MSc on Intelligent Critical Infrastructure Systems

Machine Learning

Lecture 3

Marios Polycarpou

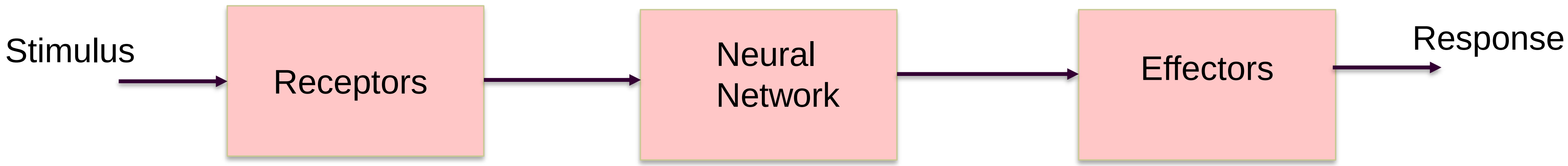
Director, KIOS Research and Innovation Center of Excellence

Professor, Electrical and Computer Engineering

University of Cyprus

funded by:

Neural Networks – Motivation



- The brain is a complex, nonlinear and distributed computer having neurons as its basic information processing units (different than traditional computers).
- The brain has the ability to perform several tasks such as pattern recognition, perception and motor control very well, despite being slow in information processing.
- Therefore, the motivation is to mimic the functioning neurons and neural networks *in-silico* so as to build machines that have very high capabilities.

Artificial Neural Networks – History

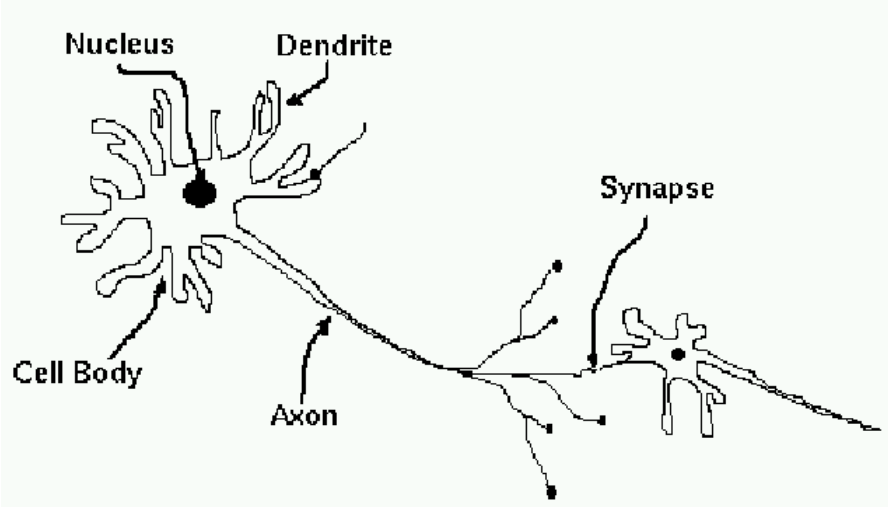
- 1943: computational model for neural networks (McCulloch &Pitts)
- 1958: Perceptron was created (Rosenblatt)
- 1969: was shown that Perceptrons were not powerful (Minsky & Papert)
- 1986: multi-layer perceptrons (Rumelhart & McClelland)
- 1990: IEEE Transactions on Neural Networks
- 2010-- : Deep learning, wide number of applications

Key Properties of Artificial Neural Networks (ANN)

- **Parallel Processing**
  - Fault Tolerance
  - Hardware implementation
  - Real time computations
- **Learning generalization**: produce reasonable outputs for inputs not encountered during training
- **Nonlinearity:** achieves better performance when modelling nonlinear phenomena
- **Adaptivity**: change weights in real time according to the evolution of the environment
- **Uniformity of analysis and design**: theory sharing, construction of modular networks
- **Neurobiological analogy**: Engineers can look for inspiration in neurobiology and neurobiologists can use ANNs as a tool to interpret neurobiological phenomena

