NEURAL NETWORKS

Learning in Neural Networks

Lecture 6

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

LEARNING TASKS

Simon Haykin "Neural Networks, A Comprehensive Foundation" , Prentice Hall

- The choice of a particular learning algorithm is influenced by the learning task that a neural network is required to perform.

- Some learing tasks that a neural network can perform are:

  1) Pattern Association
  2) Pattern Recognition
  3) Function Approximation
  4) Control

1) PATTERN ASSOCIATION

- An *associative memory* is a brainlike distributed memory that learns by association.
- Association takes one of two forms :

  - Autoassociation - Heteroassociation

Autoassociation : A neural network is required to store a set of patterns (vectors) by repeatedly presenting them to the network. The network is then presented a partial description or distorted (noisy) version of an original pattern stored in it, and the task is to retrieve (recall) that particular pattern.

Heteroassociation : Differs from autoassociation in that an arbitrary set of input patterns is paired with another arbitrary set of output patterns.

x_k : key pattern (vector) applied to an associative memory

y_k : memorized pattern (vector)

The pattern association performed by the network is described by :

$$\boxed{x_k \longrightarrow y_k \text{ , } k = 1,2,..,q.}$$

q : *Number of patterns stored in the network.(gives the storage capacity of the memory)*

In Autoassociative Memory :

$$\boxed{x_k = y_k}$$

✓ *Input Space* and *Output Space* of the network have the same dimensionality.

In Heteroassociative Memory :

$$\boxed{x_k \neq y_k}$$

✓ The dimensionality of the *Output Space* may or may not equal the dimensionality of the *Input Space*.

- Operation of an associative memory has two phases :

1) **Storage Phase :** Training of the network in accordance with

$$x_k \longrightarrow y_k \;, k = 1,2,..,q.$$

2) **Recall Phase :** Retrieval of a memorized pattern in response the presentation of a noisy version of a key pattern to the network.

- Let the input x represent a noisy version of a key pattern x_j.
- This stimulus produces a response (output) y.

- For perfect recall $\boxed{y = y_j}$

  y_j : *Memorized pattern associated with the key pattern* x_j

- When $y \neq y_j$ for $x = x_j$

  We can say memory made an error in recall.

MEMORY

- The task of pattern association leads us naturally to think about memory.
- *Memory* and *learning* are intricately connected. When a particular activity pattern is learned, it is stored in the brain where it can be recalled later when required.
- An *associative memory* is a brainlike distributed memory that learns by association.

- An *associative memory* offers the following characteristics :

- The memory is distributed.
- Both the stimulus (key pattern) and the response (stored pattern) of an associative memory consist of data vectors.
- Information is stored in memory by setting up a spatial pattern of neural activities across a large number of neurons.
- Information contained in a stimulus not only determines its storage location in memory but also an address for its retrieval.

- Learning from the association between the patterns x_k and y_k :

