

NEURAL NETWORKS

Learning in Neural Networks

Lecture 4

These lecture notes are based on the textbook;

- Simon Haykin, **Neural Networks: A Comprehensive Foundation**, Macmillan College Publishing Company, 1994.

Asst. Prof.Dr. Sibel SENAN
ssenan@istanbul.edu.tr

Introduction

- ❑ The property that is of primary significance for a neural network is the ability of the network to learn from its environment.
- ❑ Learning involves adjustments to the synaptic connections that exist between the neurons.
- ❑ Ideally the network becomes more knowledgeable about its environment after each iteration of the learning process.

❑ Definition of *learning* in the context of neural networks :

Learning is process by which the free parameters of a neural network are adapted through a process of stimulation by the environment in which the network is embedded.

The type of learning is determined by the manner in which the parameter changes take place.

❑ Definition of learning process implies the following sequence of events :

1. The neural network is *stimulated* by an environment.
2. The neural network *undergoes changes* in its free parameters as a result of this stimulation.
3. The neural network *responds in a new way* to the environment because of the changes that have occurred in its internal structure.

-
- A set of well-defined rules for the solution of a learning problem is called a learning algorithm.
 - Learning algorithms differ from each other in the way in which the adjustment to a synaptic weight of a neuron is formulated.
 - Another factor to be considered is the manner in which a neural network relates to its environment. In this context we speak of a learning paradigm that refers to a model of the environment in which the neural network operates.

➤ Two fundamental learning paradigms :

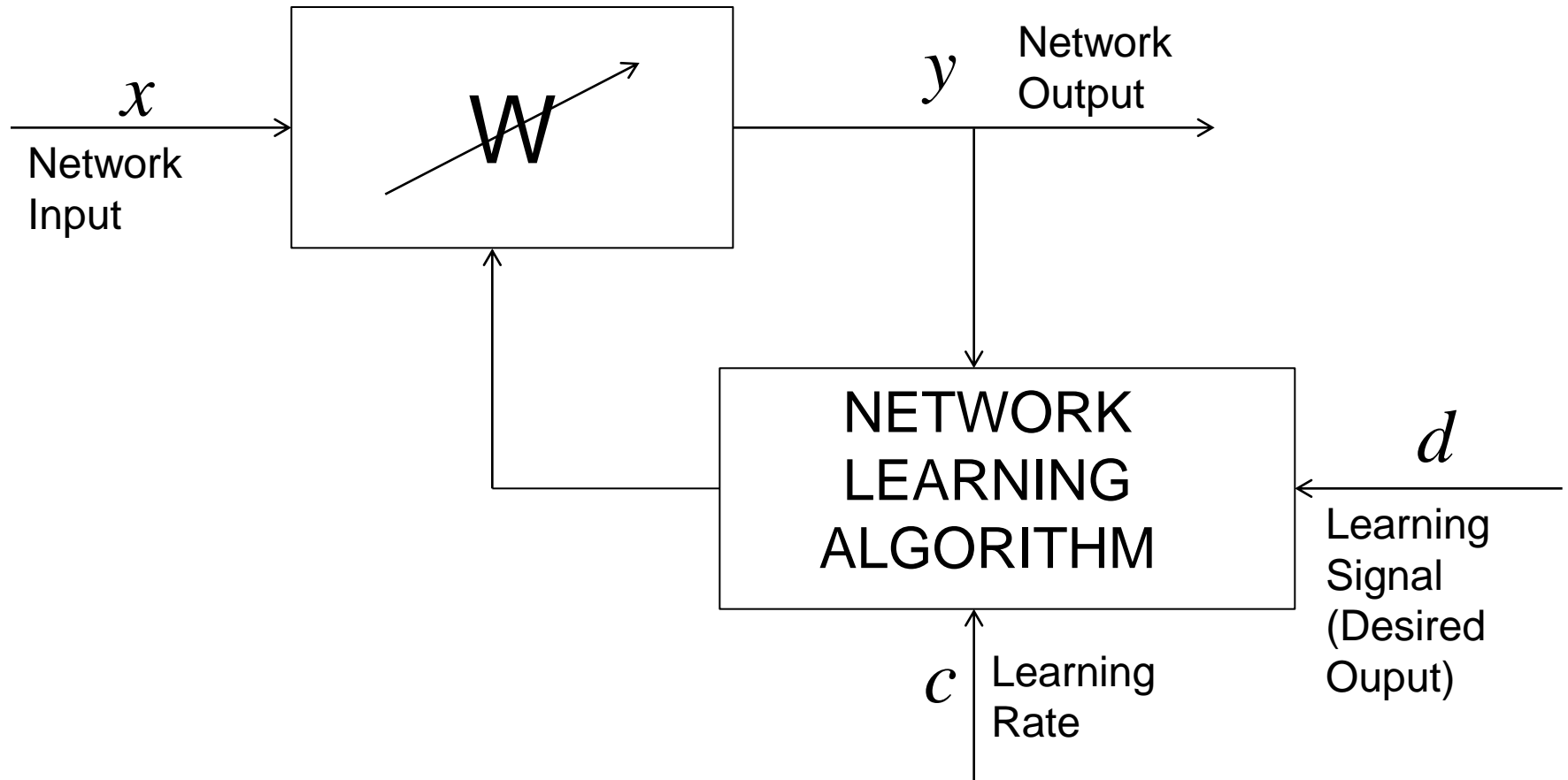
1. Learning with a teacher
(SUPERVISED LEARNING)

Desired output is given to the network

2. Learning without a teacher
(UNSUPERVISED LEARNING)

There is NO desired output

1. SUPERVISED LEARNING



The learning rule is provided with a set of training set of proper network behavior :

$$\{x_1, d_1\}, \{x_2, d_2\}, \dots, \{x_n, d_n\}$$

x_n : Input to the network

d_n : Corresponding correct desired output

As the inputs are applied to the network, the network outputs are compared with the desired outputs.

