Unit 1: Introduction[7hrs]

Unit 1: Soft Computing :

- Lect1: Introduction of soft computing, soft computing vs. hard computing, Various types of soft computing techniques, applications of soft computing
- Lect2: Characteristics of Neural Networks, Structure and Working of a biological neural network
- Lect3: Artificial Neural Network Teminology, Models of neurons: MP model,

- Lect4: Models of neurons: Perceptron model, Adaline model, Topology,
- Lect5: Basic Learning laws, What is learning, supervised and unsupervised learning,
- Lect6: Functional Units of ANN for pattern recognition task- Pattern Recognition Problem,
- Lect7: Basic functional units, Content Beyond Syllabus, Unit Test & Feedback

Lecture 1:

- Introduction of soft computing,
- Soft computing vs. Hard computing,
- Various types of soft computing techniques,
- Applications of soft computing

Introduction of soft computing

- Soft computing differs from conventional (hard) computing in that, unlike hard computing, it is tolerant of imprecision, uncertainty, partial truth, and approximation. In effect, the role model for soft computing is the human mind.

What is Hard Computing?

- Hard computing, i.e., conventional computing, requires a precisely stated analytical model and often a lot of computation time.
- Many analytical models are valid for ideal cases.
- Real world problems exist in a non-ideal environment.

Introduction of soft computing

- The principal constituents, i.e., tools, techniques, of Soft Computing (SC) are
  - Fuzzy Logic (FL), Neural Networks (NN), Support Vector Machines (SVM), Evolutionary Computation (EC), and
  - Machine Learning (ML) and Probabilistic Reasoning (PR)

Premises of Soft Computing

- The real world problems are pervasively imprecise and uncertain
- Precision and certainty carry a cost

Guiding Principles of Soft Computing

- The guiding principle of soft computing is:
  - Exploit the tolerance for imprecision, uncertainty, partial truth, and approximation to achieve tractability, robustness and low solution cost.

Unique Property of Soft computing

- Learning from experimental data.
- Soft computing techniques derive their power of generalization from approximating or interpolating to produce outputs from previously unseen inputs by using outputs from previous learned inputs.
- Generalization is usually done in a high dimensional space.

Applications using Soft Computing

- Application of soft computing to handwriting recognition
- Application of soft computing to automotive systems and manufacturing
- Application of soft computing to image processing and data compression
- Application of soft computing to architecture

Applications using Soft Computing

- Application of soft computing to decision-support systems
- Application of soft computing to power systems
- Neurofuzzy systems
- Fuzzy logic control

Soft computing vs. Hard computing

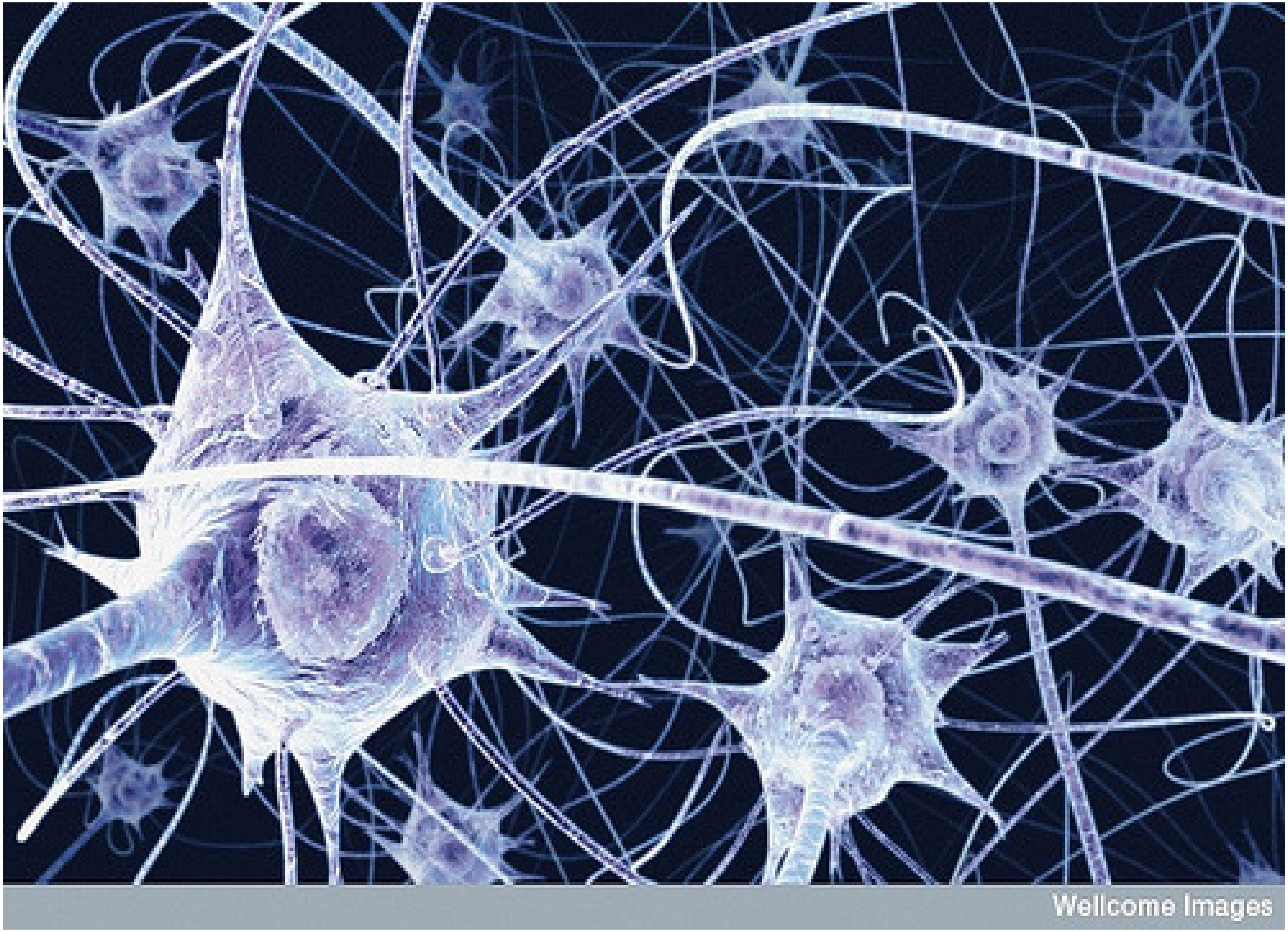
Sr. no	Hard Computing	Soft Computing
1	Hard computing, i.e., conventional computing requires a precisely stated analytical model and often a lot of computation time.	it is tolerant of imprecision, uncertainty, partial truth, and approximation.
2	based on binary logic, crisp systems, numerical analysis and crisp software	based on fuzzy logic, neural nets and probabilistic reasoning
3	has the characteristics of precision and categoricity	has the characteristics of approximation and dispositionality

Soft computing vs. Hard computing

Sr. no	Hard Computing	Soft Computing
4	requires programs to be written	can evolve its own programs
5	uses two-valued logic	can use multivalued or fuzzy logic
6	Hard computing is deterministic	soft computing incorporates stochasticity
7	requires exact input data	can deal with ambiguous and noisy data
8	is strictly sequential	allows parallel computations
9	produces precise answers	can yield approximate answers

Lecture 2:

- Characteristics of Neural Networks, Structure and Working of a biological neural network


Wellcome Images

Characteristics of Neural Networks

- Def:
  - A neural network is a system composed of many simple processing elements operating in parallel whose function is determined by network structure, connection strengths, and the processing performed at computing elements or nodes

Characteristics:

- Robustness and fault tolerance: The decay of nerve cells does not seem to affect the performance of the network significantly.
- Flexibility: The network automatically adjusts to a new environment without using any preprogrammed instruction set.

Characteristics:

- Ability to deal with a variety of data situations: The network can deal with information that is fuzzy, probabilistic, noisy or inconsistent.
- Collective computation - The network can routinely perform many operations in parallel and also a given task in a distributed manner.

