IN THE NAME OF ALLAH

Neural Networks

Shahrood University of Technology
Hossein Khosravi

Neural Networks - Introduction

References

- Judith E. Dayhoff, Neural Network Architectures: an introduction, 1990
- Simon Haykin, Neural Networks: A Comprehensive Foundation, Second edition, Prentice-Hall, 1999.
- Simon Haykin, Neural Networks and Learning Machines, 3rd Edition, 2008
- Some Papers

Evaluation

- Presentation and assignments: ~8 points
  - Present a state of the art topic or paper
  - Computer Homeworks
- Final Exam: ~12 points

Neural Networks in the Brain

- Human brain **computes** in an entirely different way from conventional digital computers
- The brain is highly **complex**, **nonlinear**, and **parallel** information processing system
- Organization of neurons to perform tasks much faster than computers.
  - Typical time taken in visual recognition tasks is 100–200 ms.
- Key features of the biological brain: **experience** and **learning.**

Case study: Bat sonar system

- Sonar is an active echo-location system:
  - Providing information about **how far away** a target (e.g. a flying insect) is
  - The relative velocity of the target
  - The size of the target
  - The size of various features of the target
  - And the azimuth and elevation of the target.

- The complex neural computations needed to extract all this information from the target echo occur **within a brain the size of a plum.**

- So blessed is Allah, the best of creators

Applications

- Pattern Recognition
  - Face recognition, character recognition, etc
- Function Approximation (Regression)
  - To find best curve representing some data
- Time Series Prediction
- Clustering
  - Divide samples into some clusters
- Pattern Completion
  - To repair an incomplete pattern

Neural Network Definition

- A neural network is a massively **parallel** distributed processor made up of **simple processing units**, which has a natural propensity for storing **experimental** knowledge and making it available for use
- Neural network resembles the brain in two respects:
  - Knowledge is acquired from the environment through a **learning process**.
  - Interneuron **connection strengths**, known as synaptic weights, are used to store the acquired knowledge.
- learning algorithm:
  - Procedure used for learning
  - Weights, or even the topology can be adjusted.

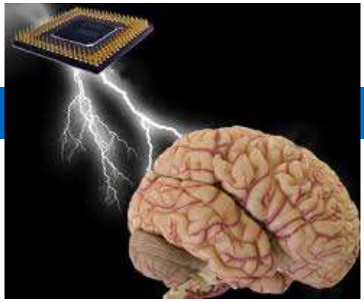
Benefits of Neural Networks

- **Nonlinearity**: neurons can be linear or nonlinear
- **Input-output mapping**: supervised learning similar to nonparametric statistical inference (model-free estimation, no prior assumptions)
- **Adaptivity**: can deal with nonstationary environments
- **Evidential response**: decision plus confidence of the decision can be provided.
- **Contextual information**: Information is distributed through the network: neurons and synapses

Benefits of Neural Networks

- **Fault tolerance:** performance degrades gracefully. There is no catastrophic failure if a neuron or its connecting links are damaged.
- **VLSI implementability:** due to parallel nature of network.
- **Uniformity of analysis and design:**
  - common components (neurons) in all application of NNs.
  - sharability of theories and learning algorithms
- **Neurobiological analogy:** Neural nets motivated by neurobiology, and neurobiology also turning to neural networks for insights and tools.