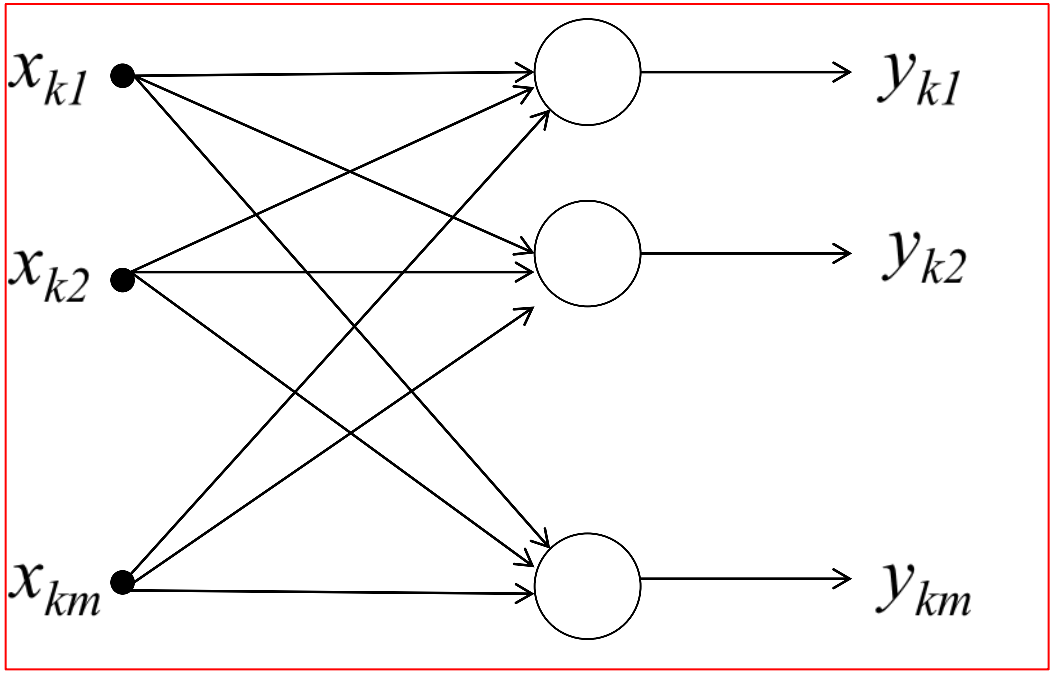
$$x_k = [x_{k1}\, x_{k2} \qquad x_{km}]^T$$
$$y_k = [y_{k1}\, y_{k2} \qquad y_{km}]^T$$

Same dimensionality

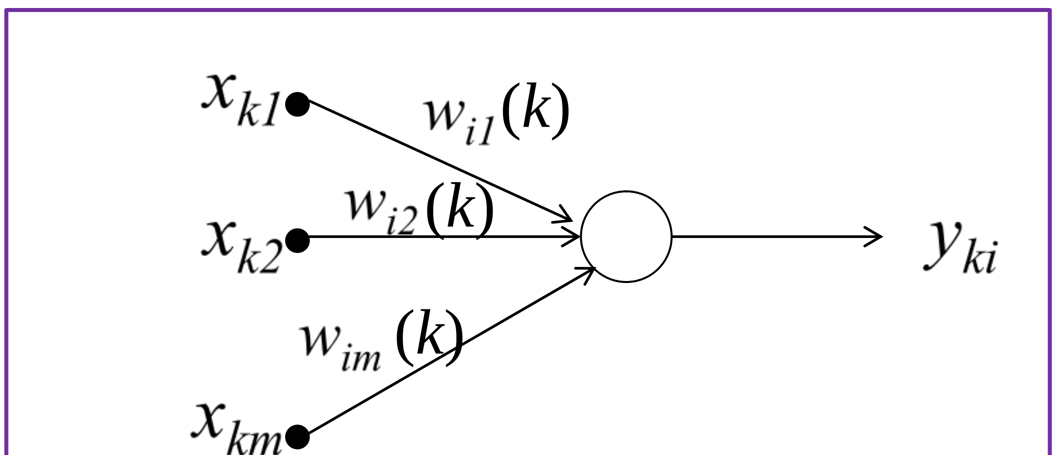
- ✓ *m : Number of source nodes in the input layer*
  Neurons in the output layer

$$x_k \longrightarrow y_k$$

Key patterns — Stored patterns



- NN diagram for association between the patterns x and y
- k : pattern
- m : components of pattern (number of nodes)



- NN model for i. neuron

2) PATTERN RECOGNITION

- Humans perform pattern recognition through a learning process. Humans can

  * Recognize the familiar face of a person even though that person has aged

  * Identify a familiar person by his/her voice on the telephone despite a bad connection

  * Distinguish a boiled egg that is good from a bad one by smelling it.

- Pattern recognition is formally defined as the process whereby a received pattern/signal is assigned to one of a prescribed number of classes (categories).

- A neural network performs pattern recognition by **first** undergoing a training session, during which the network is repeatedly presented a set of input patterns along with the category to which each particular pattern belongs.
- **Later**, a new pattern is presented to the network that has not been seen before, but which belongs to the same population of patterns used to train the network.
- **Finally**, the network is able to identify the class of that particular pattern because of the information it has extracted from the training data.

