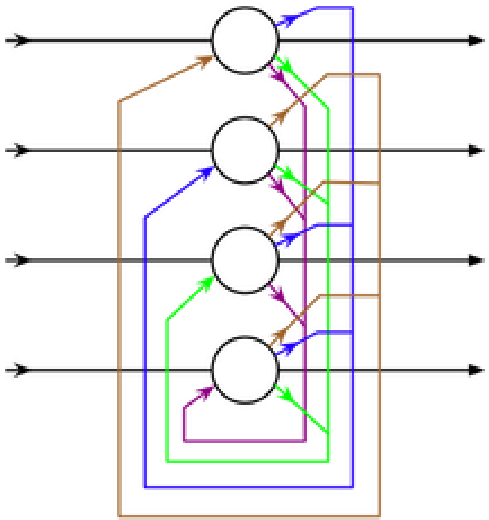


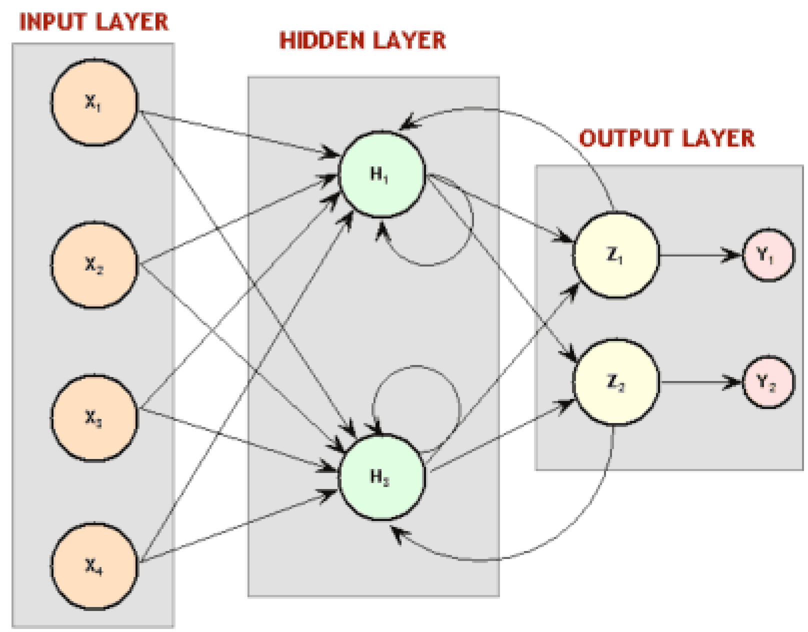
It contains one or more hidden layers (hidden neurons).

The function of hidden neuron is to intervene between input and output.

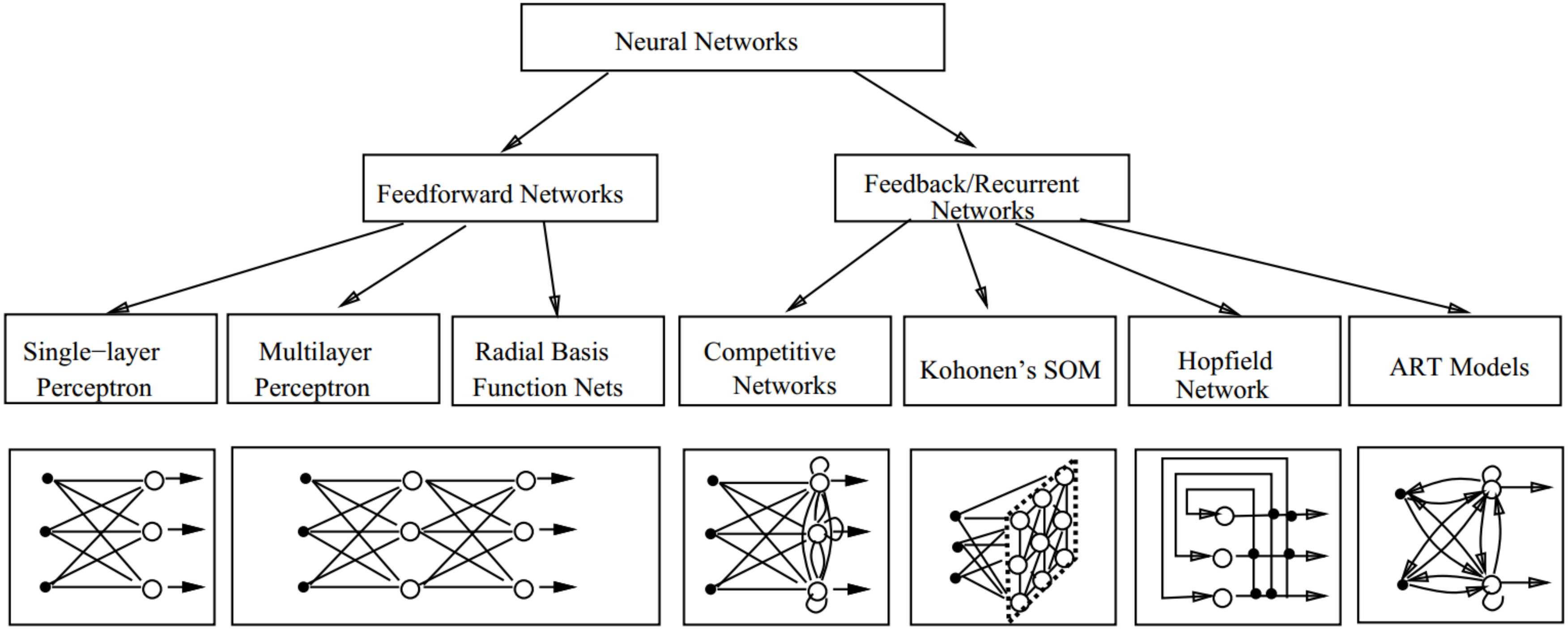
By adding one or more hidden layers, the network is able to extract higher-order statistics from input

Recurrent network





- It is different from feed forward neural network in that it has at least one feedback loop.
- Recurrent network may consist of single layer of neuron with each neuron feeding its output signal back to the inputs of all the other neurons. Note: There are no self-feedback.
- Feedback loops have a profound impact on learning and overall performance.


Neural Networks
Feedforward Networks
Feedback/Recurrent Networks
Single−layer Perceptron
Multilayer Perceptron
Radial Basis Function Nets
Competitive Networks
Kohonen's SOM
Hopfield Network
ART Models

Learning

- Supervised Learning
  - Recognizing hand-written digits, Classification, 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, clustering,
  - Unlabeled examples
    (different realizations of the input alone)
  - Neural Network models: self organizing maps, Hopfield networks.

Learning Paradigms

Supervised Learning

- Multilayer perceptrons
- Radial basis function networks
- Modular neural networks
- LVQ (learning vector quantization)

Unsupervised Learning

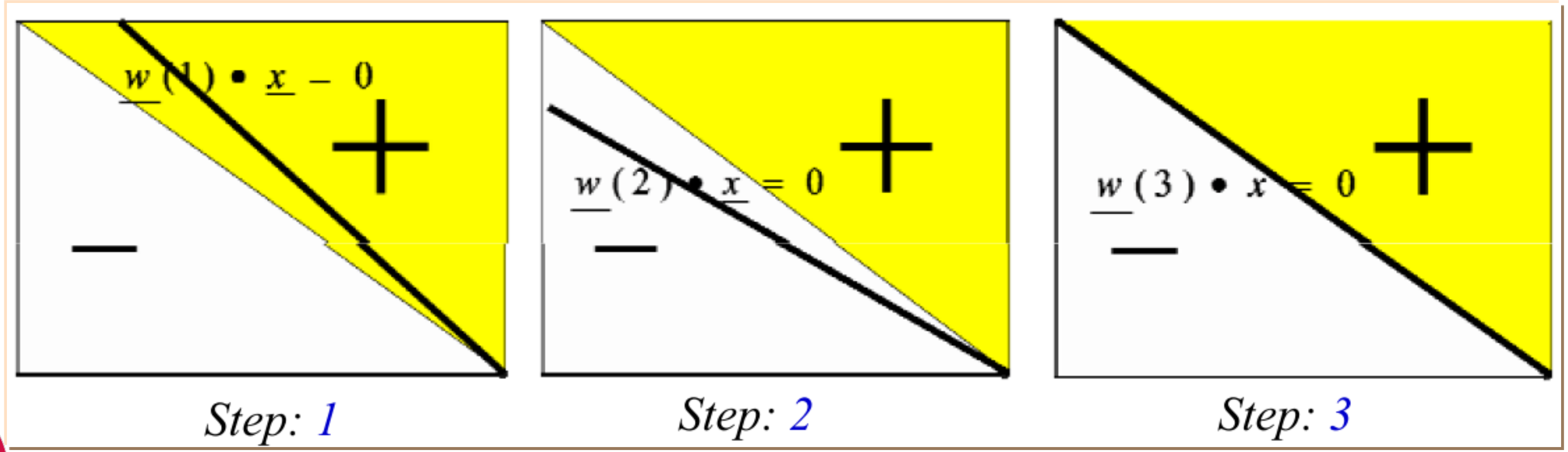
- Competitive learning networks
- Kohonen self-organizing networks
- ART (adaptive resonant theory)

Others

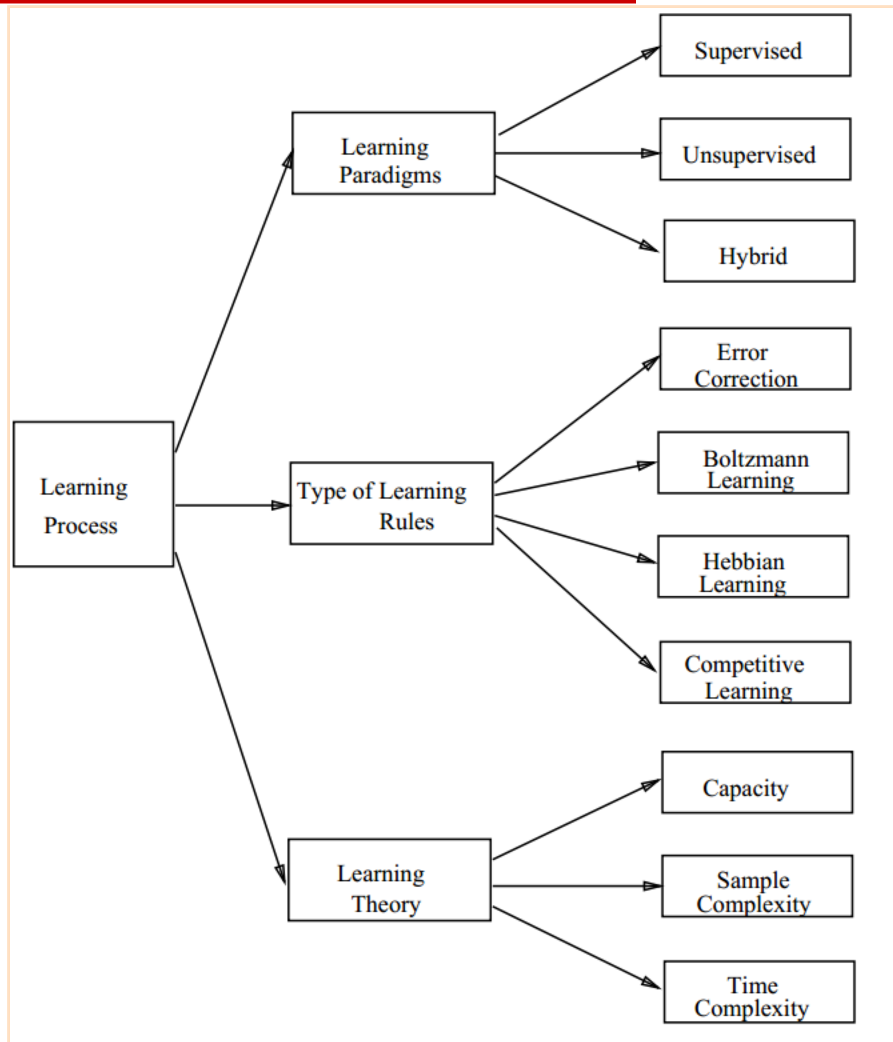
- Autoassociative memories (Hopfield networks)

Learning Algorithm

What is the goal of learning algorithm?
We need a learning algorithm which it updates the weights wi (w) so that finally (at end of learning process) the input patterns lie on both sides of the line decided by the Perceptron.