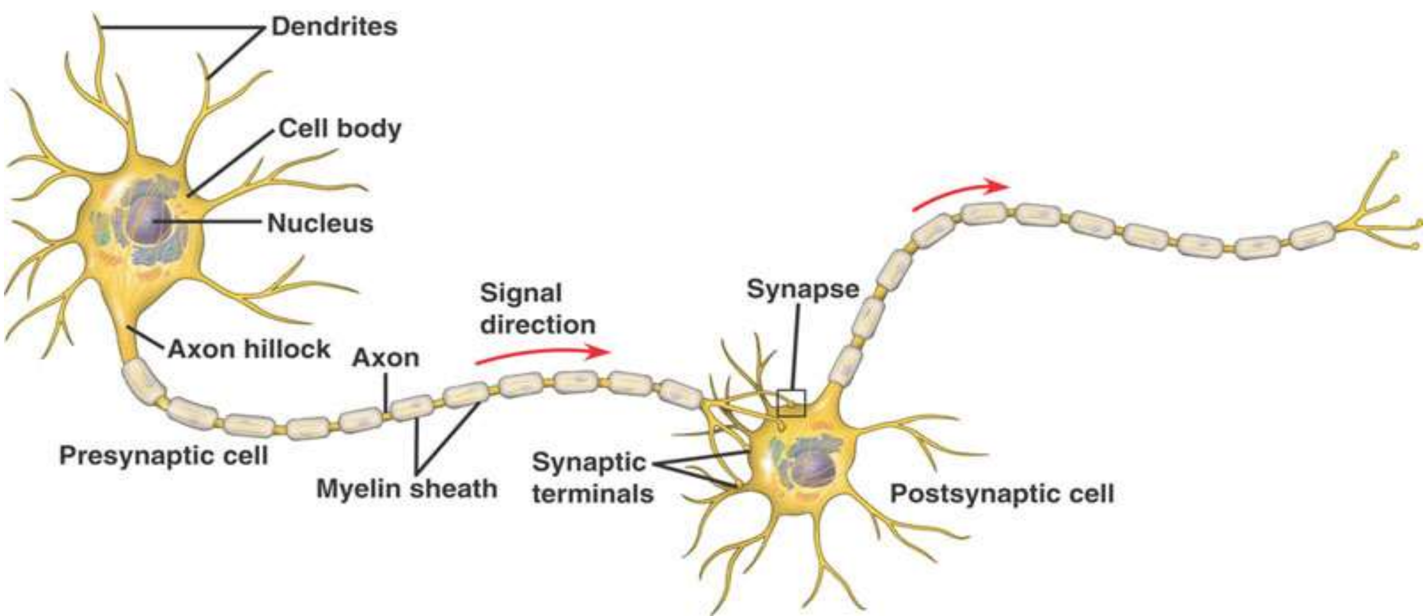
Human Nervous System



- Simple Brain Model
  - Stimulus ➢ Receptors ⮜⮞ Neural Net ⮜⮞ Effectors ➢ Response
- Neuron
  - In 1911 Santiago Ramony Cajal, a Spanish neuro-anatomist introduced neurons as a fundamental unit of brain function
- Neurons are slow
  - 1ms per operation, compared to 1ns of modern CPUs
- Huge number of neurons and connections
  - 10^{10} neurons, and 6×10^{13} connections in human brain.
- Highly energy efficient:
  - 10^{-16} J per operation per second in the brain vs. 10^{-6} J in modern computers

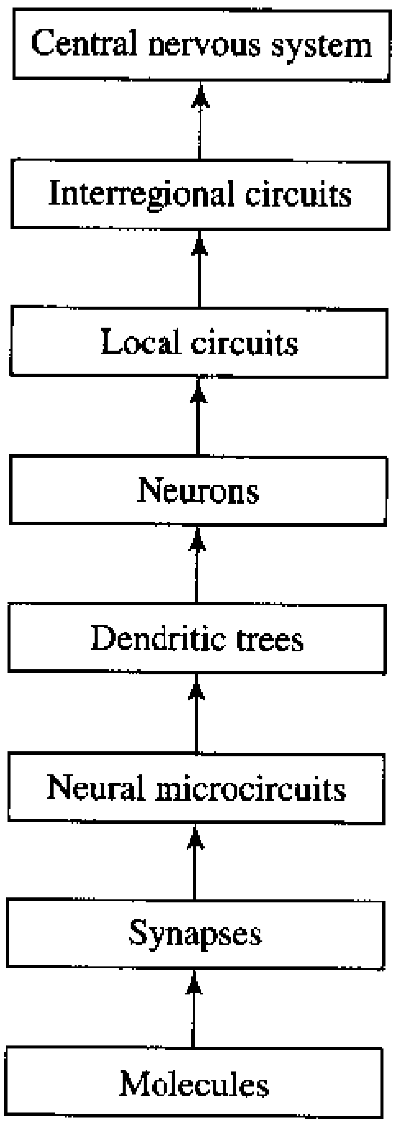
Neuron and Synapse

- Synapse: where two neurons meet.
- Presynaptic neuron: source, Postsynaptic neuron: target
- Neurotransmitters: molecules that cross the synapse (positive or excitatory, negative or inhibitory, or modulatory effect on postsynaptic activation)
- Dendrite: branch that receive input
- Axon: branch that send output (spike, or action potential)