- Patterns are represented by points in a multidimensional “*decision space*”.

- The decision space is divided into “*regions*”.

- Each region is associated with a “*class*”.

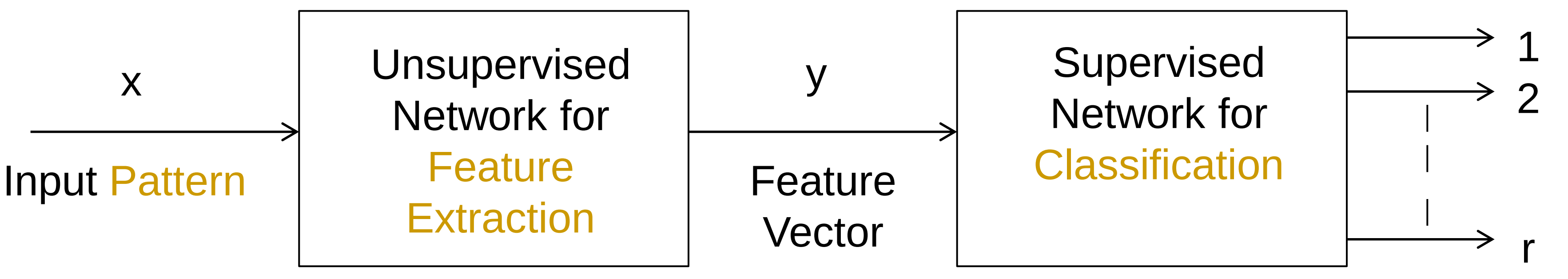
- The “*decision boundaries*” are determined by the training process.

- Pattern recognition machines using neural networks may take one of two forms :

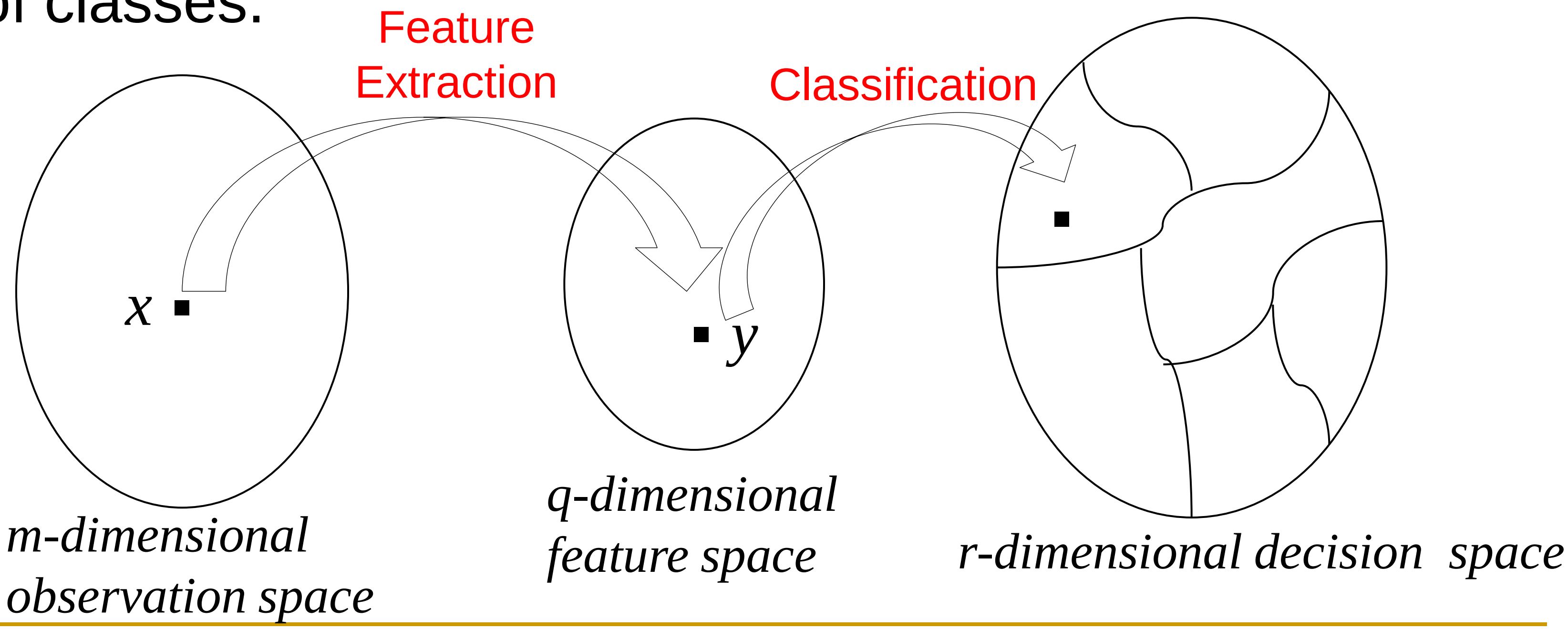
- ✓ *Which of these two approaches is adapted in practice depends on the application of interest.*