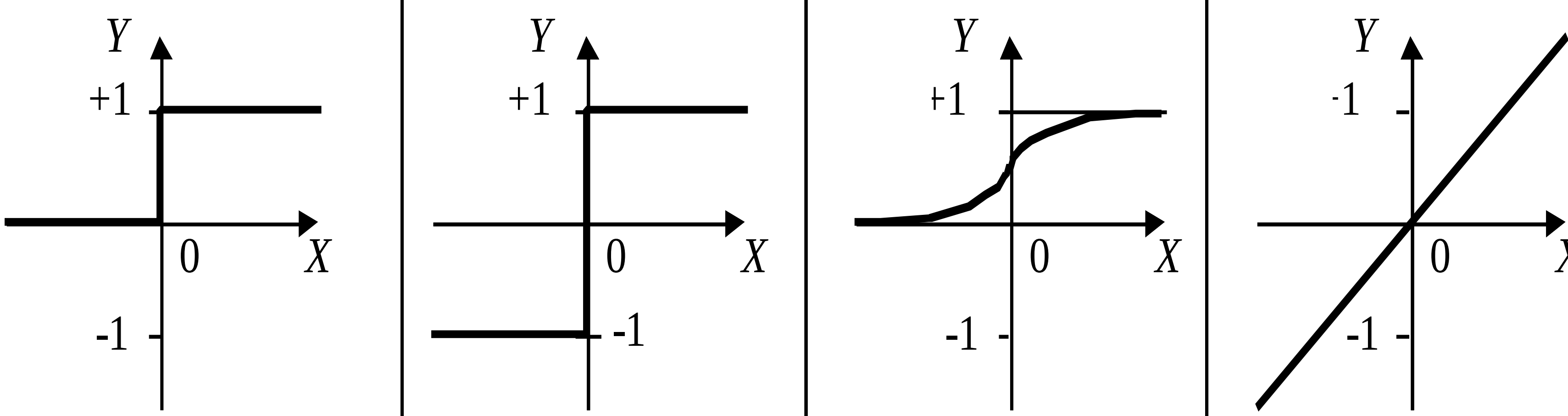
Learning



Learning Process
Learning Paradigms
Supervised
Unsupervised
Hybrid
Type of Learning Rules
Error Correction
Boltzmann Learning
Hebbian Learning
Competitive Learning
Learning Theory
Capacity
Sample Complexity
Time Complexity

Activation functions of a neuron

Step function	*Sign function*	*Sigmoid function*	*Linear function*



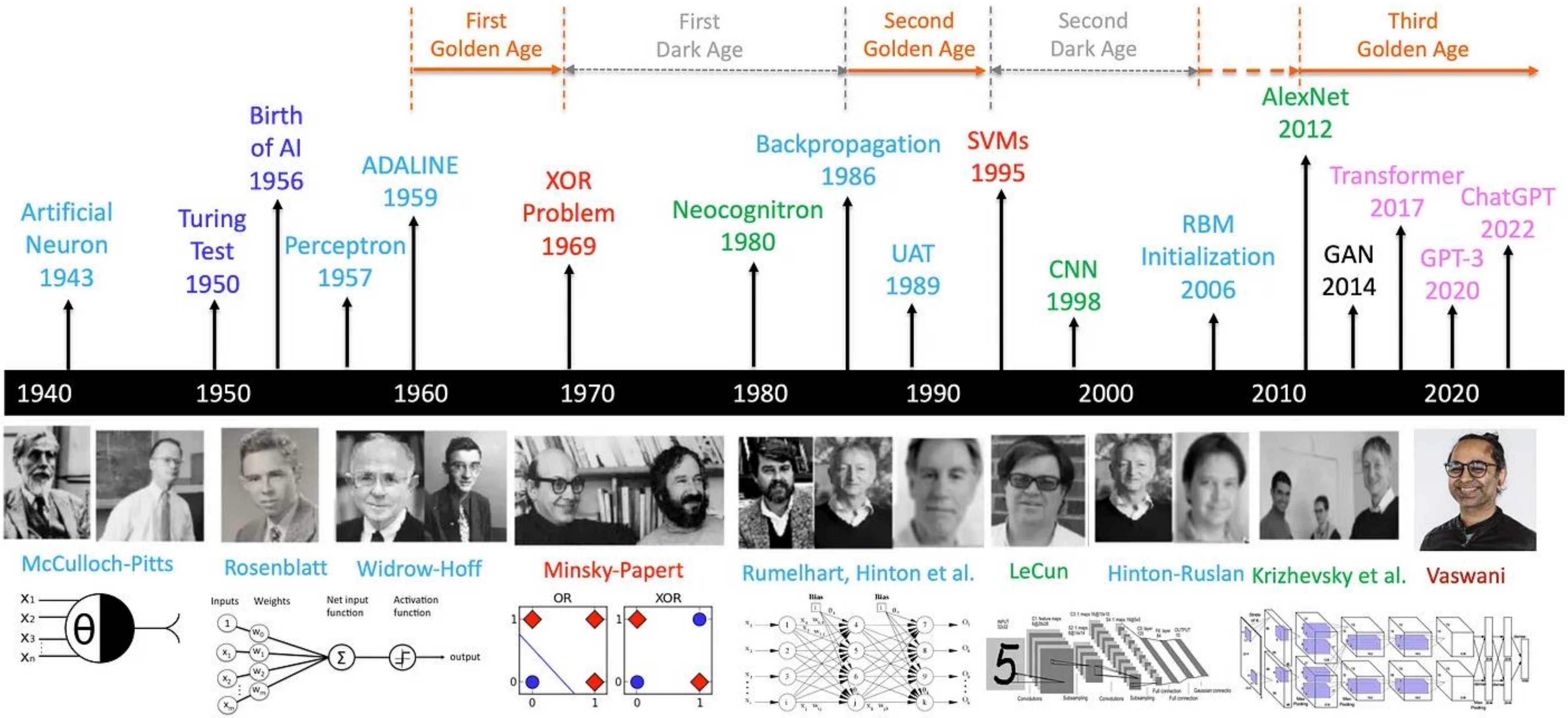
$Y^{step} = \begin{cases} 1, \text{ if } X \geq 0 \\ 0, \text{ if } X < 0 \end{cases}$	$Y^{sign} = \begin{cases} +1, \text{ if } X \geq 0 \\ -1, \text{ if } X < 0 \end{cases}$	$Y^{sigmoid} = \dfrac{1}{1 + e^{-X}}$	$Y^{linear} = X$

History

A Brief History

- **1943** McCulloch and Pitts proposed the McCulloch-Pitts neuron model
- **1949** Hebb published his book *The Organization of Behavior*, in which the Hebbian learning rule was proposed.
- **1958** Rosenblatt introduced the simple single layer networks now called Perceptrons.
- **1969** Minsky and Papert's book *Perceptrons* demonstrated the limitation of single layer perceptrons, and almost the whole field went into hibernation.
- **1982** Hopfield published a series of papers on Hopfield networks.
- **1982** Kohonen developed the Self-Organizing Maps that now bear his name.
- **1986** The Back-Propagation learning algorithm for Multi-Layer Perceptrons was re-discovered and the whole field took off again.
- **1990s** The sub-field of Radial Basis Function Networks was developed.
- **2000s** The power of Ensembles of Neural Networks and Support Vector Machines becomes apparent.

A Brief History



A Brief History

1943 McCulloch and Pitts proposed the McCulloch-Pitts neuron model

Warren S. McCulloch
(Nov., 16, 1898 – Sep., 24, 1969)

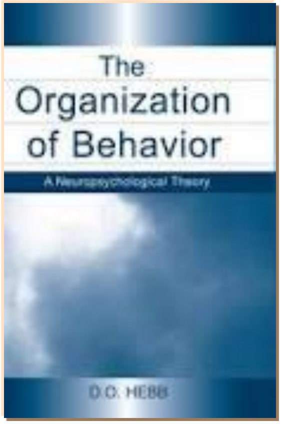
Walter Pitts (1923-1969)

W. McCulloch and W. Pitts, 1943 "A Logical Calculus of the Ideas Immanent in Nervous Activity". In :Bulletin of Mathematical Biophysics Vol 5, pp 115-133 .