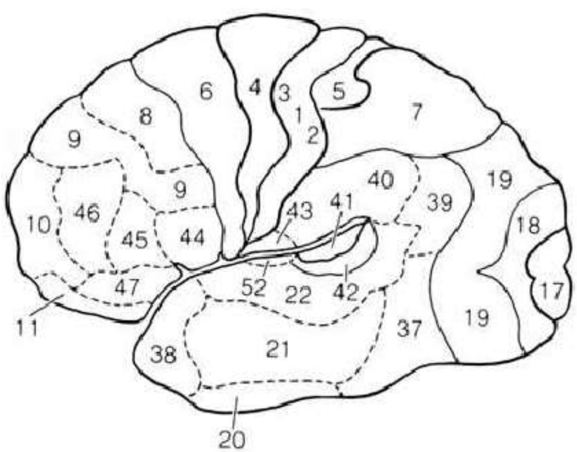
Structural Organization of the Brain

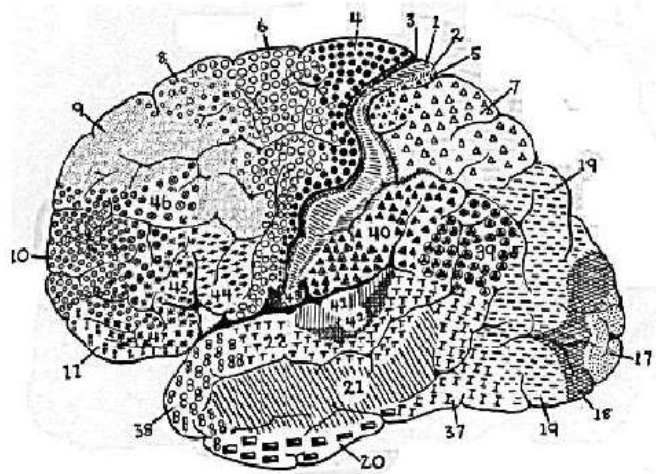
- Small to large-scale organizations
  - Molecules, Synapses,
  - Neural microcircuits: assembly of synapses
    - Like a silicon chip made up of transistors
    - size: several µm, speed: several ms
  - Dendritic trees, Neurons
  - Local circuits: assembly of local neurons
    - Size: about 1mm
  - Interregional circuits: pathways, columns, topographic maps
  - Central nervous system



Cytoarchitectural Map of the Cerebral Cortex

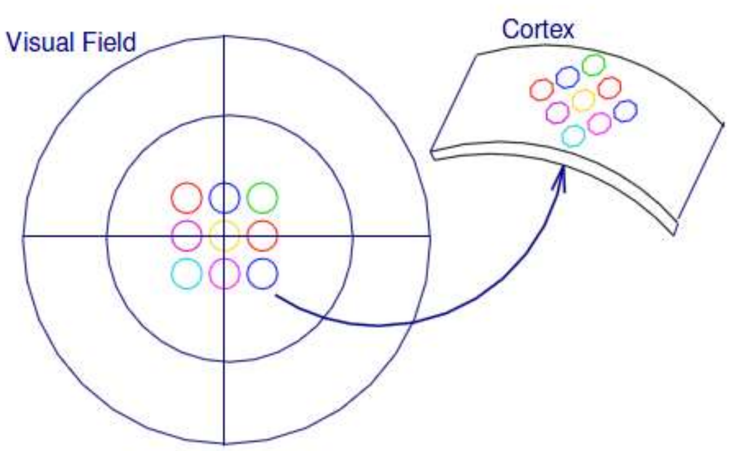
- Map-like organization:
- Area 17, 18, 19: visual cortices
- Area 41, 42: auditory cortices
- Area 1, 2, 3: somatosensory cortices (bodily sensation)