Neurons

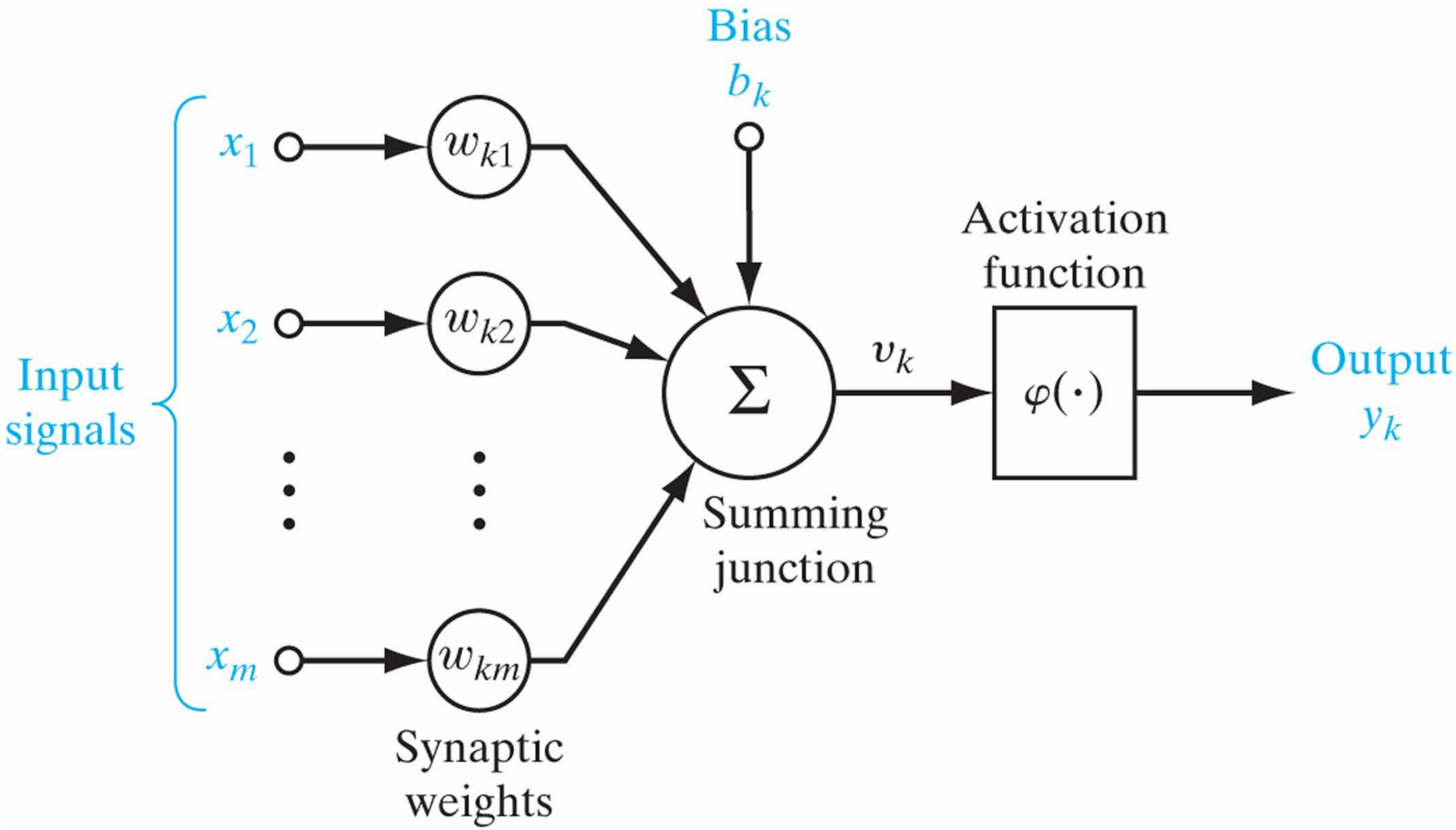
- On the average, humans have 86 billion neurons in their brain and approximately 100 trillion synapses. There are 16 billion neurons in the forebrain.
- Neuron consists of a cell body, dendrites and an axon
- Neurons are massively interconnected by *synapses*
- Synapse is an interconnection between the axon of one neuron and a dendrite of another neuron
- Information propagation is achieved via electro-chemical signals



Neurons

	Number of Neurons in brain
Lobster	100,000
Dog	2 Billion
Giraffe	10 Billion
Gorilla	33 Billion
Human	86 Billion
African Elephant	257 Billion