A Brief History

1949 Hebb published his book *The Organization of Behavior*, in which the Hebbian learning rule was proposed.

Donald Olding Hebb
(July 22, 1904 – August 20, 1985)



cells that fire together, wire together

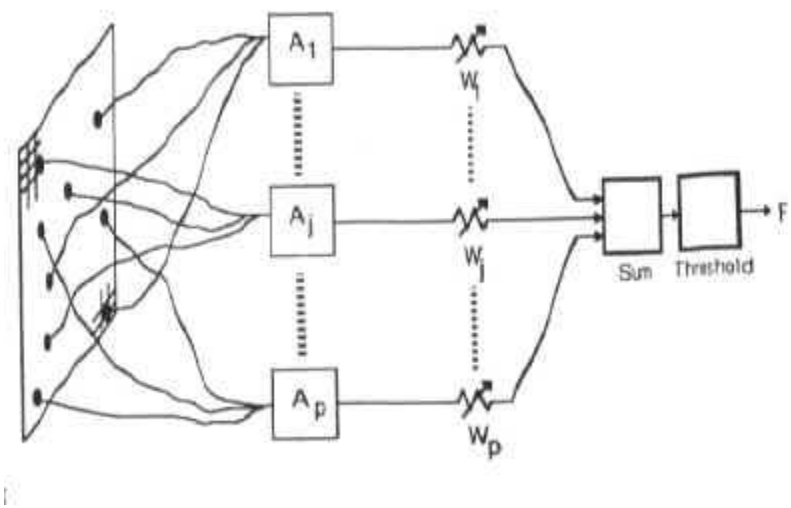
A Brief History

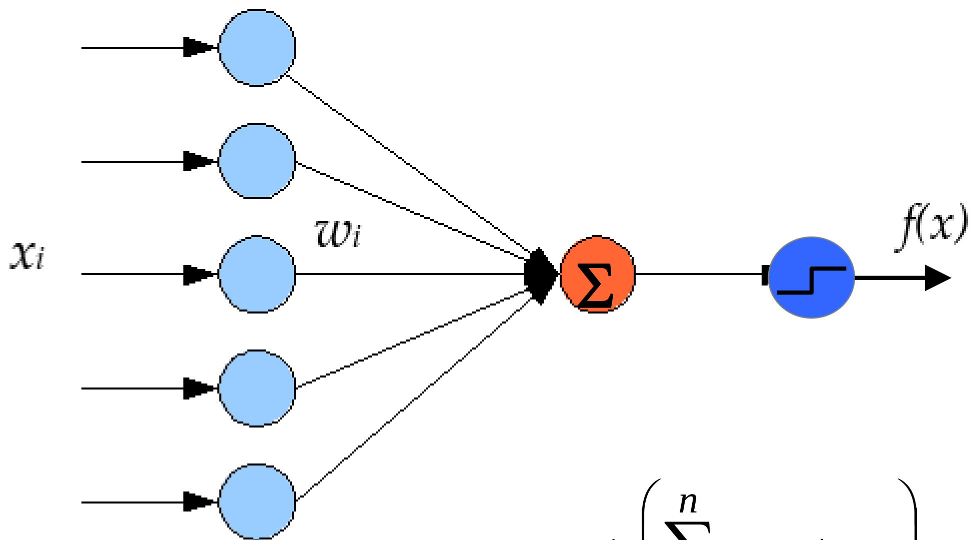
1958 Rosenblatt introduced the simple single layer networks now called Perceptrons.

Frank Rosenblatt
(11 July 1928 – 1971)

A Brief History

1958 Rosenblatt introduced the simple single layer networks now called Perceptrons.

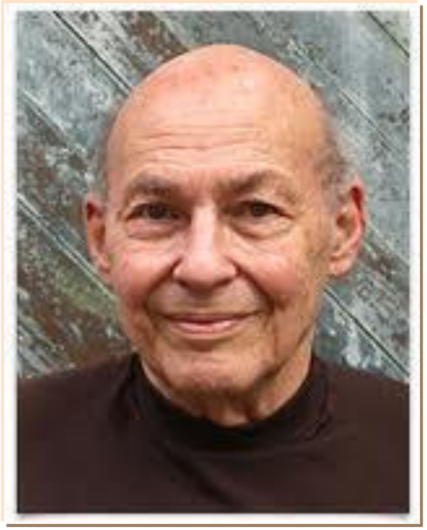


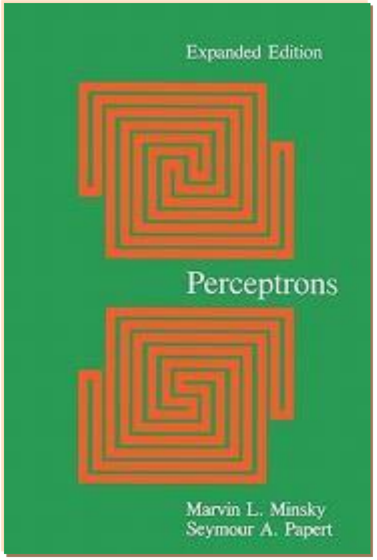


$$y = \operatorname{sgn}\left(\sum_{i=1}^{n} w_i x_i + w_0\right)$$

A Brief History

1969 Minsky and Papert's book Perceptrons demonstrated the limitation of single layer perceptrons, and almost the whole field went into **hibernation**.



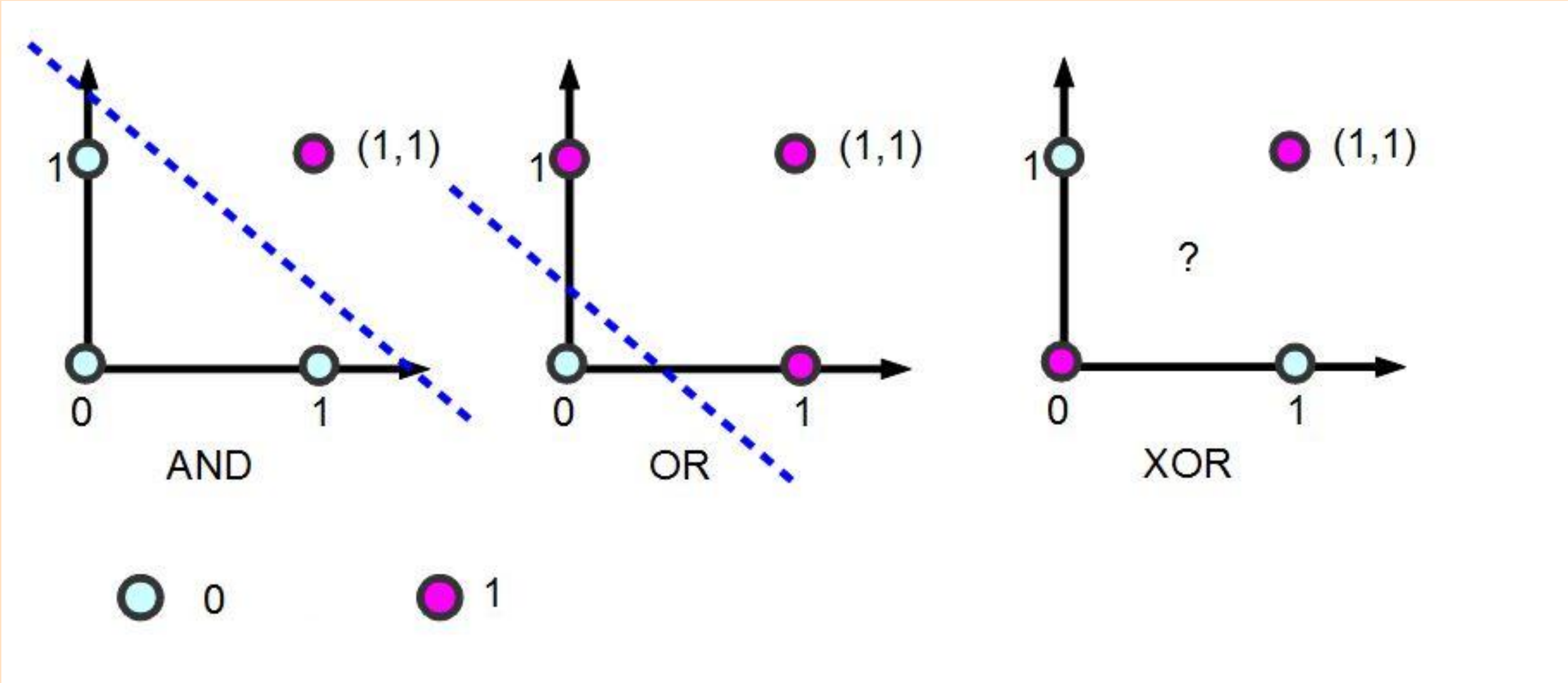


Marvin Minsky (born August 9, 1927)

Seymour Papert (born February 29, 1928)

A Brief History

limitation of single layer perceptrons



A typical example of non-linealy separable function is the XOR

History- The setback (mid 60's – late 70's)

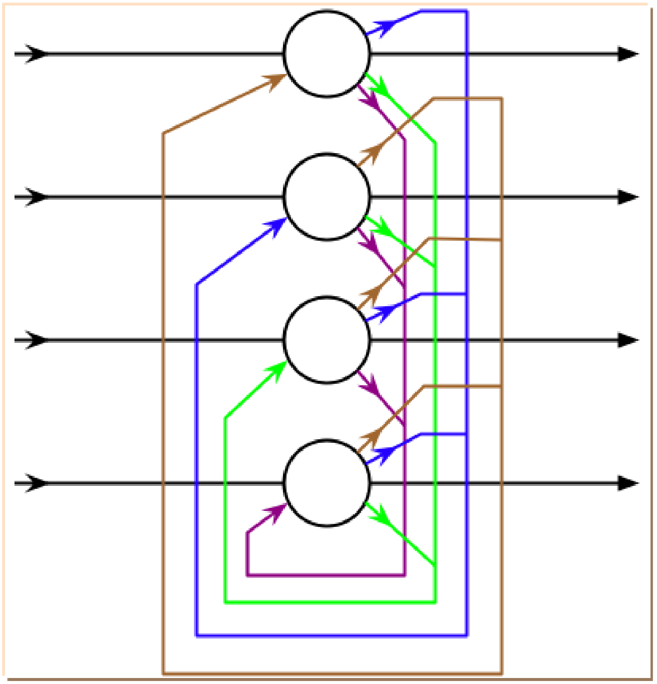
Serious problems with perceptron model (Minsky's book 1969)

- Single layer perceptron cannot represent (learn) simple functions such as **XOR**
- Multi-layer of non-linear units may have greater power but there is no learning rule for such nets
- Scaling problem: connection weights may grow infinitely

The first two problems overcame by latter effort in 80's, but the scaling problem persists

A Brief History

1982 Hopfield published a series of papers on Hopfield networks.

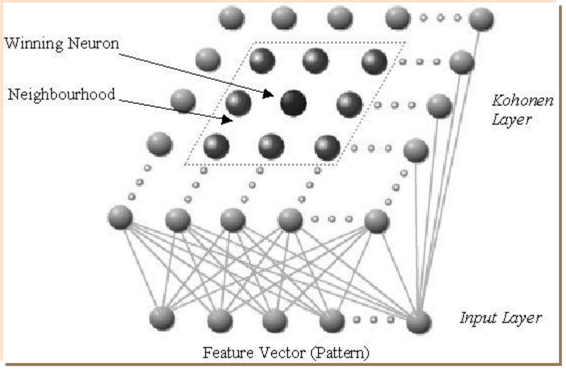


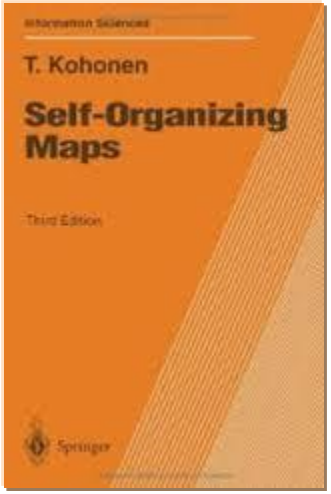
John Joseph Hopfield (born July 15, 1933)