1) The machine is designed as a single multilayer feedforward network using supervised learning algorithm. The task of feature extraction is performed by the computational units in the hidden layers of the network.

2) The machine is split into two parts :
An unsupervised network for feature extraction
A supervised network for classification



- Pattern is represented by a set of (m) observables.
- Feature extraction is described by a transformation that maps the point x, into an intermediate point y in a (q) dimensional feture space with q<m.
- The classification is described as a transformation that maps the intermediate point y into one of the classes in an r-dimensional decision space, where (r) is the number of classes.



3) FUNCTION APPROXIMATION

- Consider a nonlinear input-output mapping described by the functional relationship :

$$\boxed{d = f(x)} \quad (\, x : Input \,, \quad d : Output \,)$$

- The vector-valued function $f(.)$ is assumed to be unknown.
- To make up for the lack of knowledge about the function $f(.)$, we are given the set of labeled examples :

$$\mathrm{T} = \{(x_i, d_i)\}, \quad i = 1,2,..,N.$$

- The requirement is to design a neural network that approximates the unknown function $f\,(.)$ such that the function $F(.)$ describing the input-output mapping actually realized by the network is close enough to $f\,(.)$ in a Euclidean sense over all inputs, as shown by

$$\boxed{\| F(x) - f(x) \| < \epsilon}$$

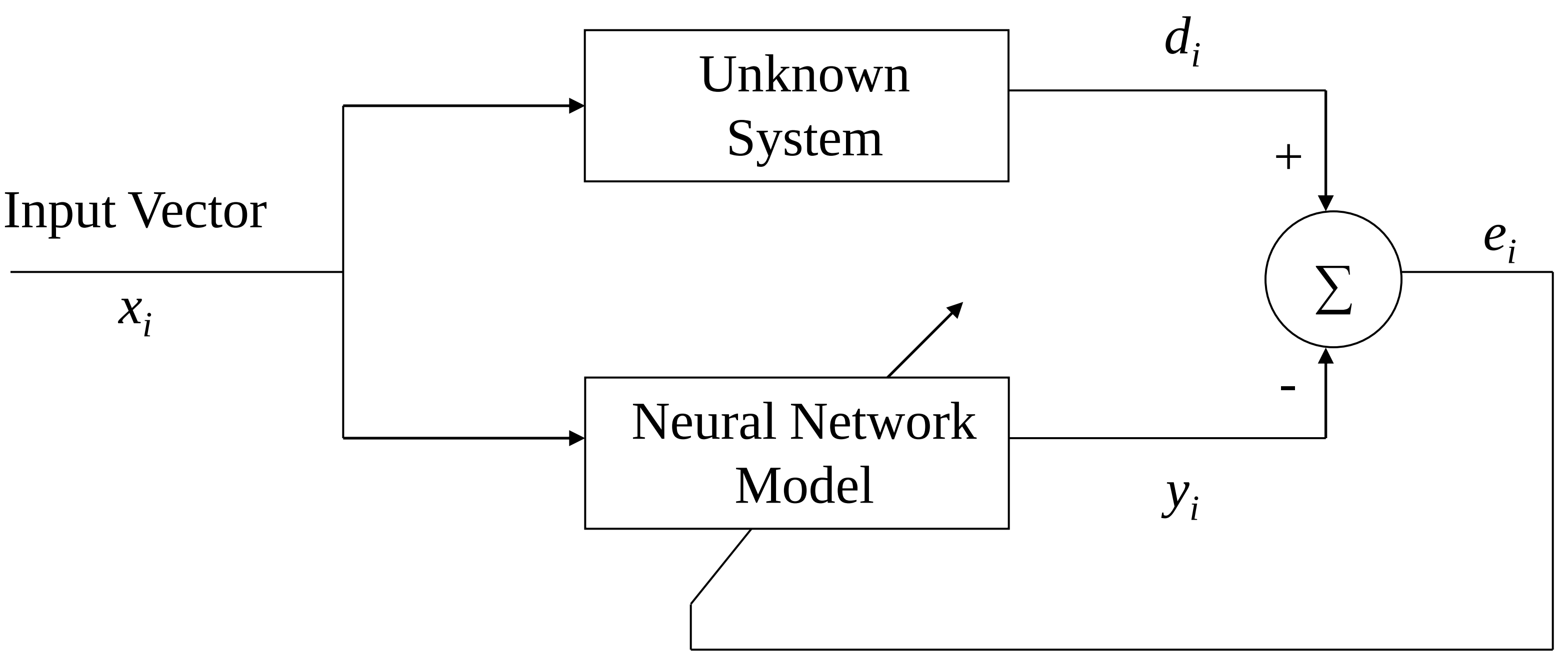
ϵ : *small positive number*

- The approximation problem here is a perfect candidate for supervised learning.
- The ability of a NN to approximate an unknown input-output mapping may be exploited in two important ways :

1) System Identification :

- ✓ Let $d = f(x)$ describe the input-output relation of an unknown memoryless (time invariant) multiple input-multiple output (MIMO) system.
- ✓ T = $\{(x_i, d_i)\}$ is used to train a neural network
- ✓ y_i denote the output of the neural network produced in response to an input vector x_i.

- ✓ The difference between d_i and y_i provides the error signal vector e_i.
- ✓ This error signal is used to adjust the free parameters of the network.