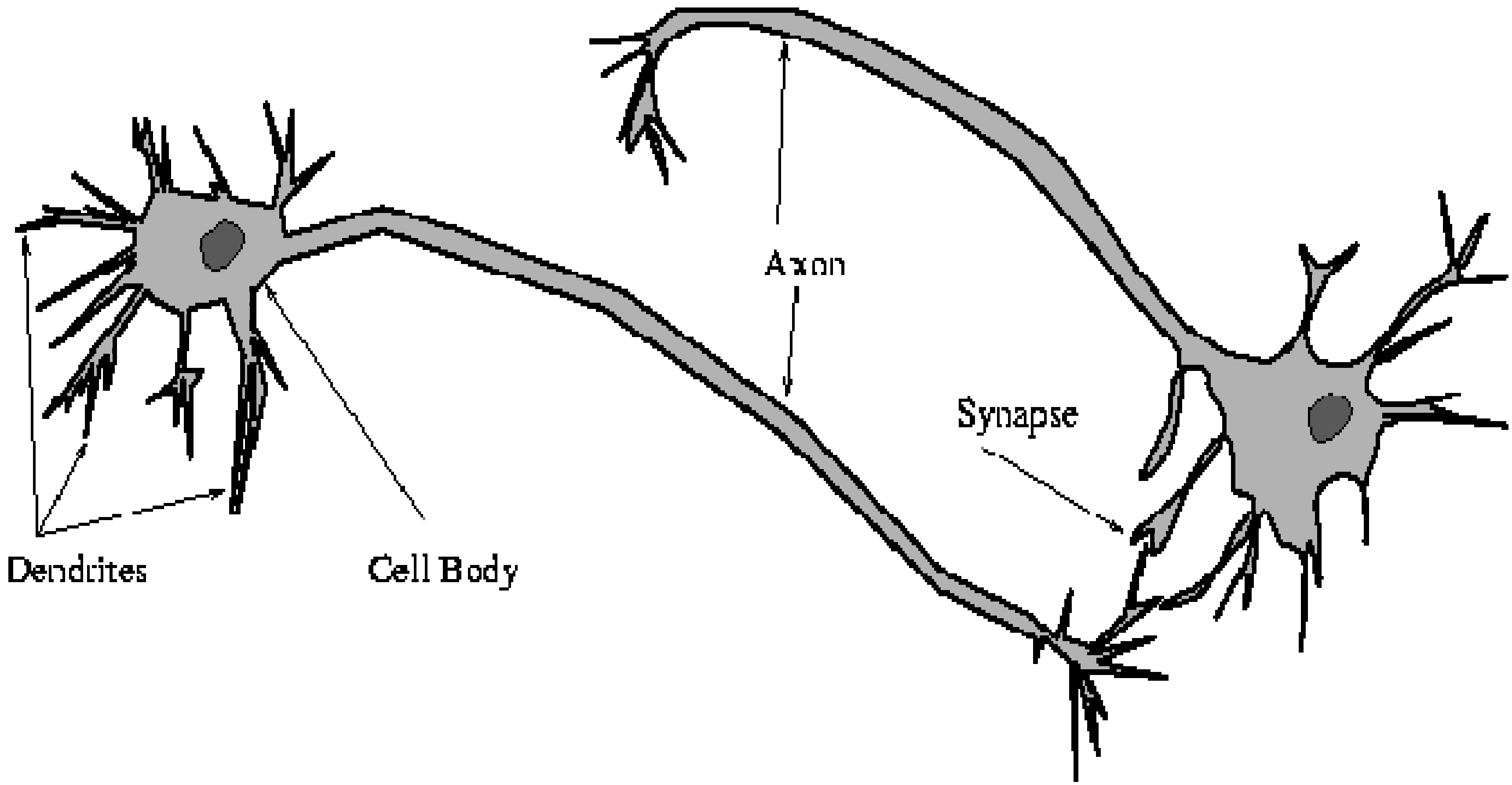
Biological Overview

- Most living creatures, which have the ability to adapt to a changing environment, need a controlling unit which is able to learn.
- Higher developed animals and humans use very complex networks of highly specialized neurons to perform this task.
- The control unit - or brain - can be divided in different anatomic and functional sub-units, each having certain tasks like vision, hearing, motor and sensor control.

- The brain is connected by nerves to the sensors and actors in the rest of the body.
- The neuron contains all structures of an animal cell.
- The complexity of the structure and of the processes in a simple cell is enormous.
- Even the most sophisticated neuron models in artificial neural networks seem comparatively toy-like.
- Structurally the neuron can be divided in three major parts: the cell body (soma), the dentrites, and the axon,

...

- ...



Axon
Synapse
Dendrites
Cell Body

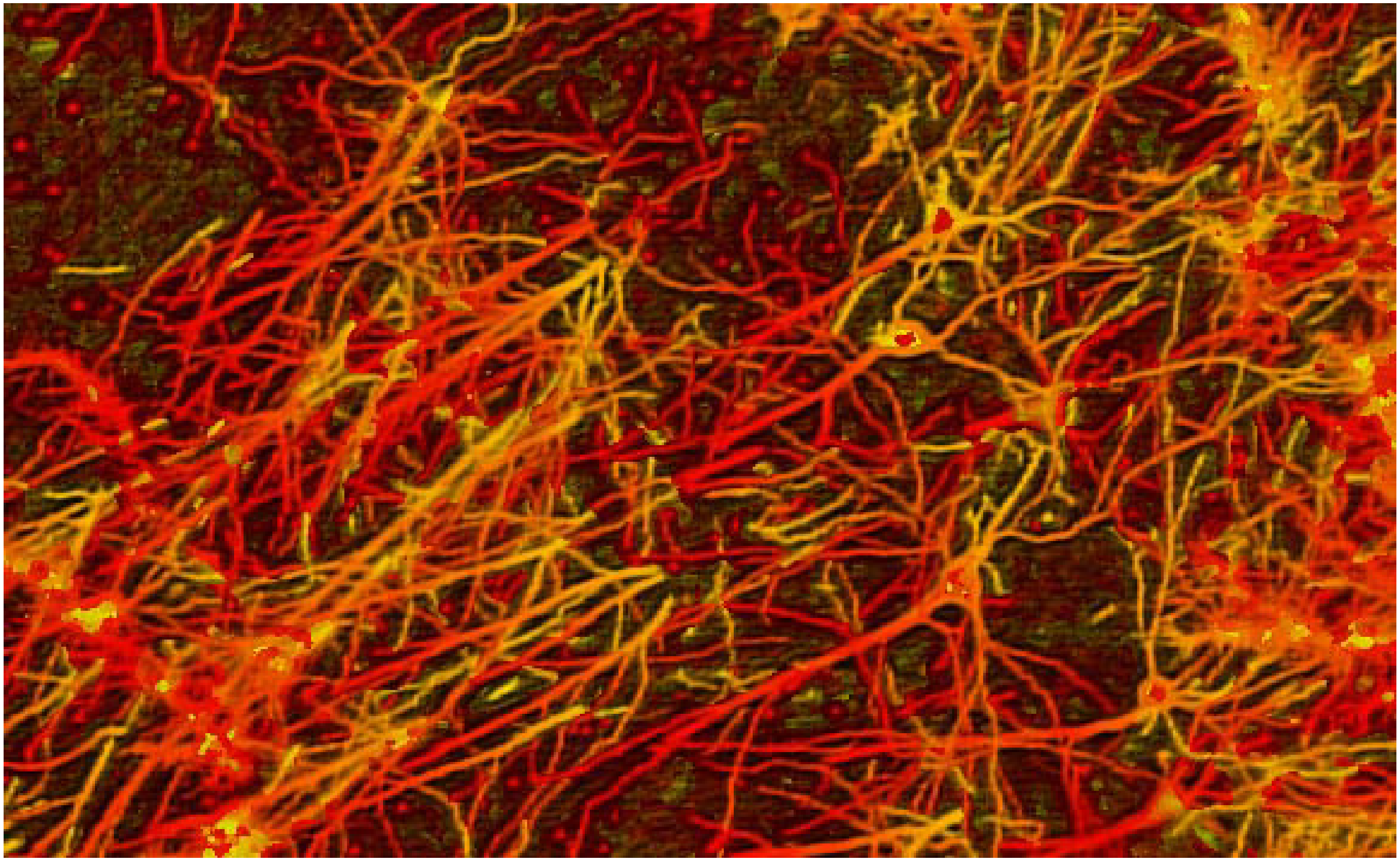
Structure and Working of a biological neural network

- The brain consists of a very large number of neurons, about 10^{11} in average.
- These can be seen as the basic building bricks for the central nervous system (CNS).
- The neurons are interconnected at points called synapses.
- The complexity of the brain is due to the massive number of highly interconnected simple units working in parallel, with an individual neuron receiving input from up to 10000 others.