A Brief History

1982 Kohonen developed the Self-Organizing Maps that now bear his name.

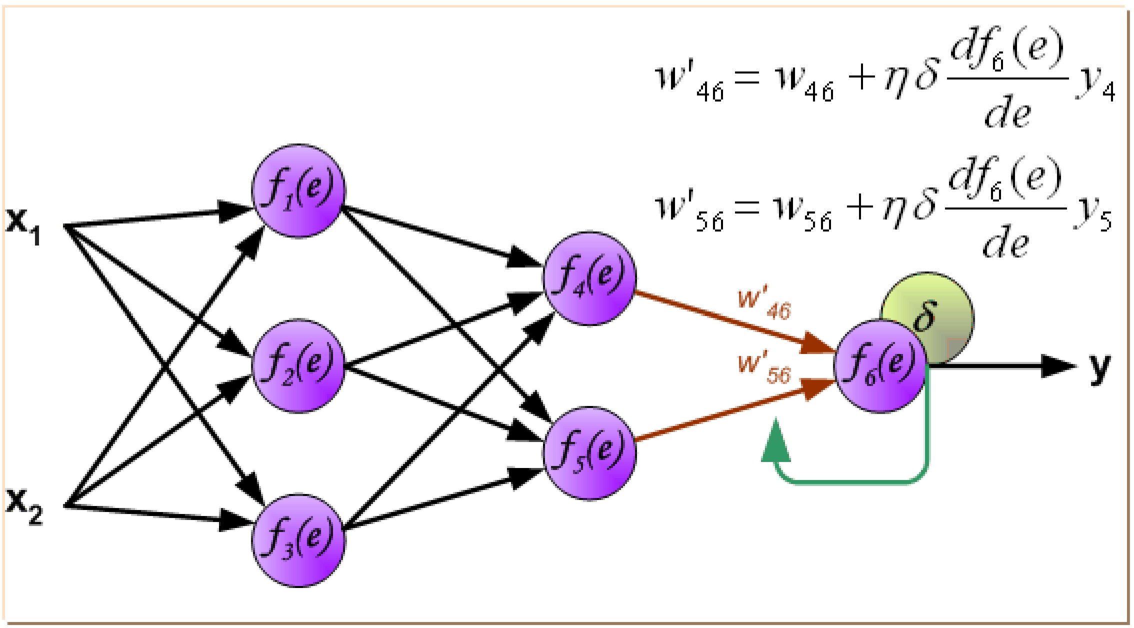
Teuvo Kohonen
(born July 11, 1934)





A Brief History

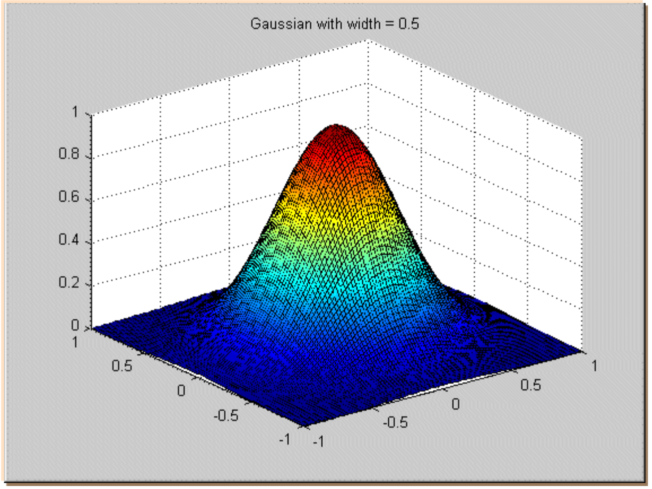
1986 The Back-Propagation learning algorithm for Multi-Layer Perceptrons was rediscovered and the whole field took off again.

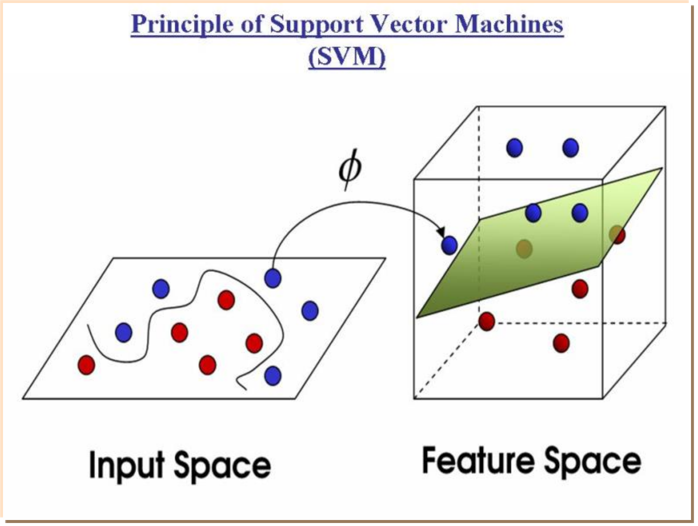


A Brief History

1990s The sub-field of Radial Basis Function Networks was developed.

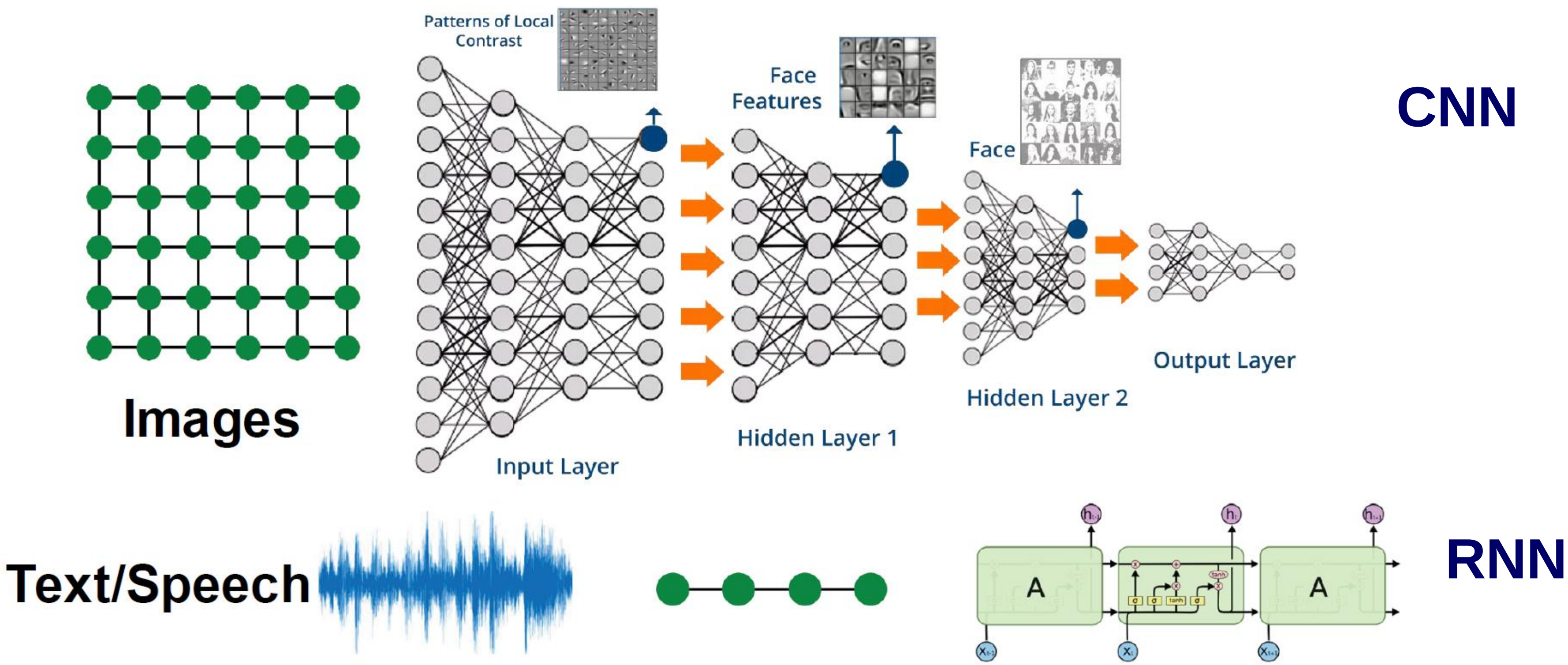
2000s The power of Ensembles of Neural Networks and Support Vector Machines becomes apparent.





A Brief History

2009- now deep learning

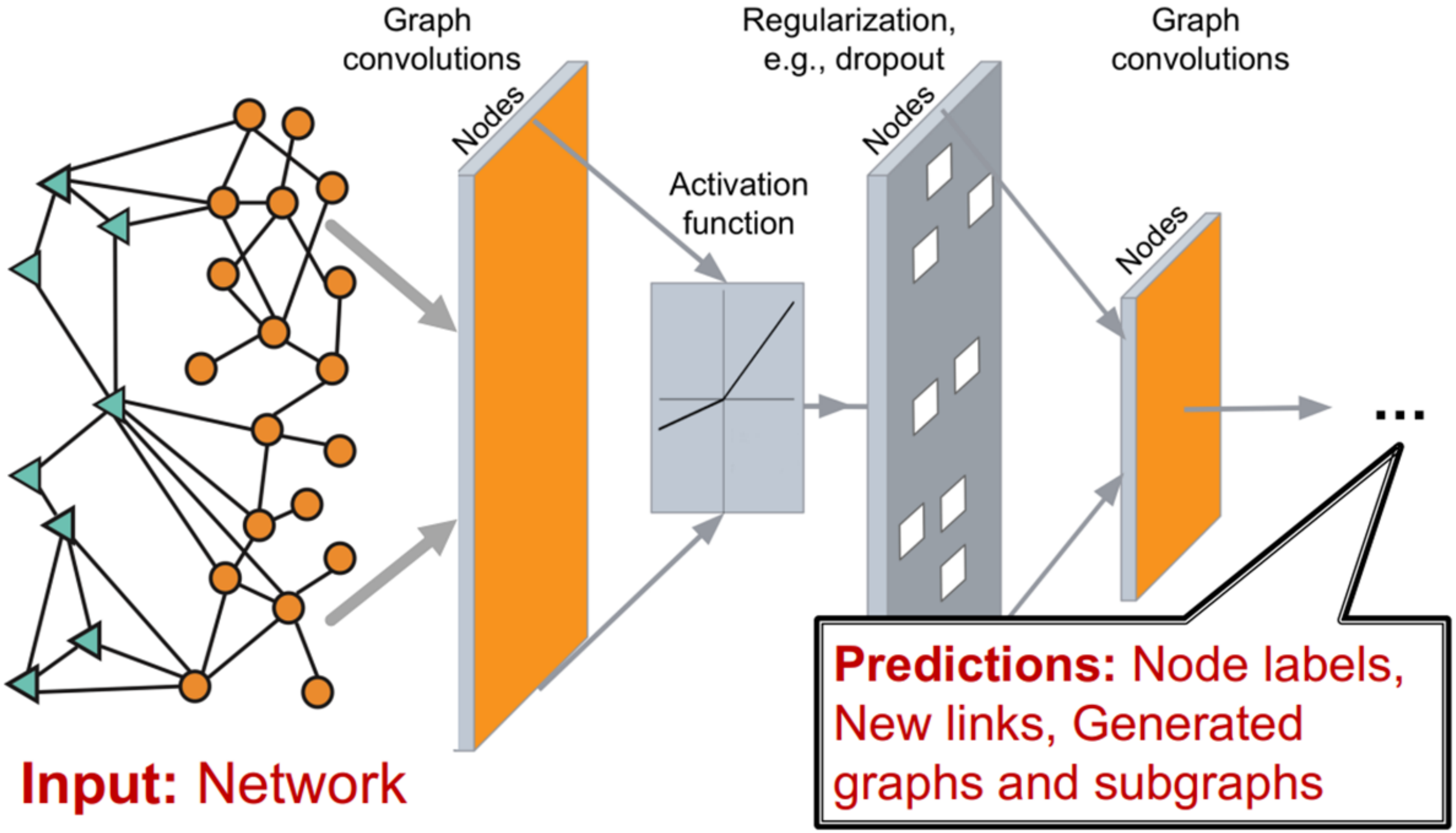


A Brief History

- 2012: AlexNet by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton
- 2014: Generative adversarial network (**GAN**) by (Ian Goodfellow et al.)
- 2017: Transformer
- 2020: GPT-3
- 2022: ChatGPT