GENERAL RULES FOR NEURON i

- The weight vector

$$w_i = [w_{i1} \quad w_{i2} \quad \dots \quad w_{in}]$$

- The learning signal

$$r = f(w_i, x, d_i)$$

- The increment of the weight vector w_i , (Δw_i) produced by the learning step at time t according to the general rule is:

$$\Delta w_i(t) = c \cdot r[w_i(t), x(t), d_i(t)] \cdot x(t)$$

- The learning rate $c > 0$.
- The adapted weight vector

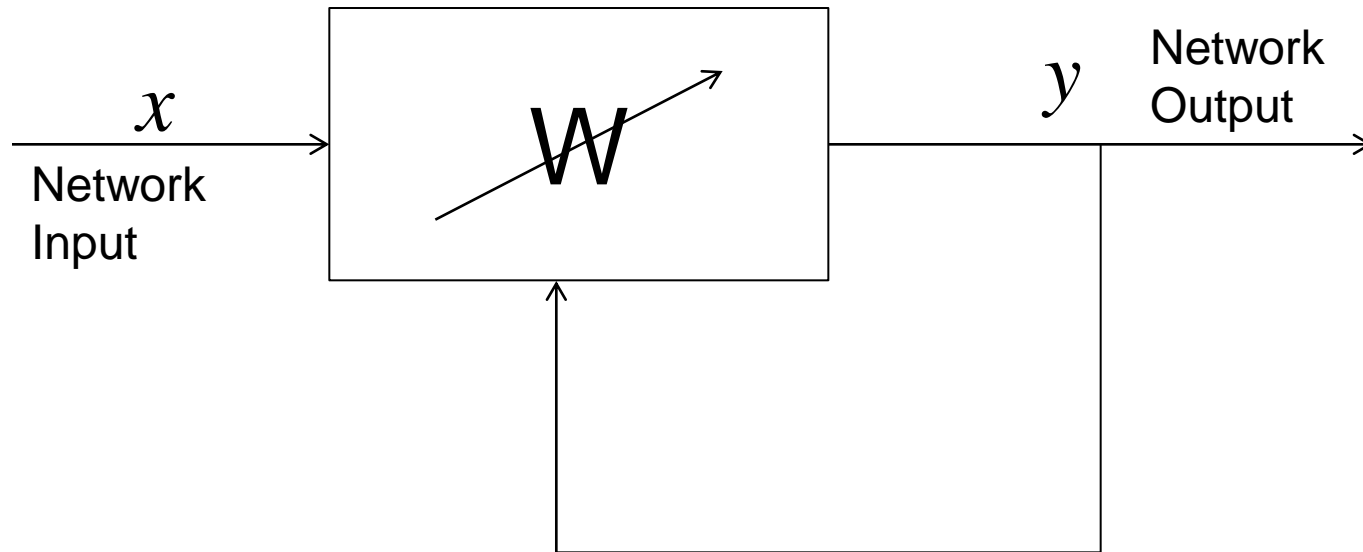
$$w_i(t + 1) = w_i(t) + \Delta w_i(t)$$

General Discrete-Time Form :

$$w_i^{k+1} = w_i^k + \Delta w^k$$

$$w_i^{k+1} = w_i^k + c \cdot r[w_i^k, x^k, d_i^k] \cdot x^k$$

2. UNSUPERVISED LEARNING



- *Learning signal is not available*

GENERAL RULES FOR NEURON i

- The weight vector

$$w_i = [w_{i1} \quad w_{i2} \quad \dots \quad w_{in}]^T$$

- The learning signal

$$r = f(w_i, x)$$

- The increment of the weight vector w_i , (Δw_i) produced by the learning step at time t according to the general rule is:

$$\Delta w_i(t) = c \cdot r[w_i(t), x(t)] \cdot x(t)$$

- The learning rate $c > 0$.
- The adapted weight vector

$$w_i(t + 1) = w_i(t) + \Delta w_i(t)$$

General Discrete-Time Form :

$$w_i^{k+1} = w_i^k + \Delta w^k$$

$$w_i^{k+1} = w_i^k + c.r[w_i^k, x^k].x^k$$

LEARNING RULES

- PERCEPTRON (PERCEPTION) LEARNING RULE (*supervised*)
- DELTA LEARNING RULE (*supervised*)
- WIDROW-HOFF LEARNING RULE (*supervised*)
- CORRELATION LEARNING RULE (*supervised*)
- HEBBIAN LEARNING RULE (*unsupervised*)