- Each neuron is a cell (Figure 1) that uses biochemical reactions to receive, process, and transmit information.
- Treelike networks of nerve fibers called *dendrites are* connected to the cell body or soma, where the cell nucleus is located.
- Extending from the cell body is a single long fiber called the *axon, which eventually branches into strands* and substrands, and are connected to other neurons through synaptic terminals or synapses.

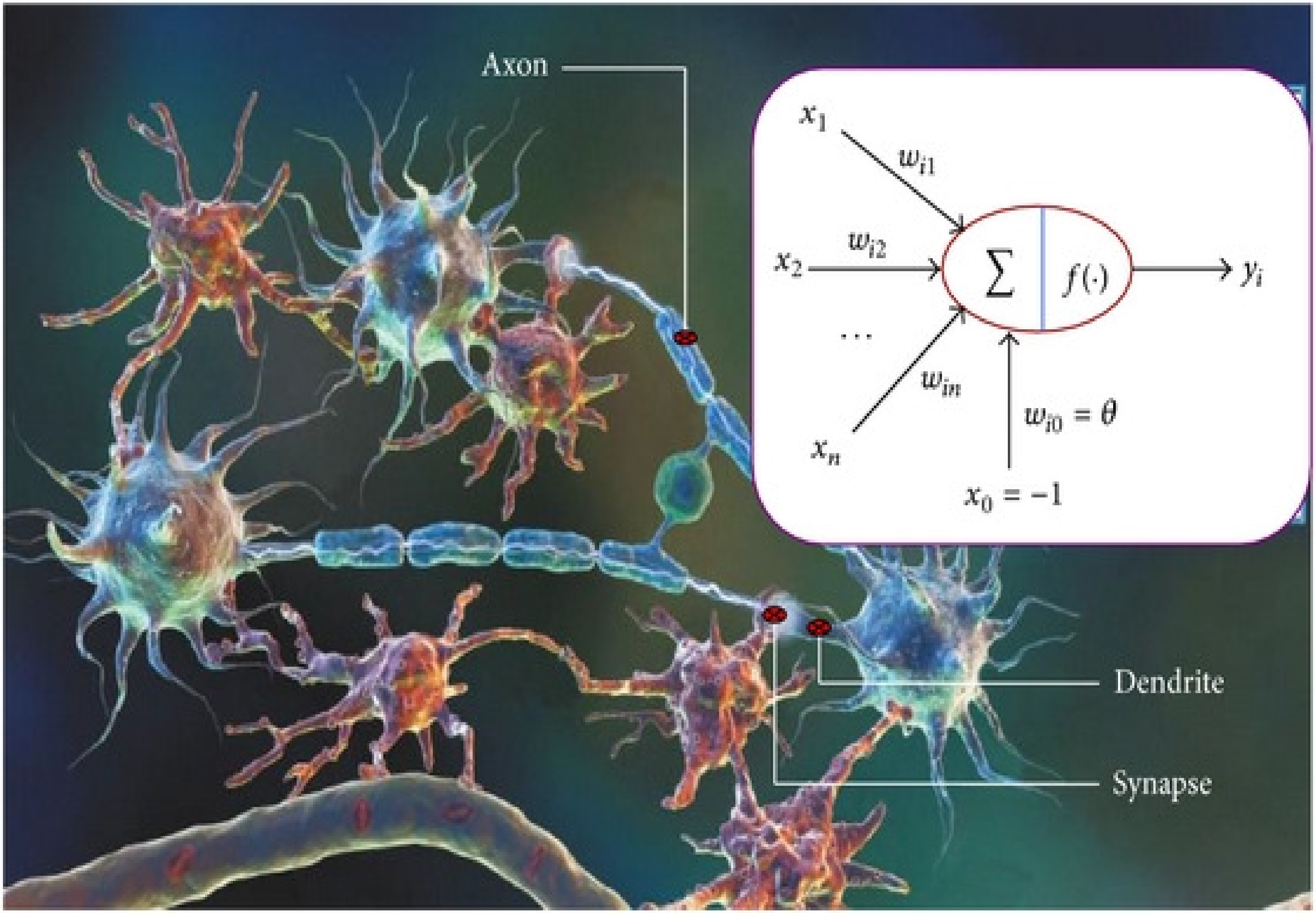
- The transmission of signals from one neuron to another at synapses is a complex chemical process in which specific transmitter substances are released from the sending end of the junction.
- The effect is to raise or lower the electrical potential inside the body of the receiving cell.
- If the potential reaches a threshold, a pulse is sent down the axon and the cell is 'fired'.

INTERCONNECTIONS IN BRAIN



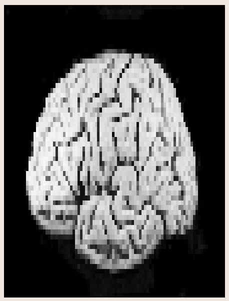
…

- Artificial neural networks (ANN) have been developed as generalizations of mathematical models of biological nervous systems.


Axon
x_1
w_{i1}
x_2
w_{i2}
Σ
$f(\cdot)$
y_i
$\cdots$
w_{in}
$w_{i0} = \theta$
x_n
$x_0 = -1$
Dendrite
Synapse

Performance comparison:

- Speed
- Processing
- Size & complexity
- storage:
- Fault tolerance
- Control mechanism

	processing elements	element size	energy use	processing speed	style of computation	fault tolerant	learns	intelligent, conscious
	10^{14} synapses	10^{-6} m	30 W	100 Hz	parallel, distributed	yes	yes	usually
	10^{8} transistors	10^{-6} m	30 W (CPU)	10^{9} Hz	serial, centralized	no	a little	not (yet)