A Brief History

Graph Neural Networks

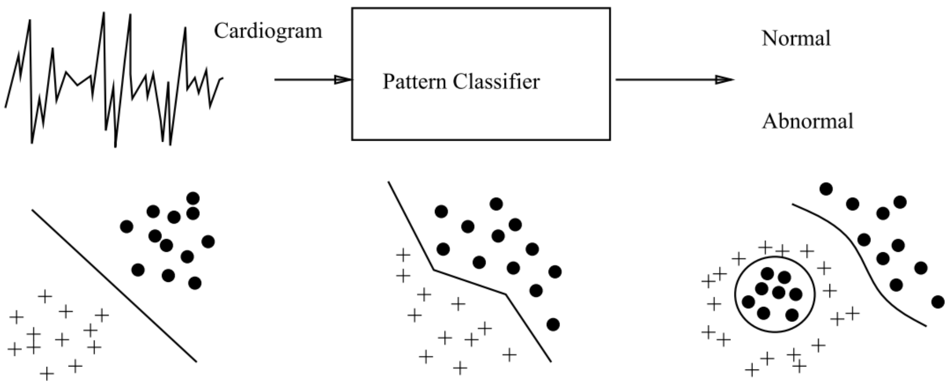


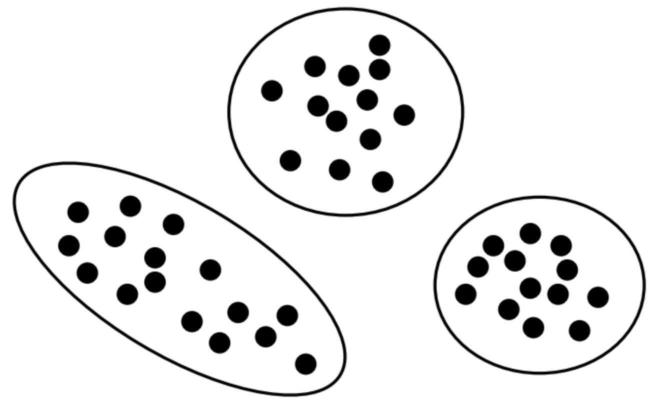
Applications

What can a ANN do?

- Pattern Classification
- Clustering
- Compute a known function
- Approximate an unknown function
- Pattern Recognition
- Signal Processing
- Content addressed memory
- Learn to do any of the above

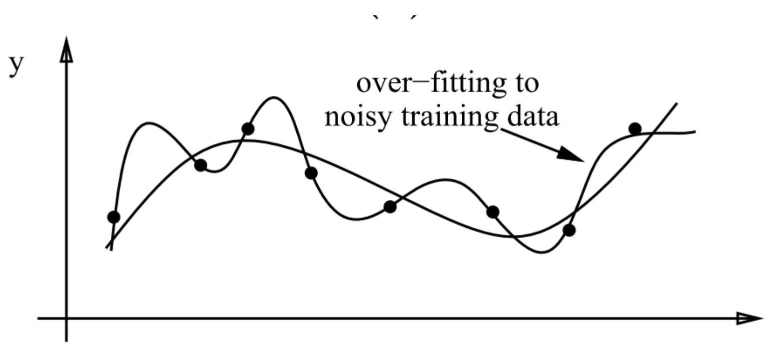
What can a ANN do?

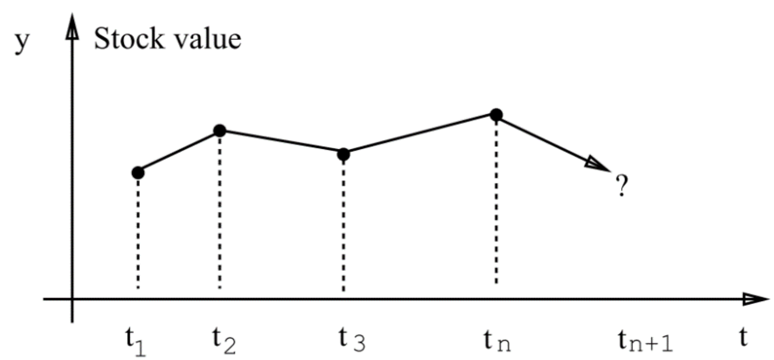




Pattern classification

clustering

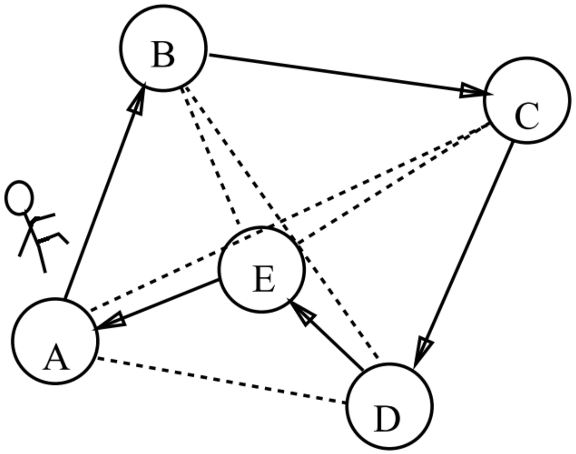




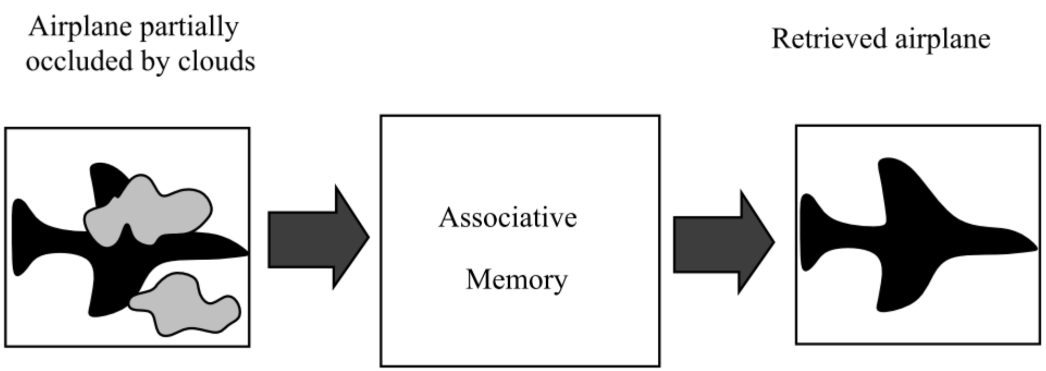
Function approximation

Prediction/forecasting

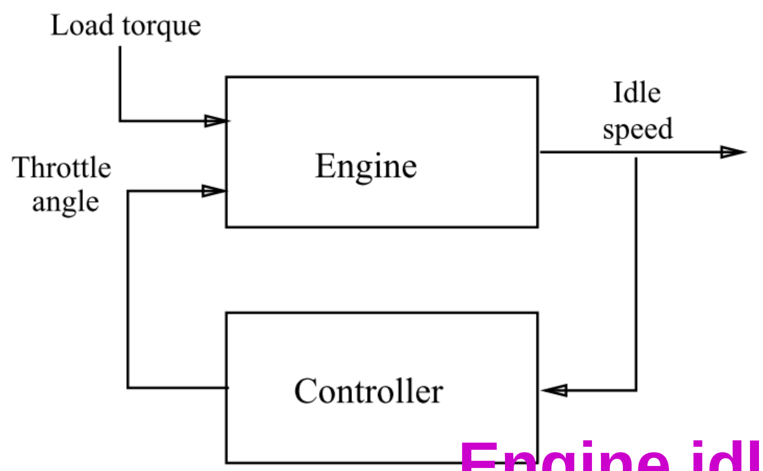
What can a ANN do?



Optimization
(TSP problem)