Artificial Neurons



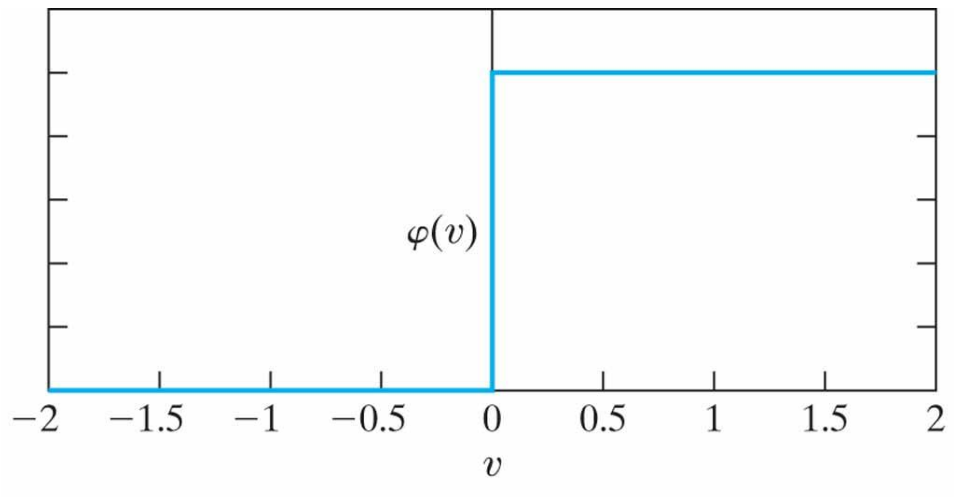
$$v_k = b_k + \sum_{i=1}^{m} w_{ki} x_i = \sum_{i=0}^{m} w_{ki} x_i, \qquad w_{k0} = b_k, \quad x_0 = 1$$

Activation Functions

Threshold function:

Neurons with this activation function can only be in two states: *“on”* and *“off”.*

$$\varphi(u)=\begin{cases}1, & u\geq u_{th}\\ 0, & u<u_{th}\end{cases}$$

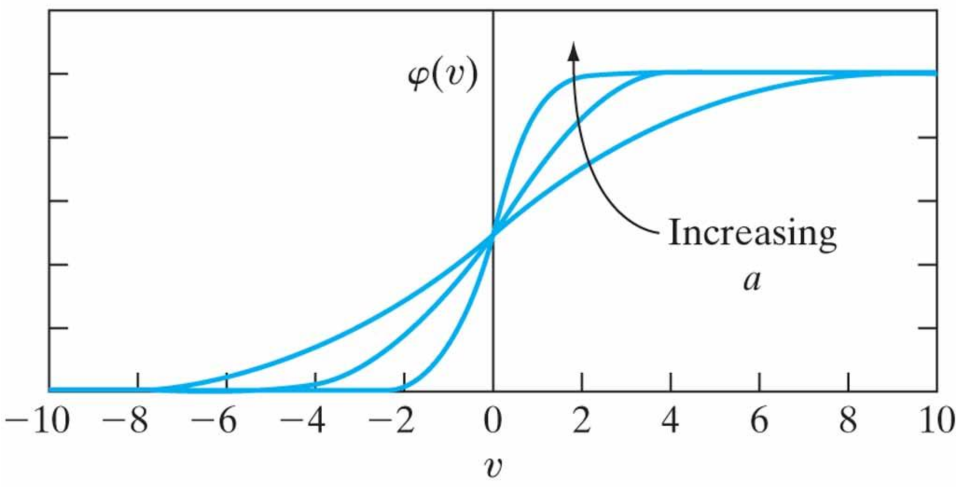


(a)

Logistic Sigmoid function:

The slope shows how fast a neuron moves from the *“off”* state to the *“on”* state:

$$\varphi(u)=\frac{1}{1+e^{-au}}$$

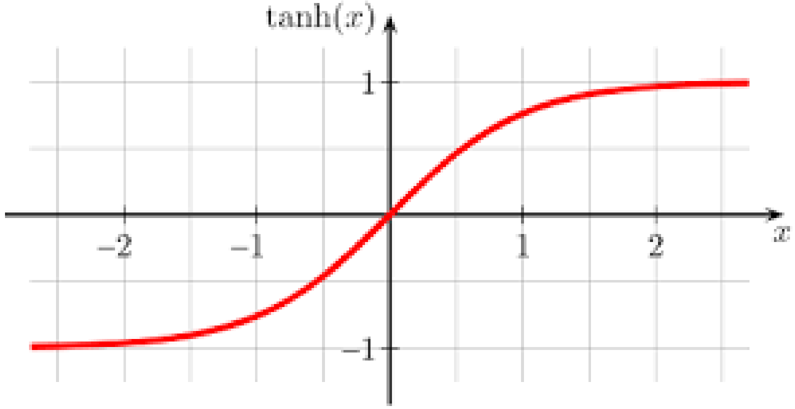


(b)