Lecture: 3

- Artificial Neural Network Terminology,
- Models of neurons: MP model

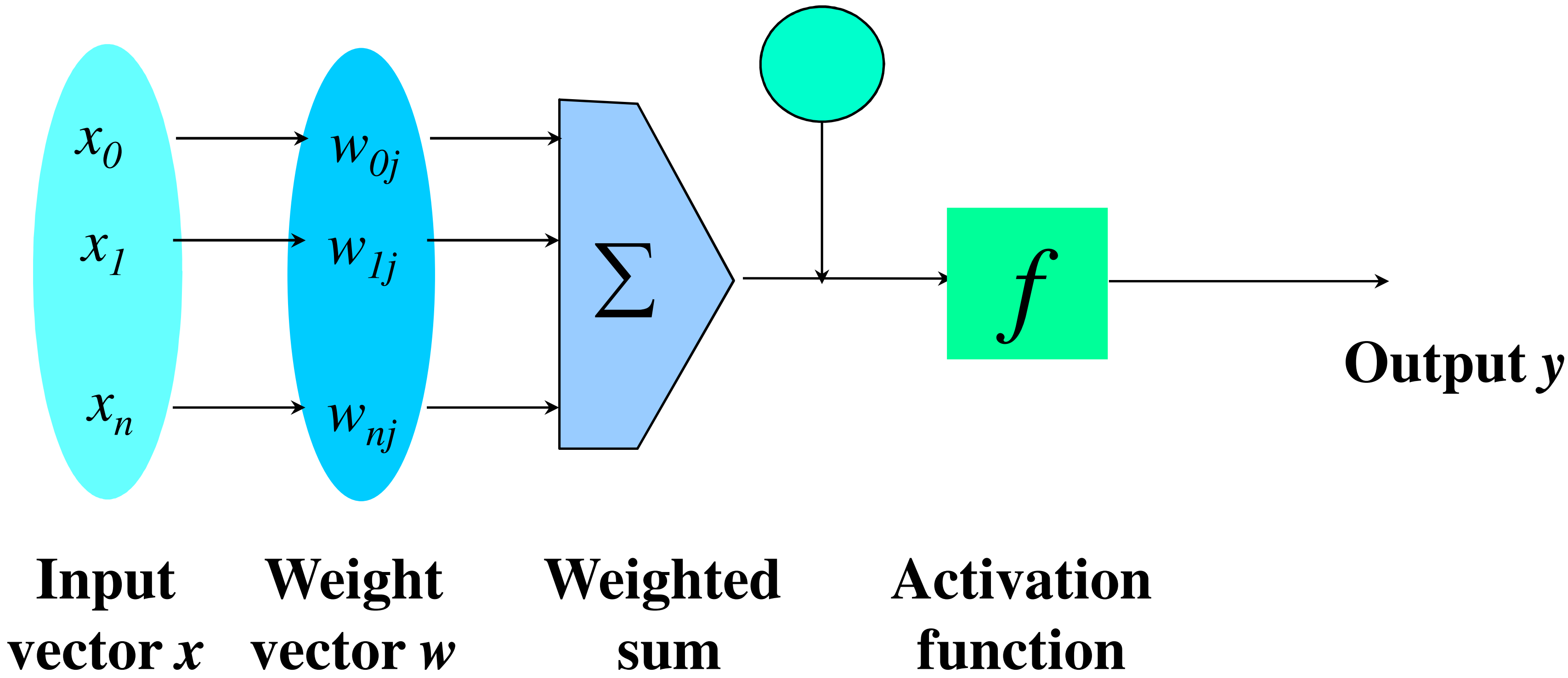

x_0
x_1
x_n
w_{0j}
w_{1j}
w_{nj}
Σ
f
Output *y*
Input vector *x*
Weight vector *w*
Weighted sum
Activation function

Fig. 1.3.1

Artificial Neural Network Terminology

- It is a computational system inspired by the
  - Structure
  - Processing Method
  - Learning Ability

  of a biological brain

Artificial Neural Network Terminology

- Characteristics of Artificial Neural Networks
  - A large number of very simple processing neuron-like processing elements
  - A large number of weighted connections between the elements
  - Distributed representation of knowledge over the connections
  - Knowledge is acquired by network through a learning process

Terminology

- Processing units:
  - artificial neural network (ANN) is a highly simplified model of the structure of the biological neural network.
  - An ANN consists of interconnected processing units.
  - The general model of a processing unit consists of a summing part followed by an output part.

Terminology

- Processing units:
  - The summing part receives n input values, weighs each value, and performs a weighted sum. The weighted sum is called the activation value.
  - The sign of the weight for each input determines whether the input is excitatory (positive weight) or inhibitory (negative weight).

Terminology

- Processing units:
  - The inputs could be discrete or continuous data values, and likewise the outputs also could be discrete or continuous.
  - The input and output may also be viewed as deterministic or stochastic or fuzzy, depending on the nature of the problem and its solution.

Interconnections

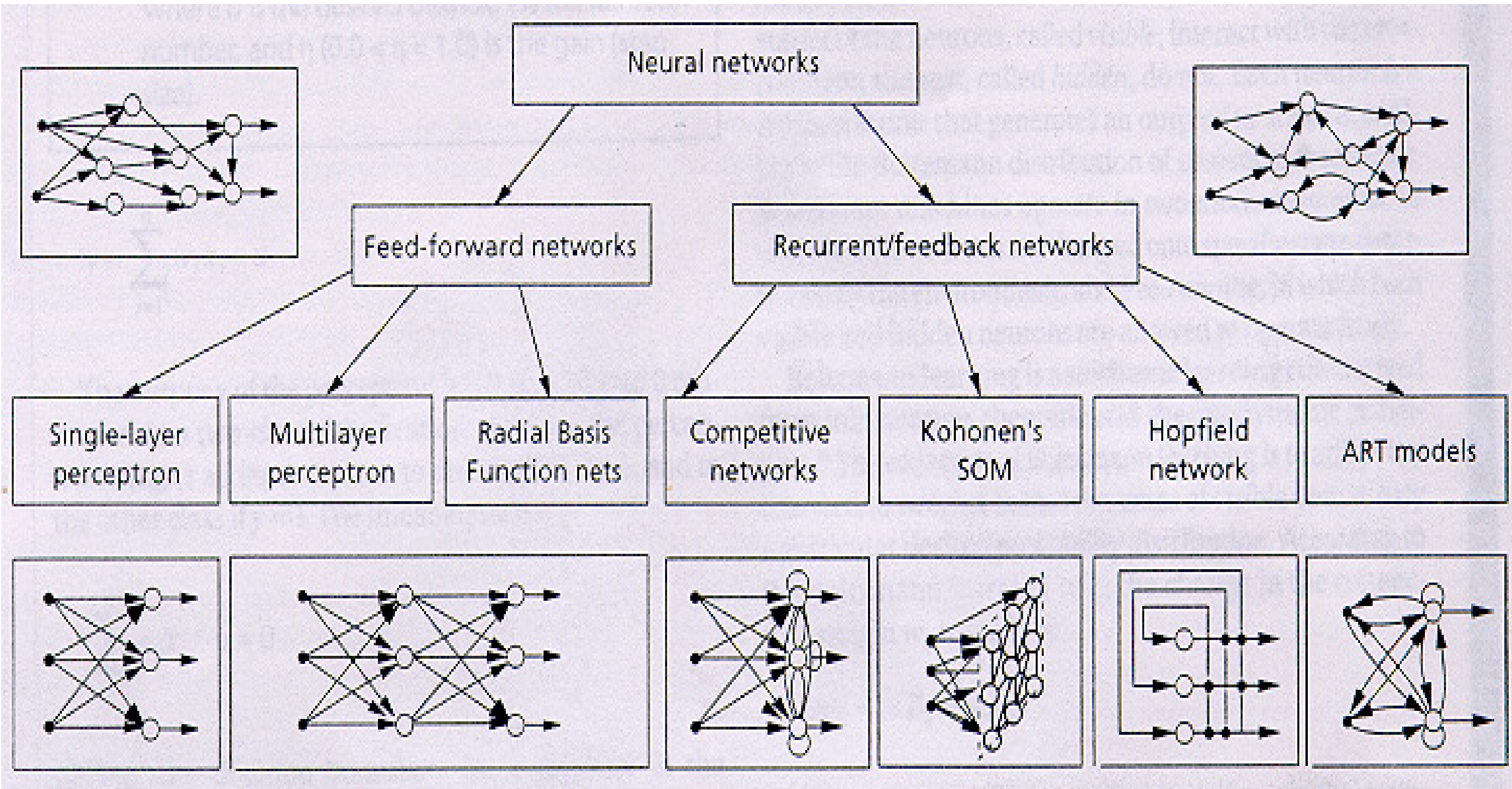


Fig. 1.3.2

Terminology: Interconnections

- In an ANN several processing units are interconnected according to some topology to accomplish a pattern recognition task.
- Therefore the inputs to a processing unit may come from outputs of other processing units, and/or from an external source.
- The output of each unit may be given to several units including itself.