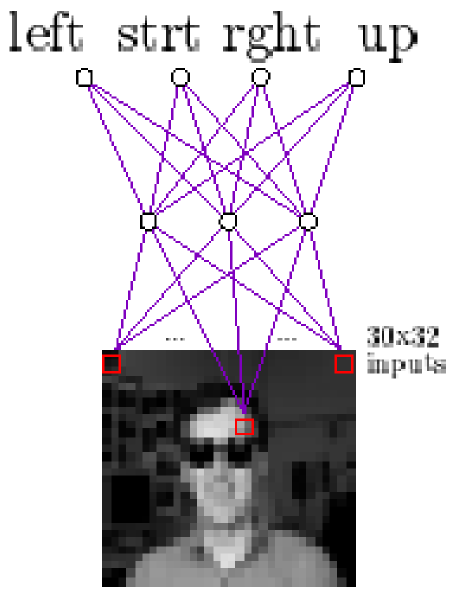


Retrieval by content



Engine idle speed control

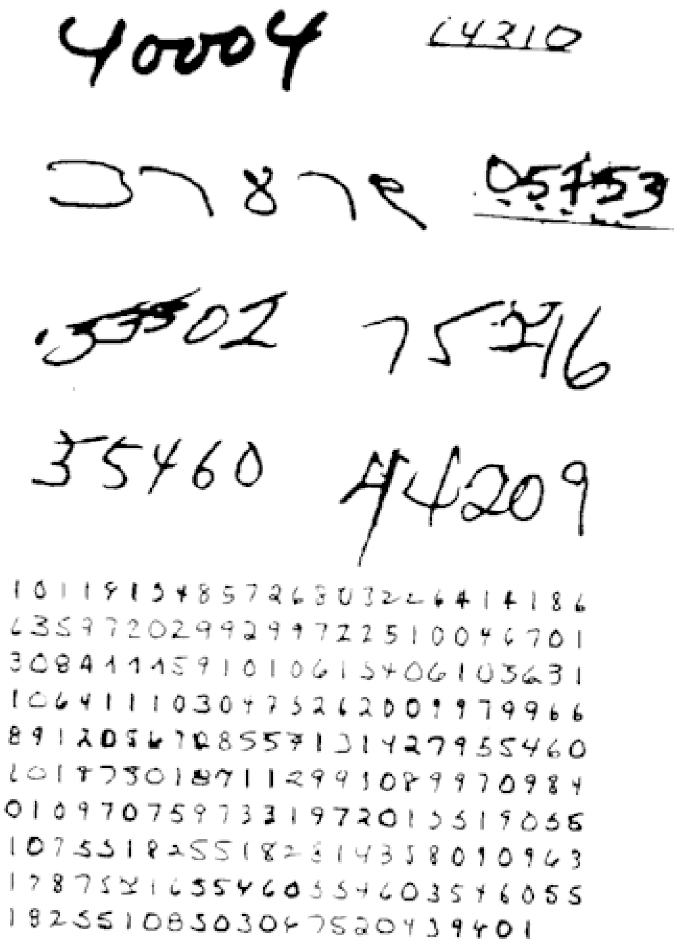
Application examples



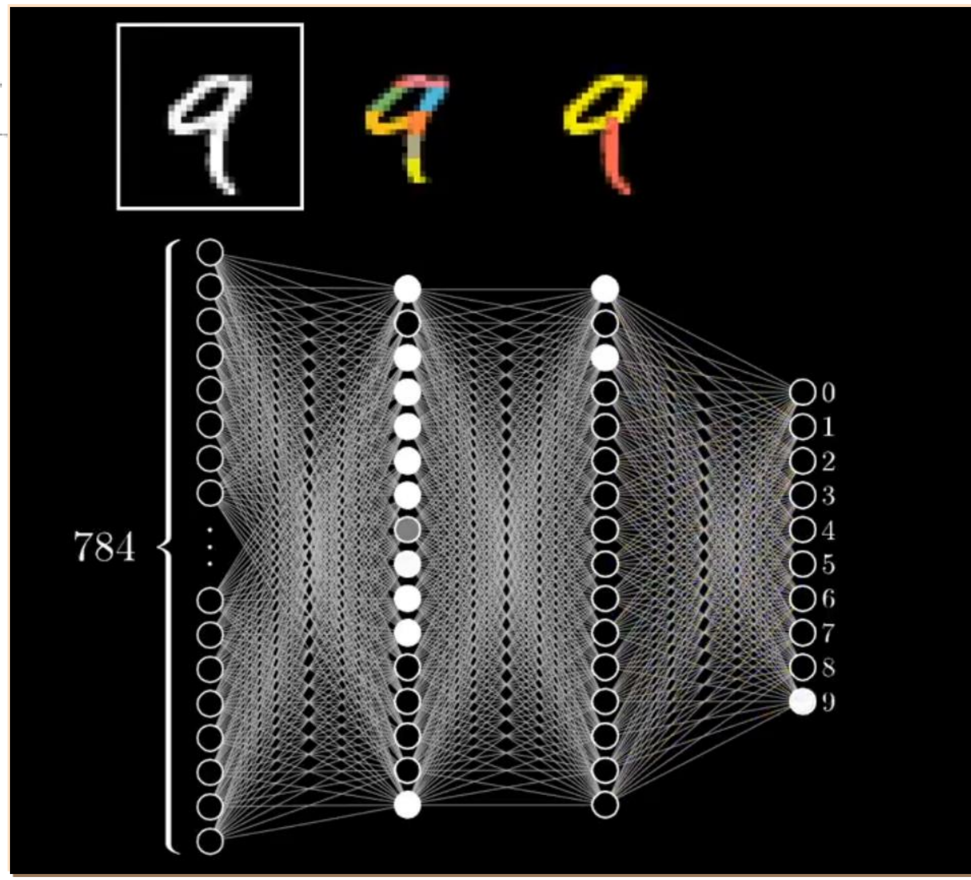
Typical input images

90% accurate learning head pose, and recognizing 1-of-20 faces

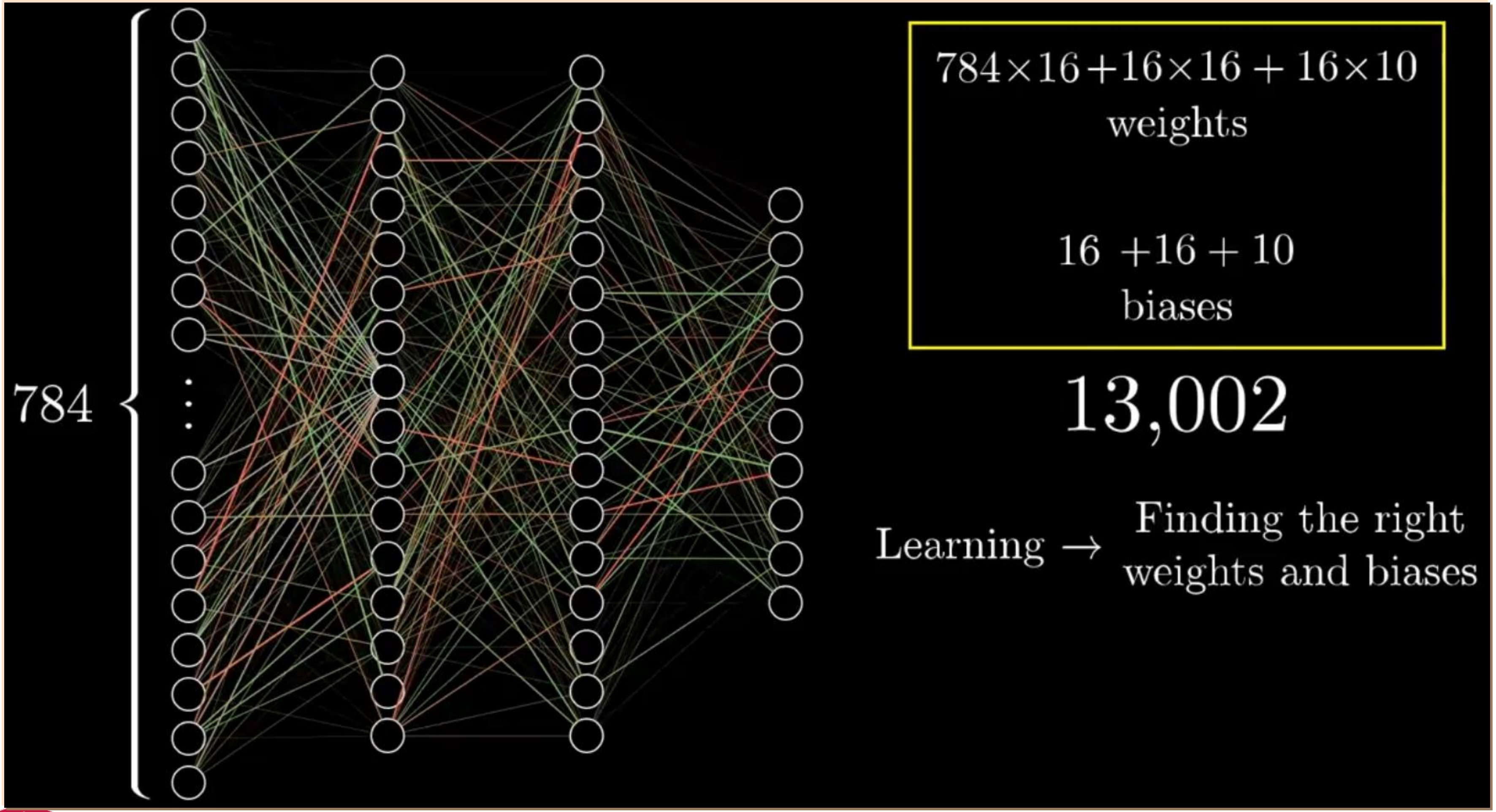
Application examples



Handwritten digit recognition

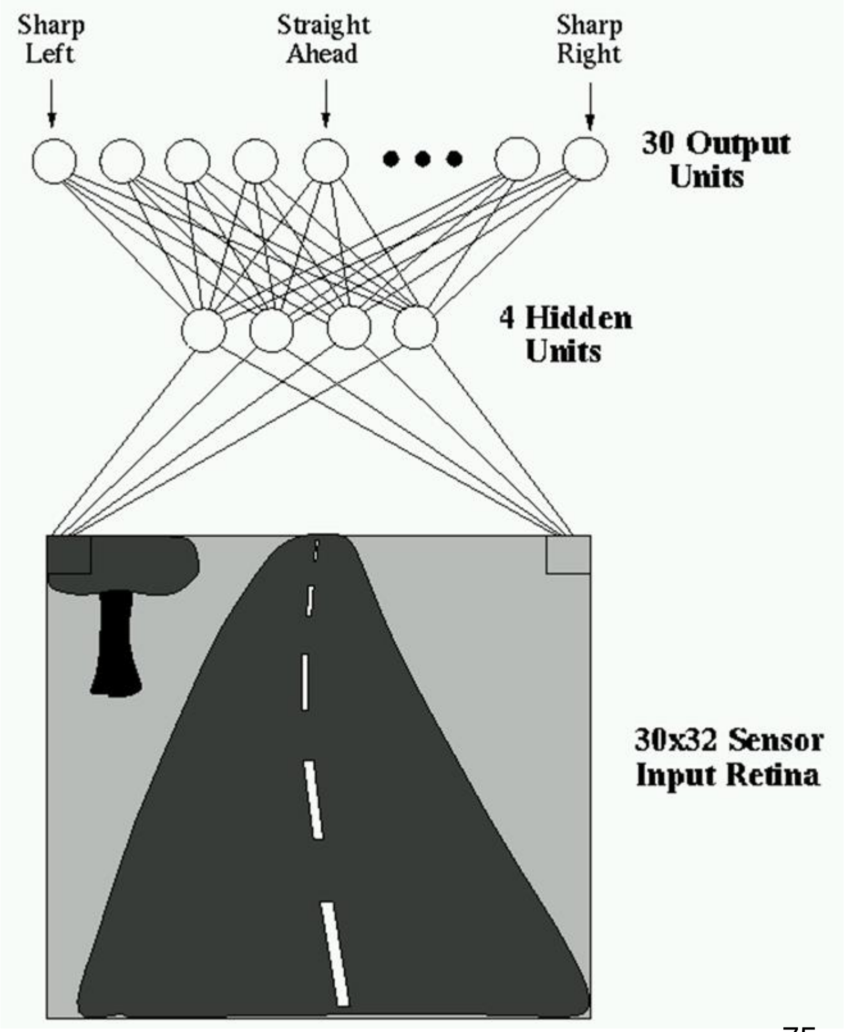


Application examples

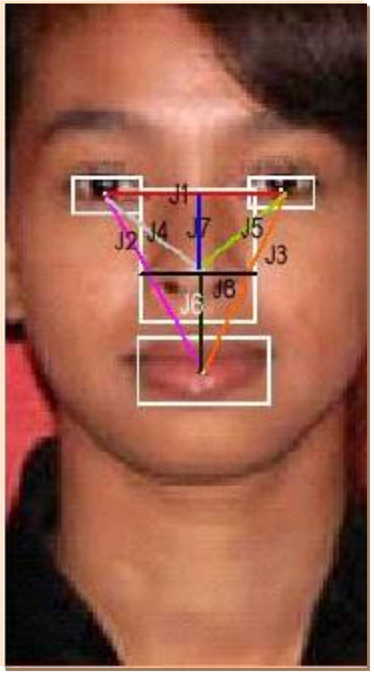


Application examples

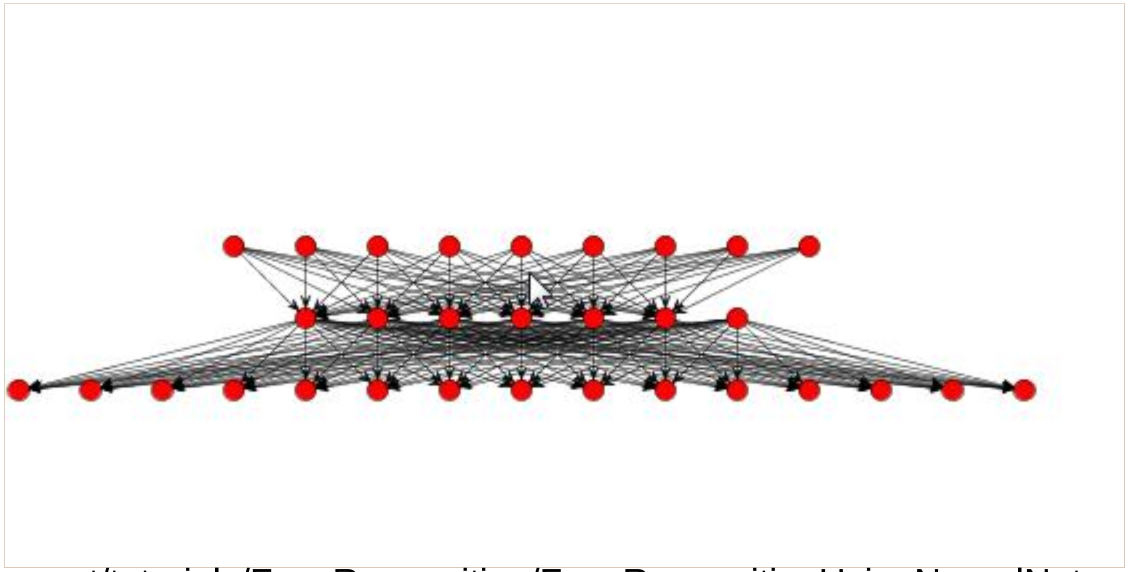
ALVINN = Autonomous Land Vehicle in a Neural Network. Was groundbreaking at the time, although it has now been superseded.
ALVINN drives 70 mp/h on highways.



Application examples



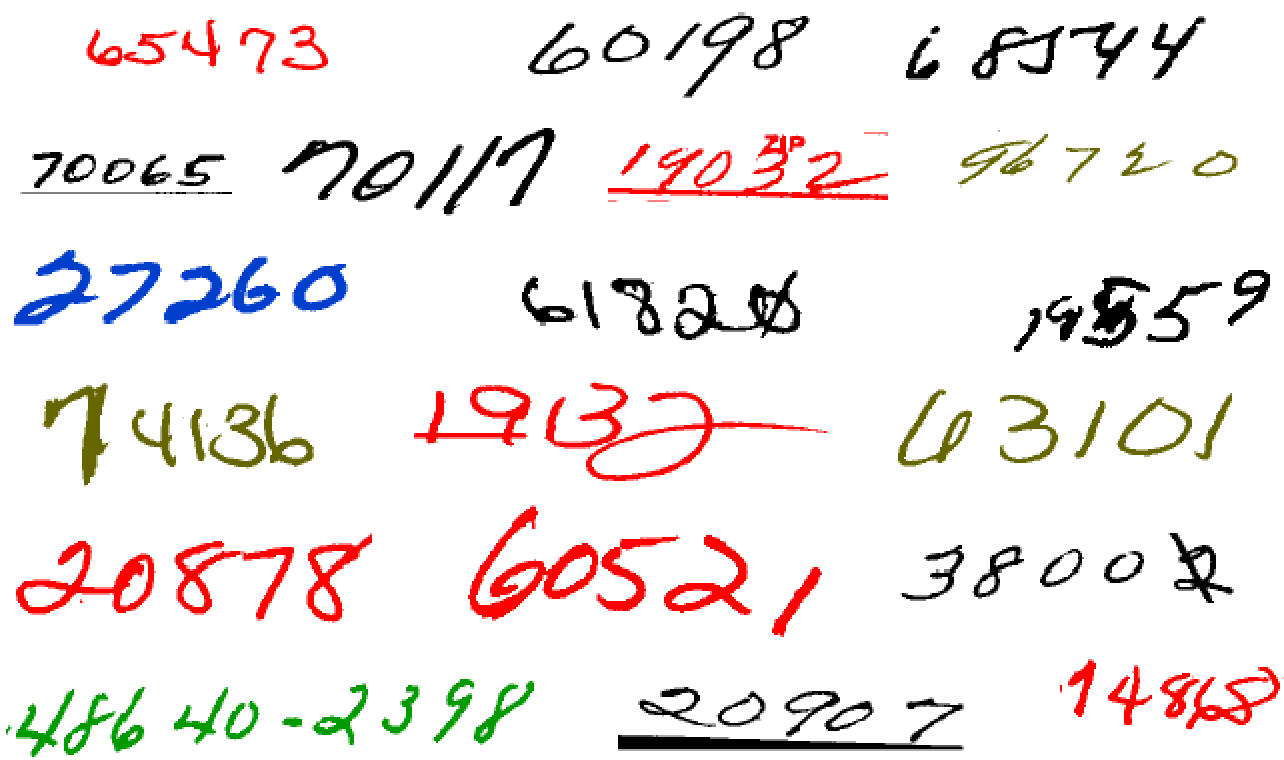
Neural Nets for Face Recognition



http://neuroph.sourceforge.net/tutorials/FaceRecognition/FaceRecognitionUsingNeuralNetwork.html

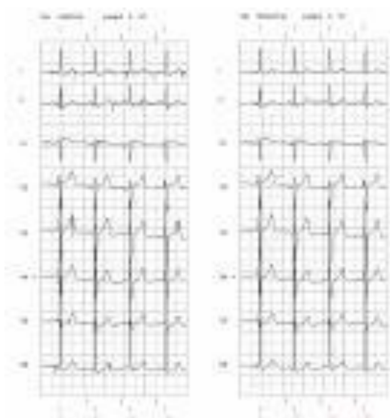
Application examples

Image Recognition and Classification of Postal Codes



Examples of handwritten postal codes drawn from a database available from the US Postal service

Application examples



- "Instant Physician" developed using neural net