Activation Functions

Hyperbolic function:

$$\varphi(u) = \tanh(u)$$

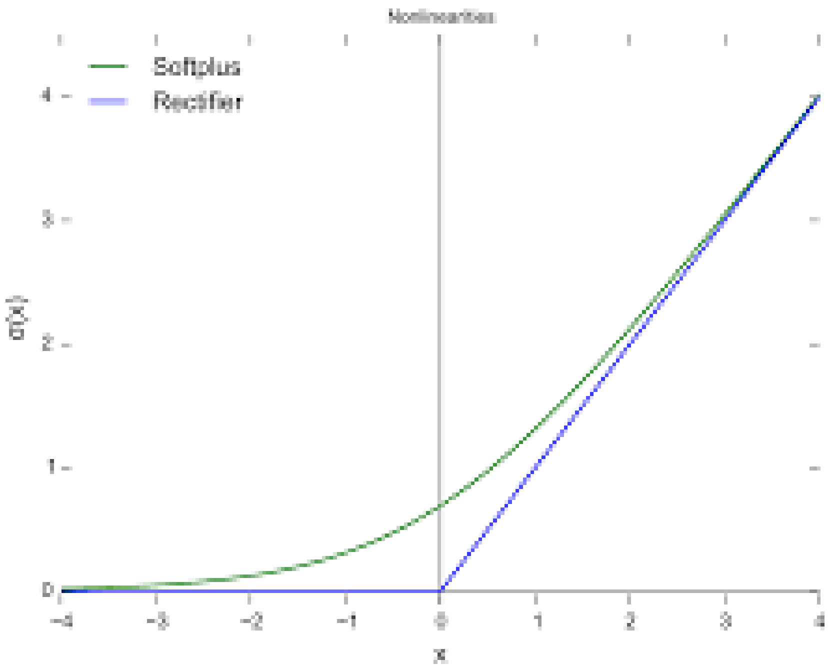


Rectifier Linear Unit (ReLU):

$$\varphi(u) = \max(0, u) \triangleq u^{+}$$

Softplus or SmoothReLU:

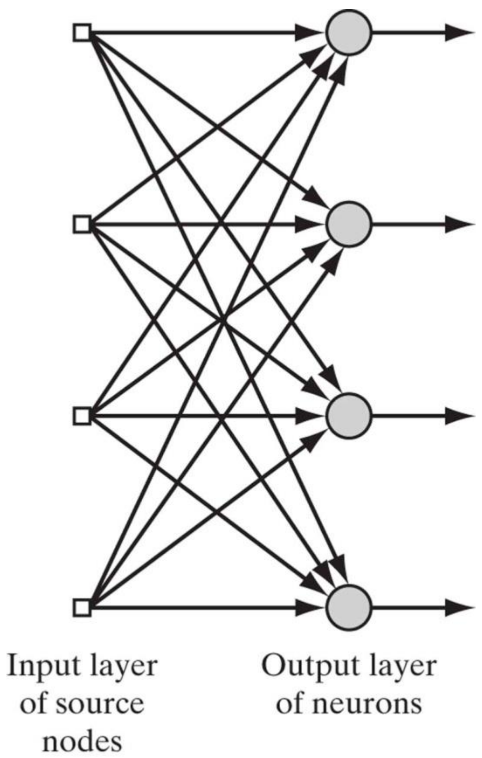
$$\varphi(u) = \ln(1 + e^{u})$$



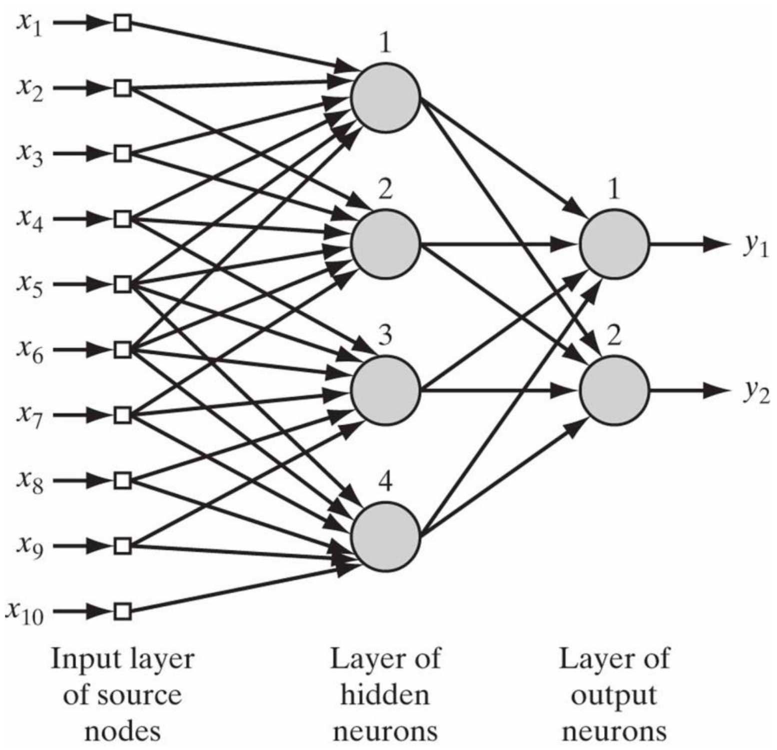
Building Artificial Neural Networks (ANNs)