Terminology: Interconnections

- The amount of the output of one unit received by another unit depends on the strength of the connection between the units, and it is reflected in the weight value associated with the connecting link.
- If there are N units in a given ANN then at any instant of time each unit will have a unique activation value and a unique output value.
- The set of the N activation values of the network defines the activation state of the network at that instant.

Terminology: Operations

- In operation, each unit of an ANN receives inputs from other connected units and/or from an external source.
- A weighted sum of the inputs is computed at a given instant of time.
- The resulting activation value determines the actual output from the output function unit, i.e., the output state of the unit.

Terminology: Operations

- The output values and other external inputs in turn determine the activation and output states of the other units.
- The activation values of the units (activation state) of the network as a function of time are referred to as activation dynamics.

Terminology: Operations

- The activation dynamics also determine the dynamics of the output state of the network.
- The set of all activation states defines the state space of the network.
- The set of all output states defines the output or signal state space of the network.
- Activation dynamics determines the trajectory of the path of the states in the state space of the network.

Terminology: Operations

- For a given network, defined by the units and their interconnections with appropriate weights, the activation states refer to the short term memory function of the network.
- Generally the activation dynamics is followed to recall a pattern stored in a network.

Terminology: Operations

- In order to store a pattern in a network, it is necessary to adjust the weights of the network.
- The sets of all weight values (corresponding to strengths of all connecting links of an ANN) defines the weight space.
- If the weights are changing, then the set of weight values as a function of time defines the synaptic dynamics of the network.

Terminology: Operations

- Synaptic dynamics is followed to adjust the weights in order to store given patterns in the network.
- The process of adjusting the weights is referred to as learning.
- Once the learning process is completed, the final set of weight values corresponds to the long term memory function of the network.
- The procedure to incrementally update each of the weights is called a learning law or learning algorithm.

Terminology: Update

- In implementation, there are several options available for both activation and synaptic dynamics.
- In particular, the updating of the output states of all units could be performed synchronously.
- In this case, the activation values of all units are computed at the same time assuming a given output state throughout.
- From these activation values the new output state of the network is derived.

Terminology: Update

- In an asynchronous update, on the other hand, each unit is updated sequentially, taking the current output state of the network into account each time.
- For each unit, the output state can be determined from the activation value either deterministically or stochastically.

Terminology: Update

- In practice, the activation dynamics, including the update, is much more complex in a biological neural network.
- The ANN models along with the equations governing the activation and synaptic dynamics are developed according to the complexity of the pattern recognition task to be handled.

Terminology:

- Bias
- Threshold
- Learning rate

Models of neurons: MP model

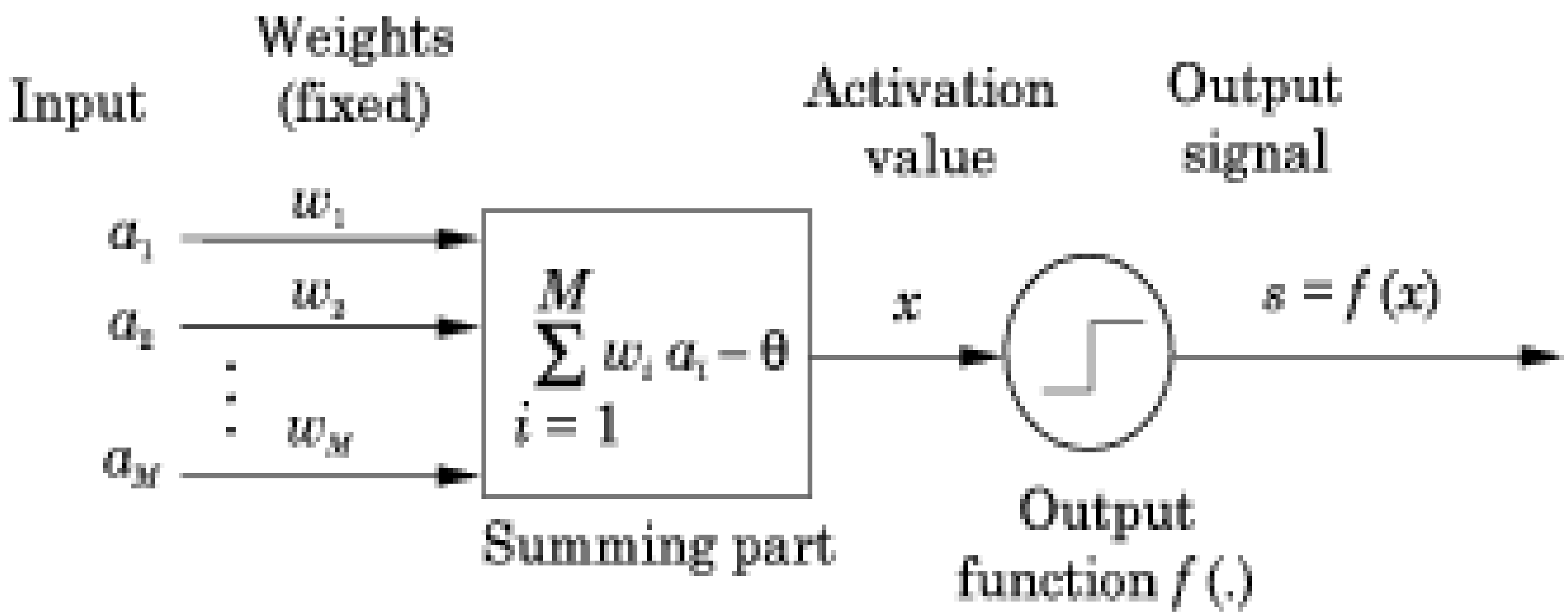


Fig. 1.3.3

McCulloch-Pitts model

- In the McCulloch-Pitts (MP) model (fig. 1.3.3) the activation (x) is given by a weighted sum of its n-input signal values {a_i) and a bias term.
- The activation could have an additional absolute inhibition term, which can prevent excitation of the neuron.
- The output signal (s) is typically a nonlinear function of the activation value.
- Three common nonlinear functions (binary, ramp and sigmoid) are shown in figure, although the binary function was used in the original MP model.