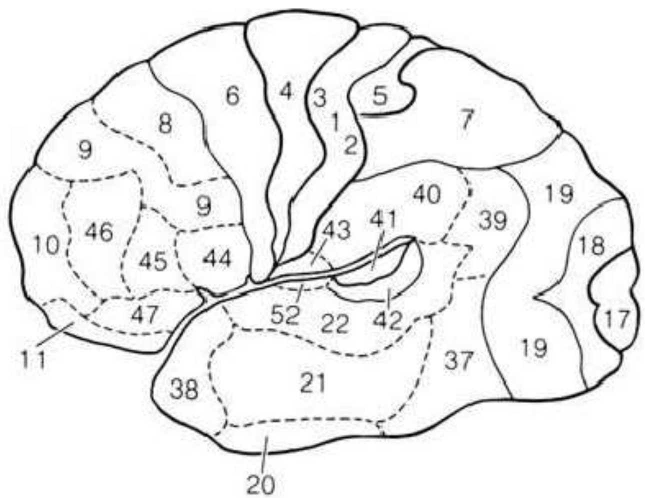




Topographic Maps in the Cortex

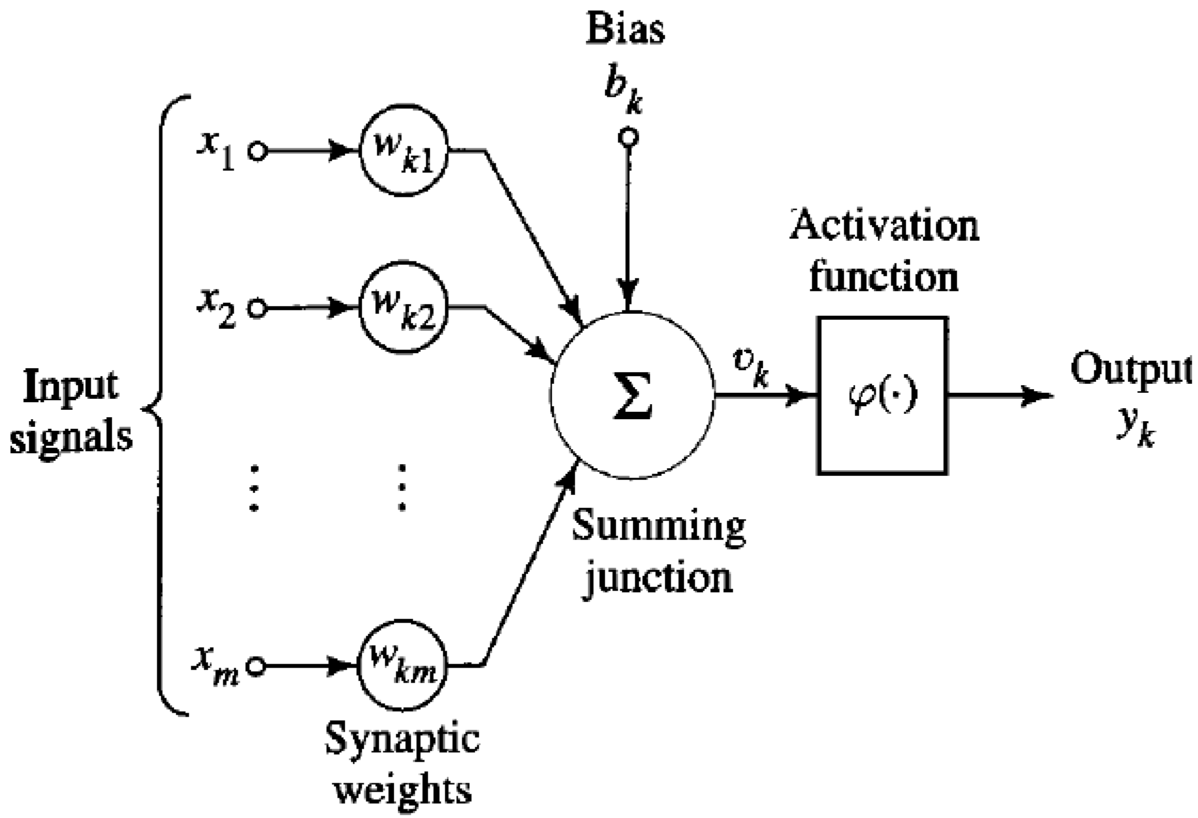
- Nearby location in the stimulus space are mapped to nearby neurons in the cortex.
- Thus, it is like a map of the sensory space, thus the term topographic organization.
- Many regions of the cortex are organized this way: visual (V1), auditory (A1), and somatosensory (S1) cortices