- Net presented with a set of symptoms, medical records
- Output is best diagnosis and treatment

- Finger prints never change. Bifurcations or "Minutae"
- Minutiae-based techniques find minutiae points and map their relative placement on the finger
- Large volumes of fingerprints are collected and stored everyday in a wide range of applications including forensics, access control, and driver license registration
- FBI database contains 70 million fingerprints!

Application examples

System Control & Reliability

- Backing Up a truck to a loading dock is a difficult problem for a novice, easy for an experienced driver
- Very difficult problem mathematically
- Can train a neural net to solve it.

- **Automobile airbags can do serious damage**
- **System reliability continuously assessed & failure pre-empted by correct interpretation of data from accelerometers**

Application examples Business

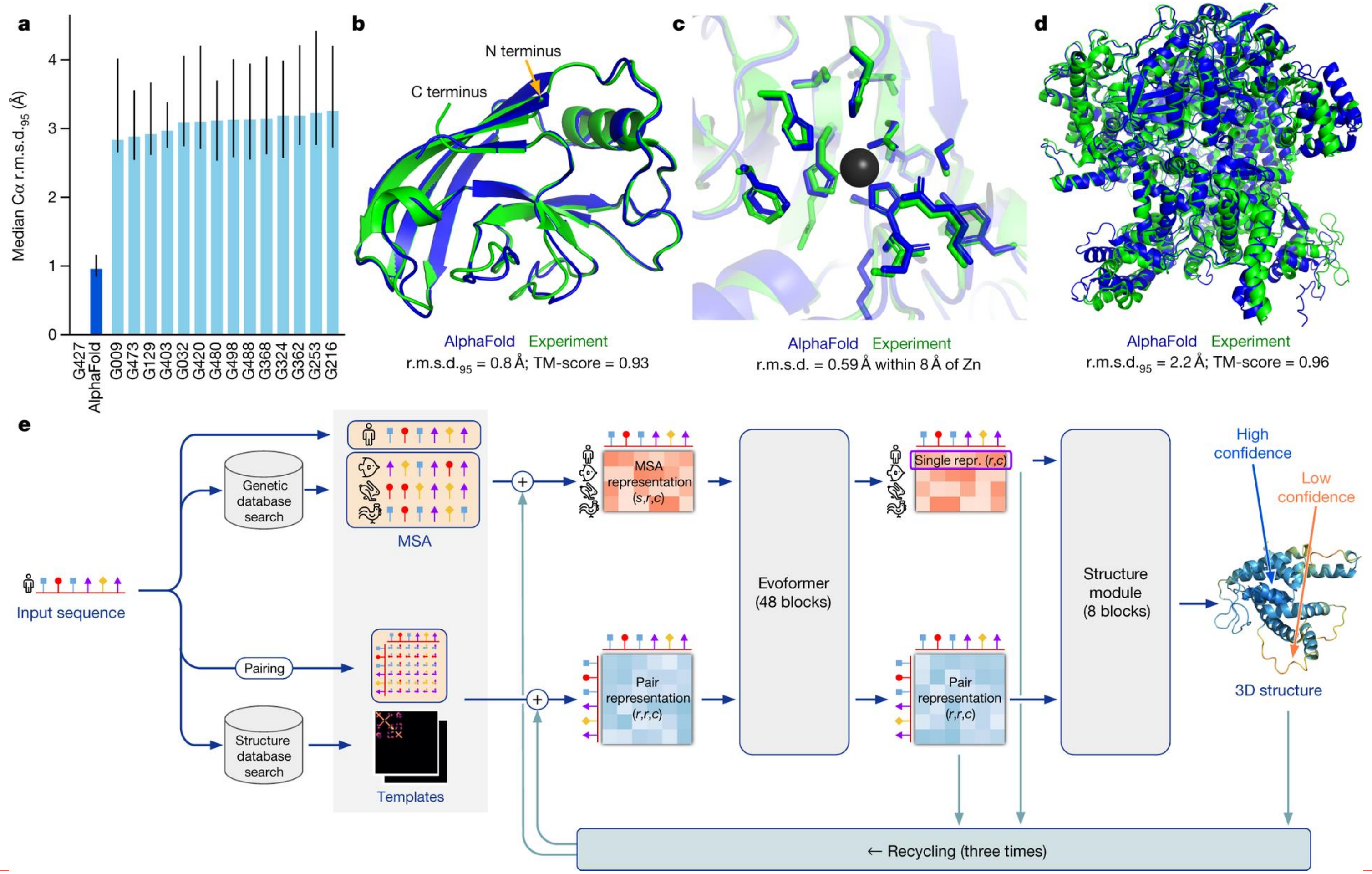
- Mortgage Risk Assessment – reduces delinquency rates
- Inputs include years of employment, # of dependents, property info, income, loan-to-value-ratio
- Output is: Loan or not

- **Prediction of of behavior of stock market indices**
- **Requires knowledge of market history**