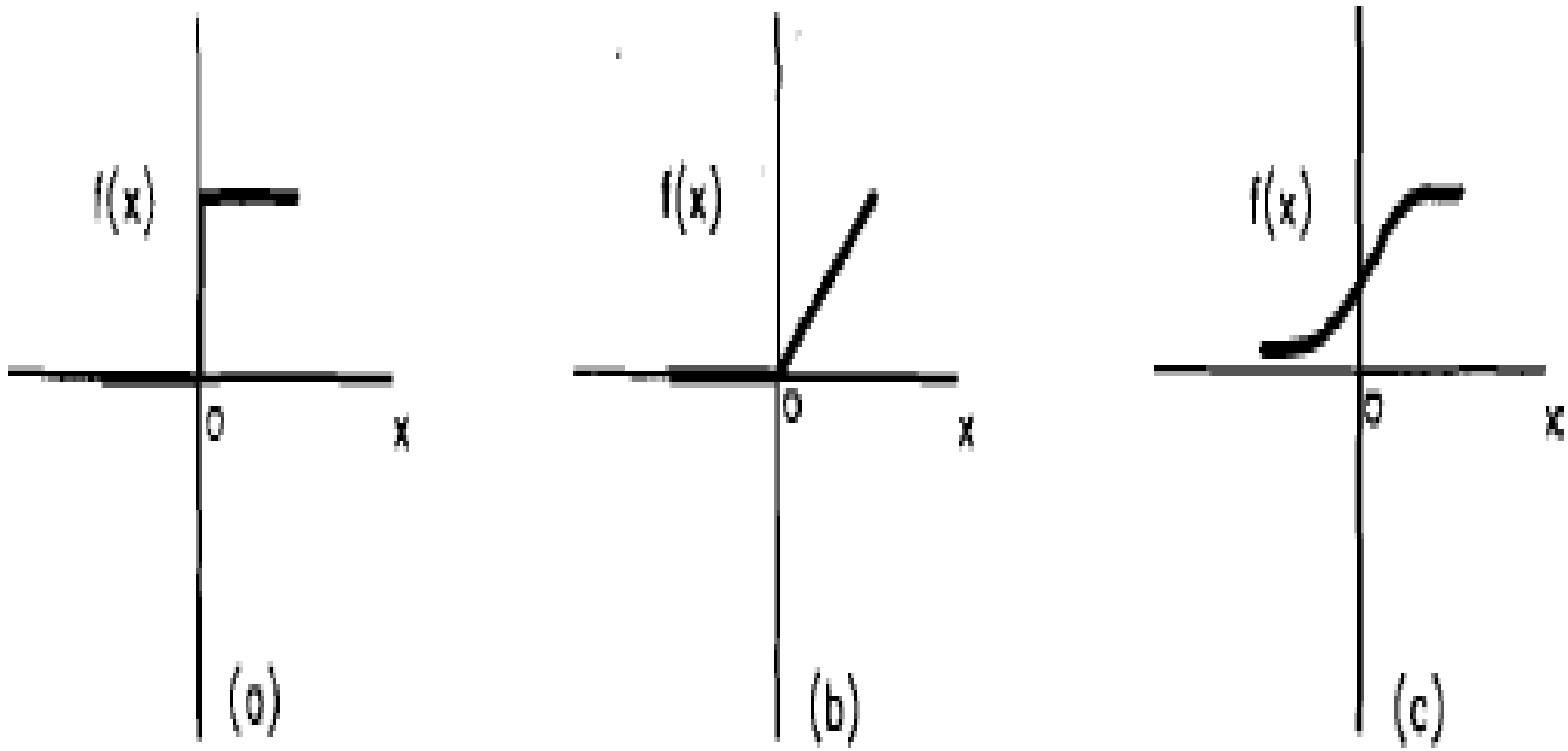


Fig. 1.3.4 **(a) Binary, (b) ramp and (c) sigmoid**

McCulloch-Pitts model

- activation: $x = \sum_{i=1}^{n} w_i a_i - \theta - [\text{inhibition}],$
- output signal: $s = f(x).$
- In this model the weights w_i are constant.
- That means there is no learning.
- Networks consisting of MP neurons with binary (on-off) output signals can be configured to perform several logical functions

Problems:

- Generate AND/OR function using McCulloch-Pitts neural net
- Generate ANDNOT function using McCulloch-Pitts neural net

Solution 01:

- AND function using McCulloch-Pitts neural net

Solution 02:

- ANDNOT function using McCulloch-Pitts neural net

Lecture 4:

- Models of neurons:
  - Perceptron model,
  - Adaline model,
- Topology
- Basic Learning laws