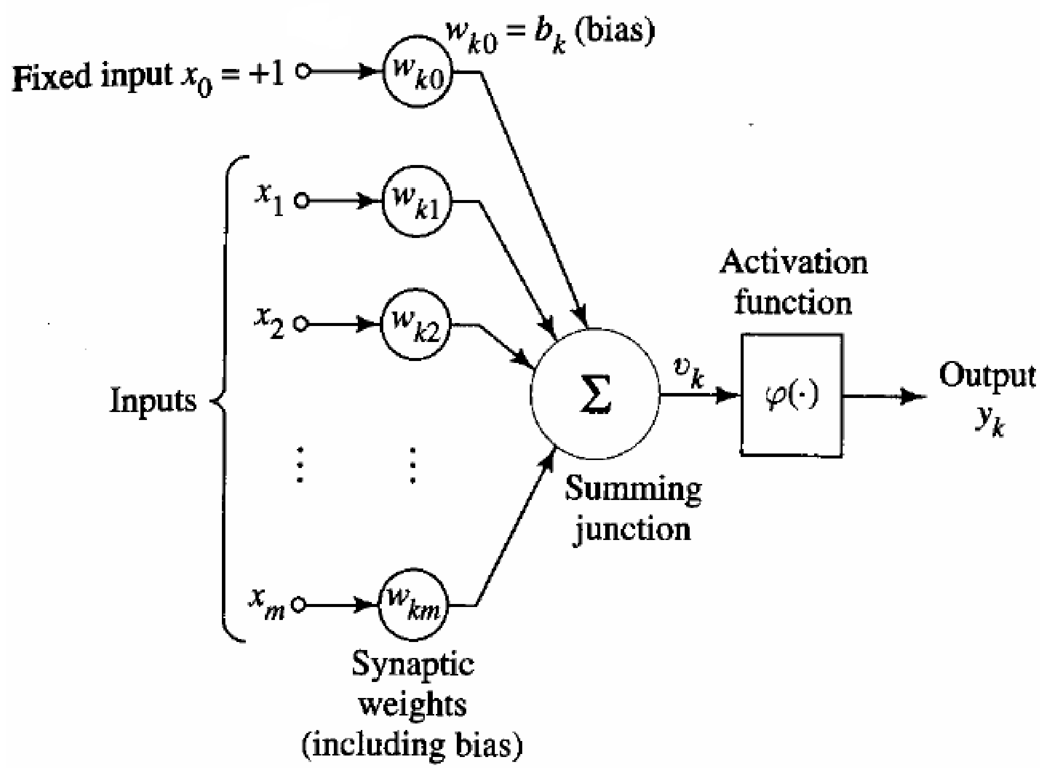




Models of (artificial) neuron

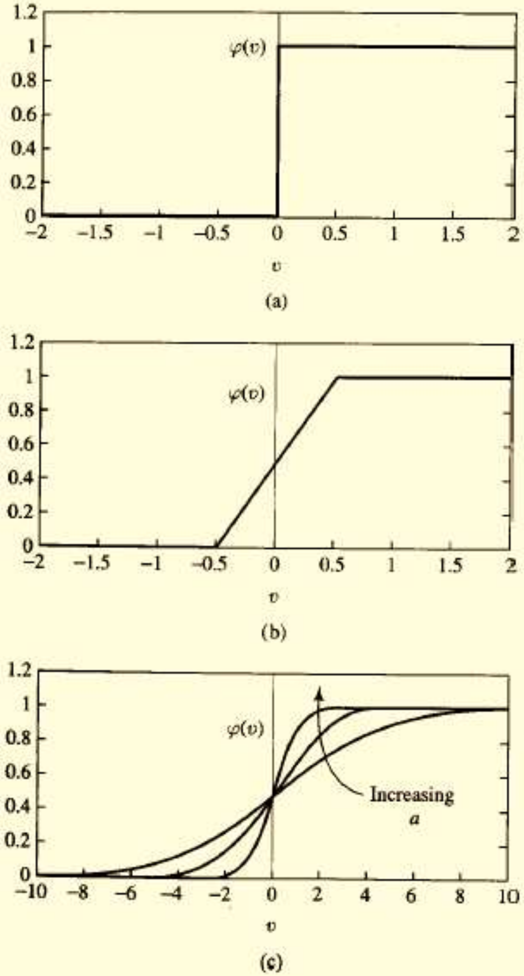
- Neuron: information processing unit fundamental to neural network operation.





- Synapses with associated **weights**: j to k denoted w_{kj}.
- Summing function: $u_k = \sum_{j=1}^{m} w_{kj} x_j$
- Activation function: $y_k = \phi(u_k + b_k)$
- Bias b_k: $v_k = u_k + b_k$, or $v_k = \sum_{j=0}^{m} w_{kj} x_j$

Activation functions



- Threshold unit:

$$\phi(v) = \begin{cases} 1 & \text{if } v \geq 0 \\ 0 & \text{if } v < 0 \end{cases}$$

- Piece-wise linear:

$$\phi(v) = \begin{cases} 1 & \text{if } v \geq +\frac{1}{2} \\ v & \text{if } +\frac{1}{2} > v > -\frac{1}{2} \\ 0 & \text{if } v \leq -\frac{1}{2} \end{cases}$$