- **Time series forecasting**
- **Short and long term predictions**

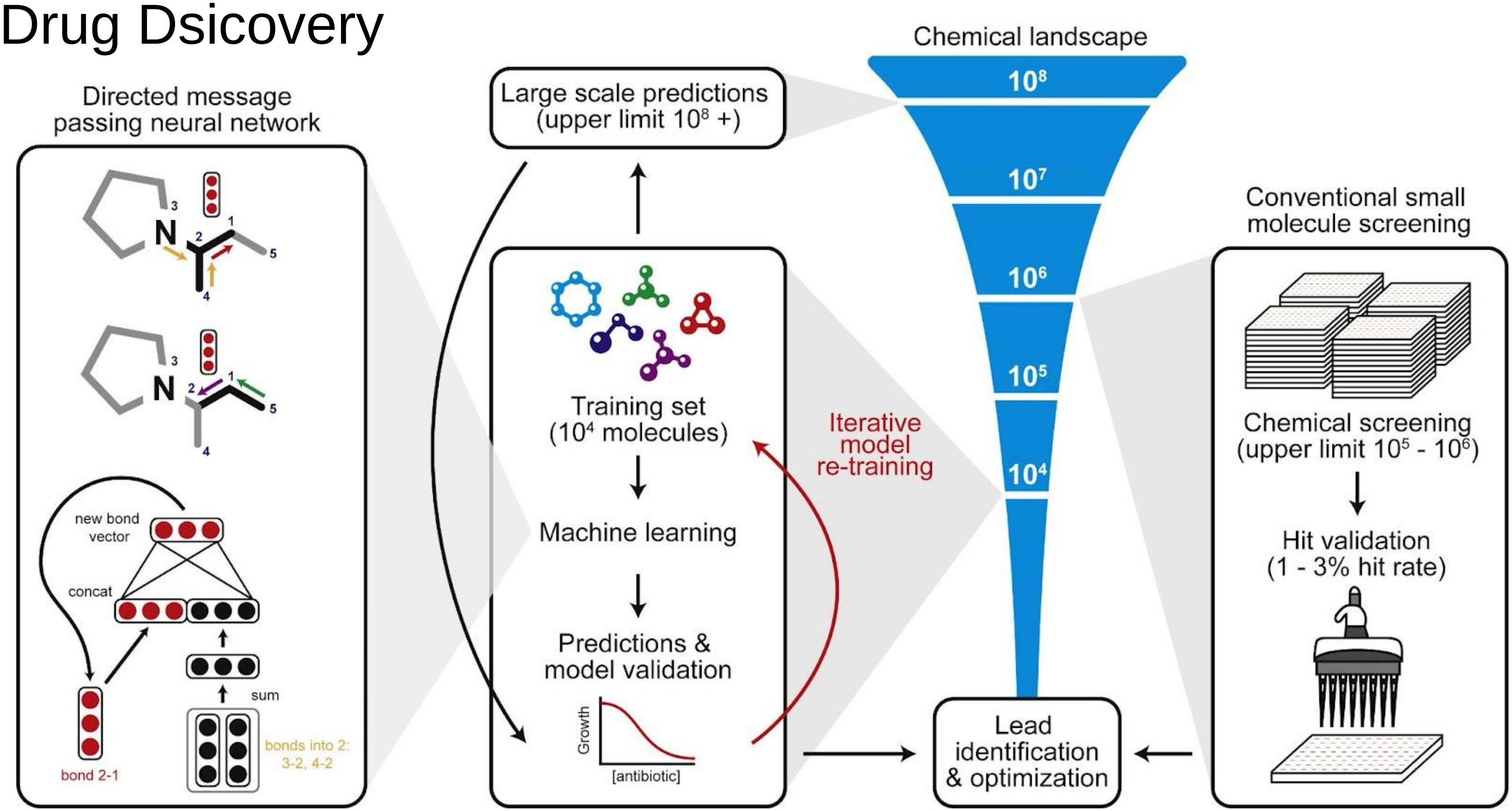
Application examples

DeepMind's AlphaFold (Protein Folding)



Application examples

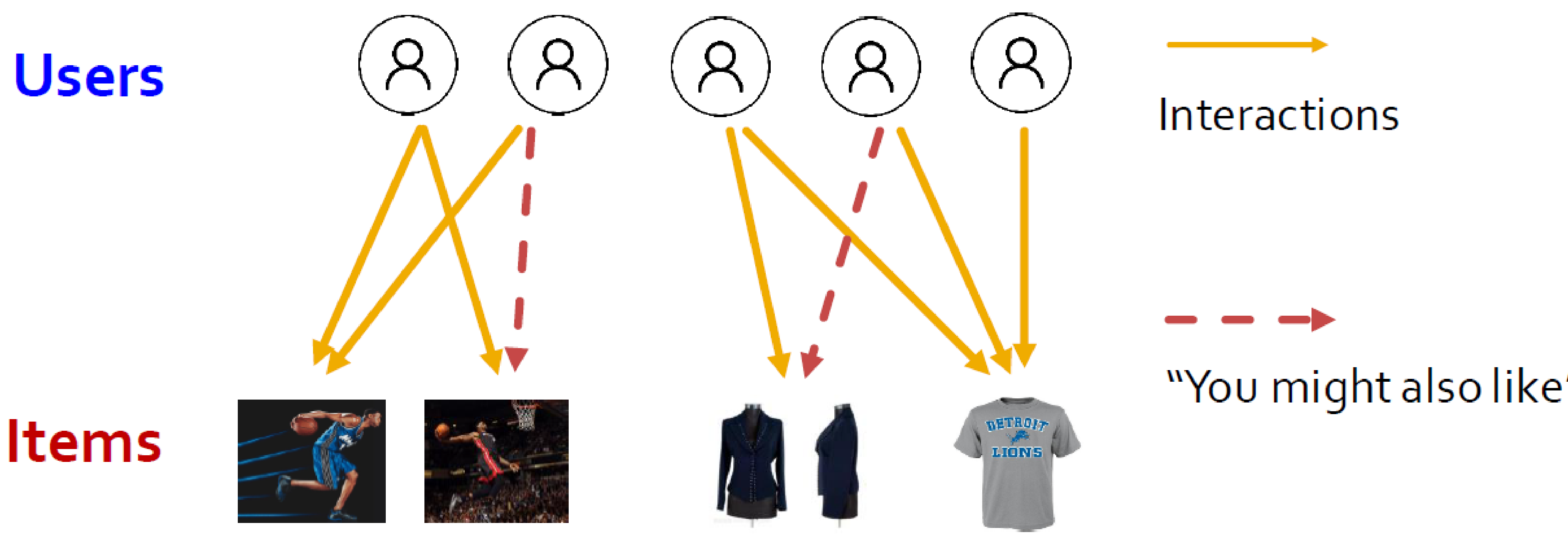
Drug Dsicovery



Application examples

Recommender systems

Goal: Recommend items users might like



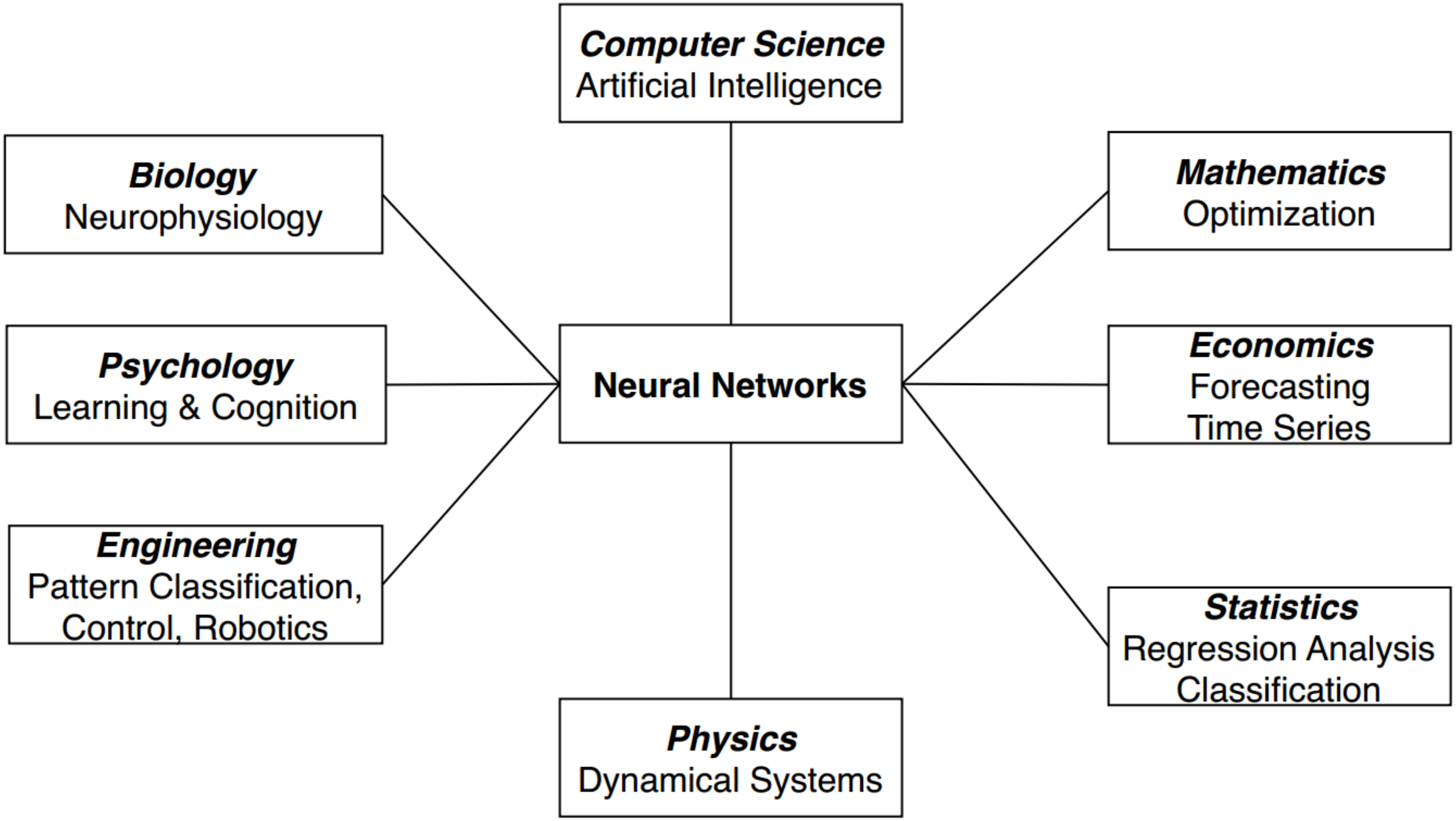
Who is concerned with NNs?

- Computer scientists want to find out about the properties of non-symbolic information processing with neural nets and about learning systems in general.
- Statisticians use neural nets as flexible, nonlinear regression and classification models.
- Engineers of many kinds exploit the capabilities of neural networks in many areas, such as signal processing and automatic control.
- Cognitive scientists view neural networks as a possible apparatus to describe models of thinking and consciousness (High-level brain function).
- Neuro-physiologists use neural networks to describe and explore medium-level brain function (e.g. memory, sensory system, motorics).

Who is concerned with NNs?

- Physicists use neural networks to model phenomena in statistical mechanics and for a lot of other tasks.
- Biologists use Neural Networks to interpret nucleotide sequences.
- Philosophers and some other people may also be interested in Neural Networks for various reasons

Multi-disciplinary area




Questions