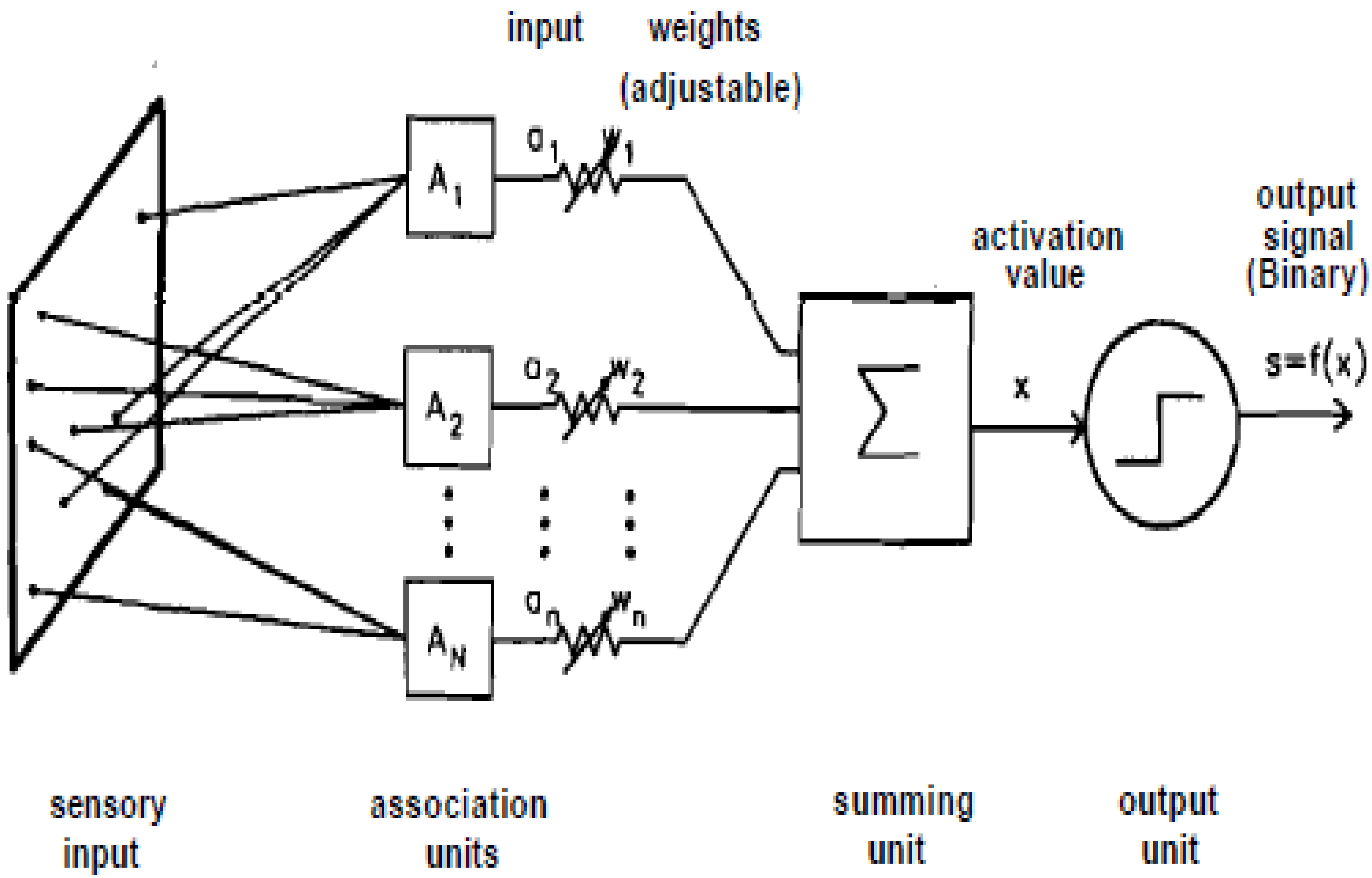


Fig. 1.4.1 **Rosenblatt's model of a neuron**

Perceptron

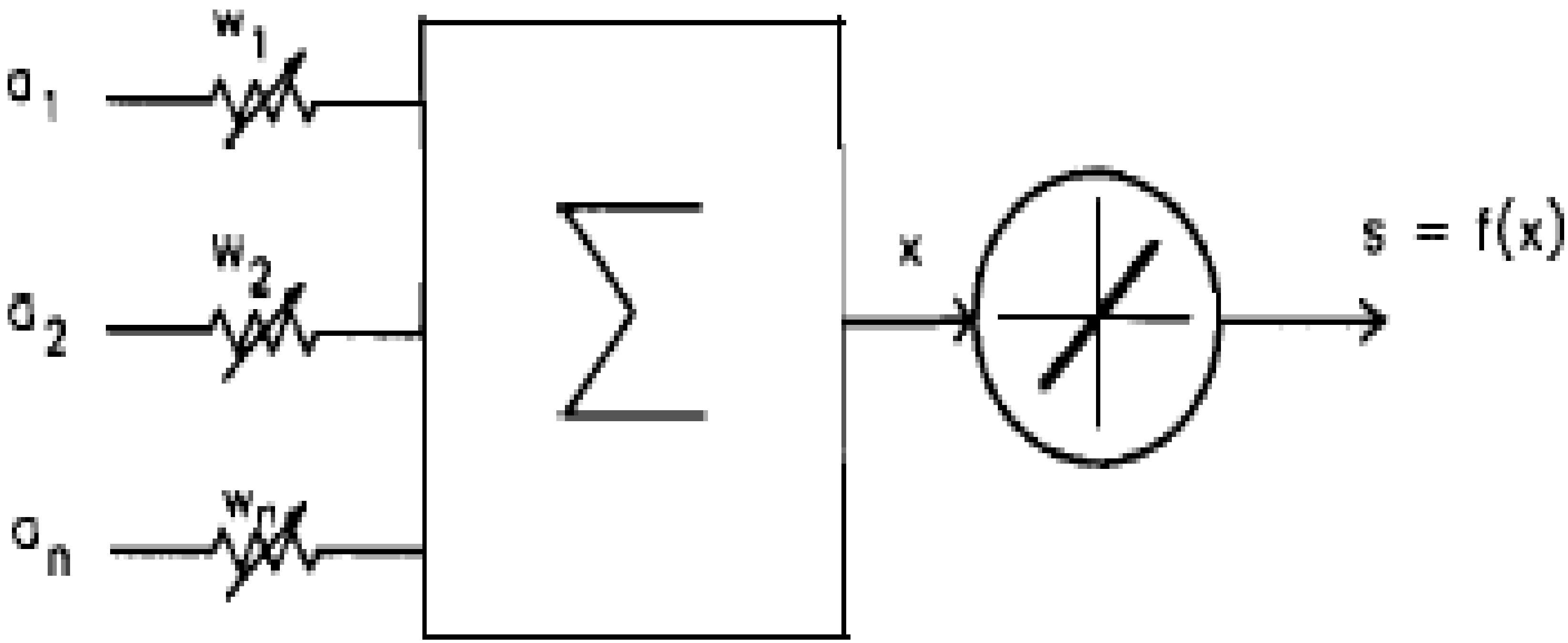
- Rosenblatt's perceptron model (1958)
- It consists of outputs from sensory units to a fixed set of association units, the outputs of which are fed to an MP neuron.
- The association units perform predetermined manipulations on their inputs.
- The main deviation from the MP model is that here learning (i.e., adjustment of weights) is incorporated in the operation of the unit.
- The target output (b) is compared with the actual binary output (s) and the error is used to adjust the weights

Perceptron

- activation: $$x = \sum_{i=1}^{n} w_i a_i - \theta,$$
- output signal: $$s = f(x),$$
- error: $$\delta = b - s,$$
- weight update: $$\frac{dw_i}{dt} = \eta \delta a_i,$$

ADALINE

- Adaptive Linear Element proposed by Widrow



ADALINE

- In the adaline model the analog activation value (x) is compared with the target output (b).
- In other words, the output is a linear function of the activation value (x).

ADALINE

- activation: $$\mathrm{x} = \sum_{i=1}^{n} w_i a_i - \theta,$$
- output signal: $$s = \mathrm{f}(\mathrm{x}) = \mathrm{x},$$
- error: $$\delta = b - s = b - x,$$
- weight update: $$\mathrm{d}w_i/\mathrm{d}t = \eta\delta a_i.$$

Topology of ANN

- The way interconnections and nodes are arranged
- Choice of topology is determined by the problem being considered
- The arrangement of the processing units, connections, and pattern input/output is referred to as topology

Topology of ANN

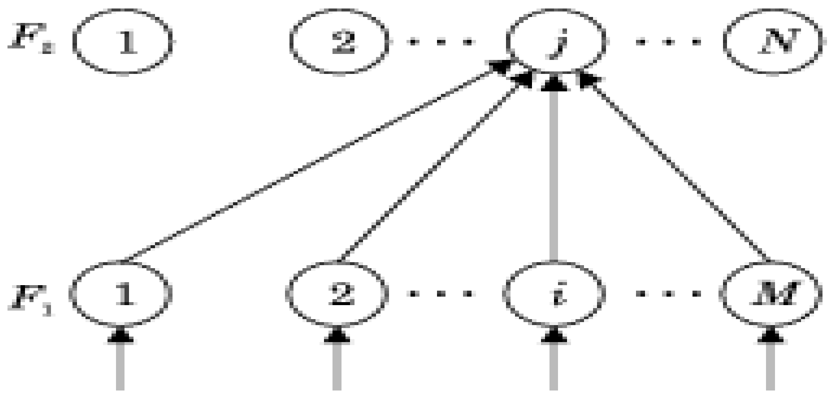
- Artificial neural networks are useful only when the processing units are organized in a suitable manner to accomplish a given pattern recognition task.
- Artificial neural networks are normally organized into layers of processing units.

Topology of ANN

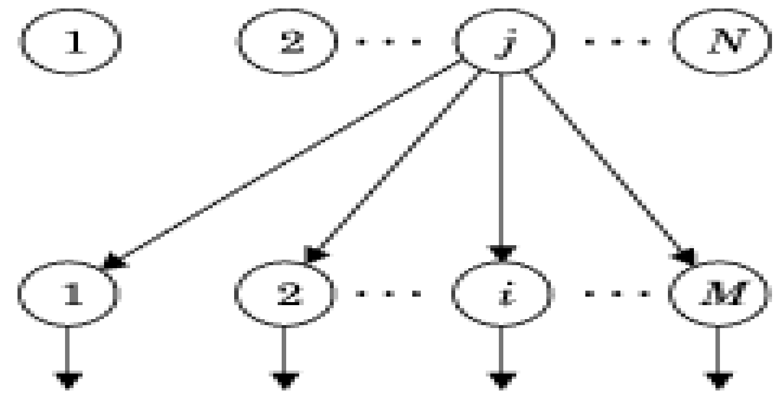
- Connections can be made either from units of one layer to units of another (interlayer connections) or from the units within the layer (intralayer connections) or both inter and intralayer connections.
- Further, the connections among the layers and among the units within a layer can be organized either in a feedforward manner or in a feedback manner.

Topology

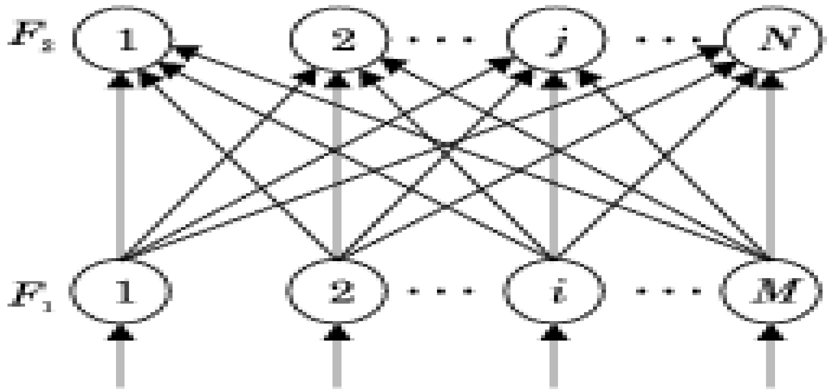
- Instar
- Outstar
- Group of instars
- Group of outstars
- Bidirectional Associative Memory
- Auto-associative Memory