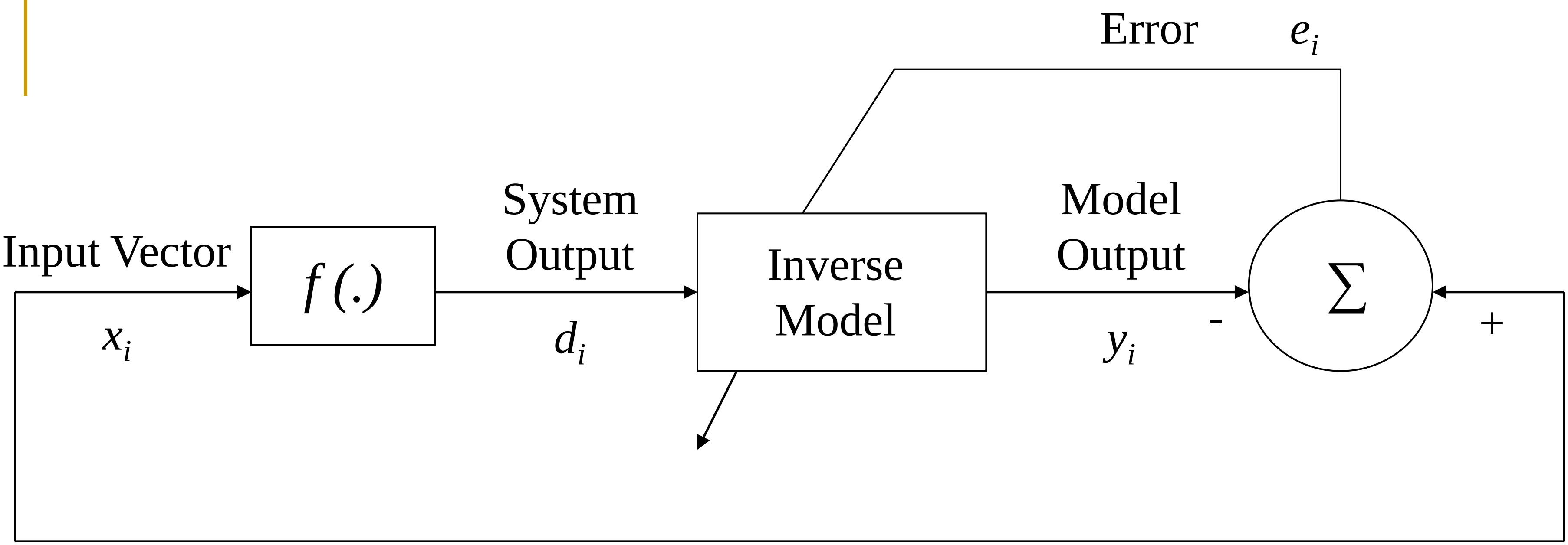


- Block Diagram of System Identification

2) Inverse System :

- Let $d = f(x)$ describe the input-output relation of an known memoryless (time invariant) multiple input-multiple output (MIMO) system.
- The requirement is to construct an inverse system that produces the vector x in response to the vector d.

$$\textit{Inverse System}: \quad \boxed{x = f^{-1}(d)}$$



- Block Diagram of Inverse System

x_i : *Desired Response*

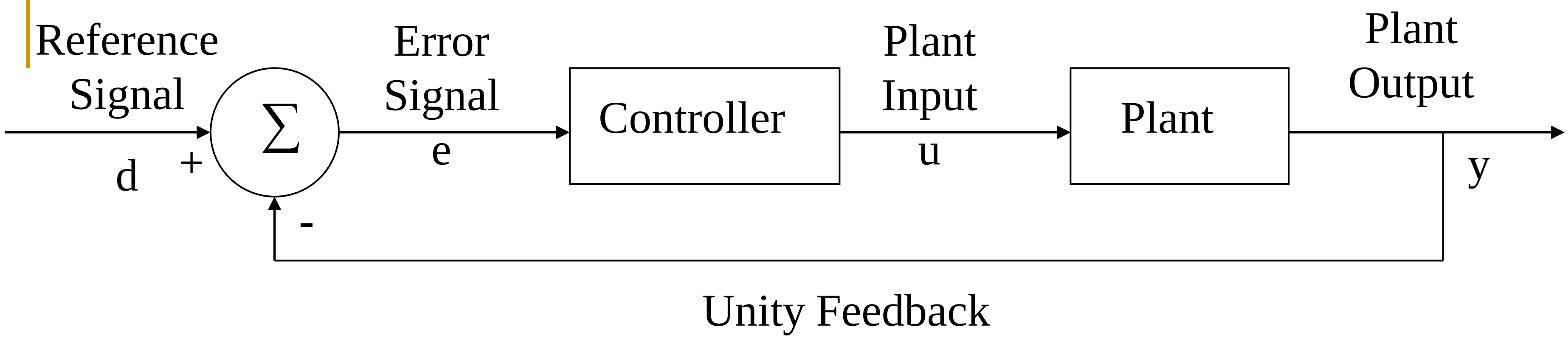
d_i : *Input*

e_i : *Difference between* x_i *and output of the neural network produced in response to* d_i

4) CONTROL

- The *control of a plant* is another learning task that can be performed by a neural network.

Plant : Process or a critical part of a system that is to be maintained in a controlled condition.



> ***Consider the above feedback control system :***

- ✓ Plant Output is fed back directly to the input
- ✓ Thus, the plant output y is subtracted from a reference signal d supplied from an external source
- ✓ The error signal so produced is applied to a neural controller for the purpose of adjusting its free parameters
- ✓ The primary objective of the controller is to supply appropriate inputs to the plant to make its output y track the reference signal d.