Feedforward neural networks

Mostly applied to the broad areas of function approximation, classification and data processing (e.g. compression)

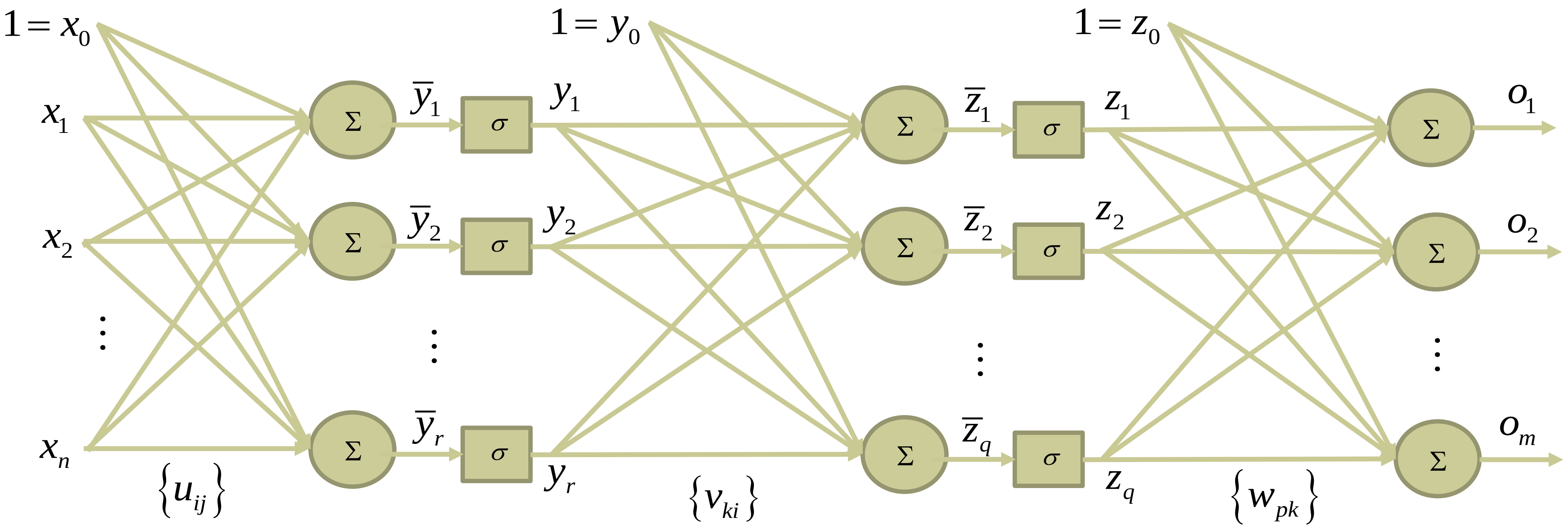


Single Layer



Multilayer

Multilayer Neural Networks



$$\underset{[m\times 1]=[m\times(q+1)]\times[(q+1)\times 1]}{O = WZ}$$

$$z_i = \sigma(\bar{z}_i)\,,\ i=1,\ldots,q \quad , \quad \underset{[(q+1)\times 1]=[q\times(r+1)]\times[(r+1)\times 1]}{\bar{Z} = VY}$$

$$y_j = \sigma(\bar{y}_j)\,,\ j=1,\ldots,r \quad , \quad \underset{[(r+1)\times 1]=[r\times(n+1)]\times[(n+1)\times 1]}{\bar{Y} = UX}$$

Weights: W, V, U

Input: X

Output: O

Activation or squashing function: $\sigma(.):\mathbb{R}\mapsto\mathbb{R}$

Building Artificial Neural Networks (ANNs)

Recurrent neural networks

Applied to areas related to optimization and temporal processing such as time series prediction, speech recognition, etc.