- Sigmoid: logistic function (a: slope parameter)

$$\phi(v) = \frac{1}{1 + \exp(-av)}$$

It is differentiable: $\phi'(v) = a\phi(v)(1 - \phi(v)$.

Other activation functions

- Signum function:

$$\phi(v) = \begin{cases} 1 & \text{if } v > 0 \\ 0 & \text{if } v = 0 \\ -1 & \text{if } v < 0 \end{cases}$$

- Sign function:

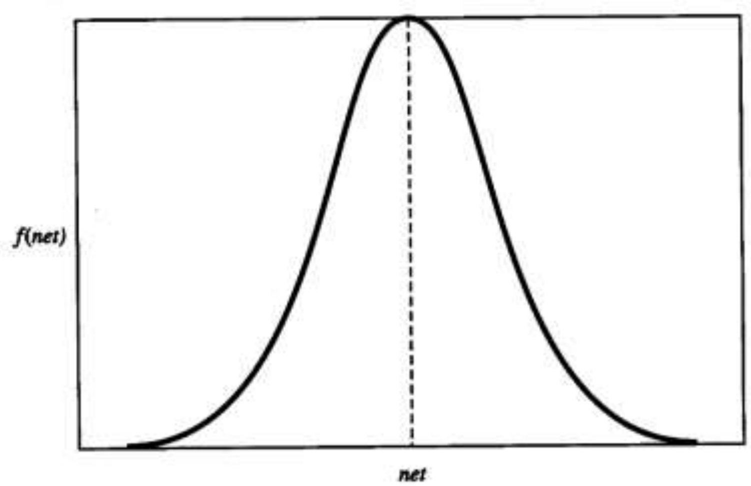
$$\phi(v) = \begin{cases} 1 & \text{if } v \geq 0 \\ -1 & \text{if } v < 0 \end{cases}$$

- Hyperbolic tangent function:

$$\phi(v) = \tanh(v)$$

$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{e^{2x} - 1}{e^{2x} + 1}$$

Radial Basis Function



Signal-flow Graphs