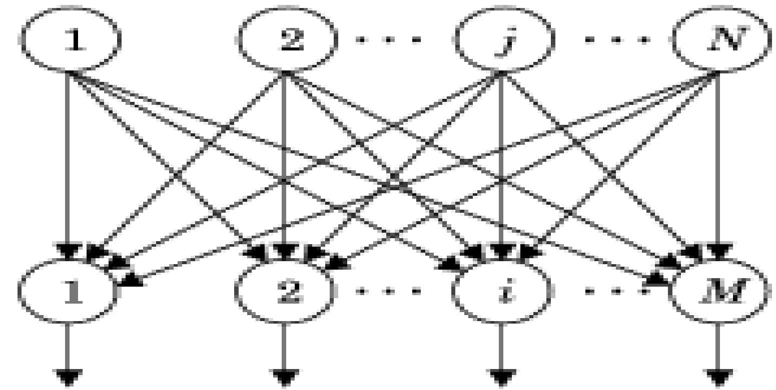
(a) Instar



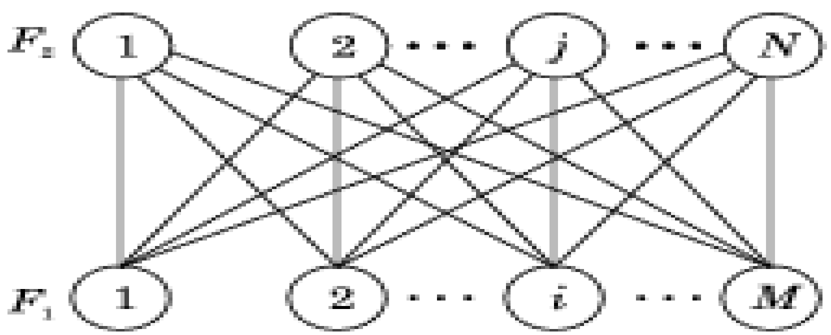
(b) Outstar



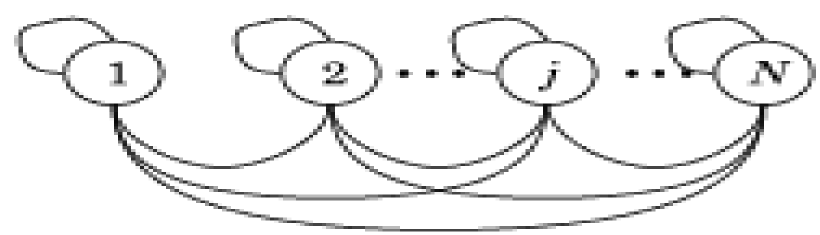
(c) Group of instars



(d) Group of outstars



(e) Bidirectional associative memory



(f) Autoassociative memory

Basic Learning Laws

- Operation of neural network is governed by neuronal dynamics.
- Neuronal dynamics consists of two parts: one corresponding to the dynamics of activation state and other corresponding to the dynamics of the synaptic weights.
- STM in neural network is modeled by activation state of the network
- LTM corresponds to the encoded pattern information in the synaptic weights due to learning.

Lecture 5:

- Basic Learning laws
- What is learning,
- supervised and unsupervised learning

Learning law	Weight adjustment Δw_{ij}	Initial weights	Learning
Hebbian	$\Delta w_{ij} = \eta f(\mathbf{w}_i^T \mathbf{a}) a_j = \eta s_i a_j$, for $j = 1, 2, ..., M$	Near zero	Unsupervised
Perceptron	$\Delta w_{ij} = \eta [b_i - \mathrm{sgn}(\mathbf{w}_i^T \mathbf{a})] a_j = \eta (b_i - s_i) a_j$, for $j = 1, 2, ..., M$	Random	Supervised
Delta	$\Delta w_{ij} = \eta [b_i - f(\mathbf{w}_i^T \mathbf{a})] \dot{f}(\mathbf{w}_i^T \mathbf{a}) a_j = \eta [b_i - s_i] \dot{f}(x_i) a_j$, for $j = 1, 2, ..., M$	Random	Supervised
Widrow-Hoff	$\Delta w_{ij} = \eta [b_i - \mathbf{w}_i^T \mathbf{a}] a_j$, for $j = 1, 2, ..., M$	Random	Supervised
Correlation	$\Delta w_{ij} = \eta b_i a_j$, for $j = 1, 2, ..., M$	Near zero	Supervised
Winner-take-all	$\Delta w_{kj} = \eta (a_j - w_{kj})$, k is the winning unit, for $j = 1, 2, ..., M$	Random but normalised	Unsupervised
Outstar	$\Delta w_{jk} = \eta (b_j - w_{jk})$, for $j = 1, 2, ..., M$	Zero	Supervised

What is Learning

- Learning is a 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 the learning is determined by the manner in which the parameter changes take place. (Mendel & McClaren 1970)

Learning

- Learning is used to update the weights w
- Neural networks can be classified by the algorithm used for learning
- Two most used learning mechanisms are
  - Supervised
  - Unsupervised

Supervised learning

- In supervised learning the weight changes are determined by the difference between the desired output and the actual output.
- Some of the supervised learning laws are: error correction learning or delta rule, stochastic learning, and hardwired systems
- Supervised learning may be used for structural learning or for temporal learning.

Supervised learning

- Structural learning is concerned with capturing in the weights the relationship between a given input-output pattern pair.
- Temporal learning is concerned with capturing in the weights the relationship between neighbouring patterns in a sequence of patterns.

Unsupervised learning

- Unsupervised learning discovers features in a given set of patterns and organizes the patterns accordingly.
- There is no externally specified desired output as in the case of supervised learning.
- Examples of unsupervised learning laws are: Hebbian learning, differential Hebbian learning, principle component learning and competitive learning.

Unsupervised learning

- Unsupervised learning uses mostly local information to update the weights.
- The local information consists of signal or activation values of the units at either end of the connection for which the weight update is being made.

Supervised vs Unsupervised

- Unsupervised learning uses mostly local information to update the weights.
- The local information consists of signal or activation values of the units at either end of the connection for which the weight update is being made.

Lecture 6:

- Functional Units of ANN for pattern recognition task
  - Pattern Recognition Problem
  - Basic functional units

Pattern Recognition Problem

- Pattern Recognition task,

Basic functional units

Lecture 7:

- Pattern Recognition task,
- Content Beyond Syllabus,
- Unit Test &
- Feedback

Basic Functional Units

- Feed Forward Network
- Feedback
- Combination of Both

1. Feedforward ANN
   (a) Pattern association
   (b) Pattern classification
   (c) Pattern mapping/classification

2. Feedback ANN
   (a) Autoassociation
   (b) Pattern storage (LTM)
   (c) Pattern environment storage (LTM)

3. Feedforward and Feedback ANN
   (a) Pattern storage (STM)
   (b) Pattern clustering
   (c) Feature map

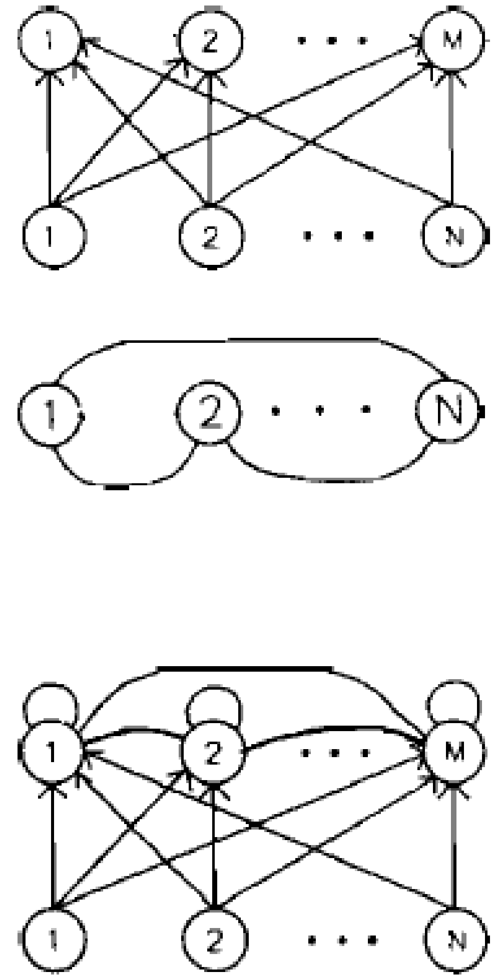


Figure **7.** Summary of ANN for pattern recognition problems.

Feed Forward

Feedback

Competitive Network

PR Task By Functional Units

- PR By Feed Forward

References

- Artificial neural networks by B YEGNANARAYANA
- http://www.burtchiropractic.com/tag/biological-neural-network/