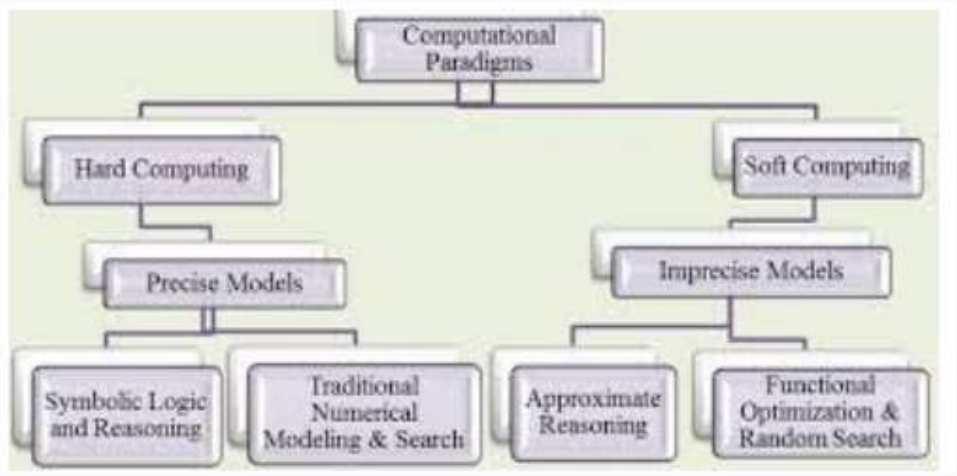
Module 1

Module – 1 (Introduction to Soft Computing & Artificial Neural Network)

Introduction to Soft Computing. Difference between Hard Computing & Soft Computing. Applications of Soft Computing. Artificial Neurons Vs Biological Neurons. Basic models of artificial neural networks – Connections, Learning, Activation Functions. McCulloch and Pitts Neuron. Hebb network.

Computational Paradigms



Computational paradigm is classified into two viz: Hard computing and soft computing.

Hard Computing:- Hard computing is conventional computing. It is based on the principles of precision, certainty, and inflexibility. It requires mathematical model to solve problems. It deals with the precise models. This model is further classified into

1. Symbolic logic and reasoning, and
2. Traditional numerical modelling and search methods.

The basics of traditional artificial intelligence are utilized by these methods. It consumes a lot of time to deal with real life problems which contain imprecise and uncertain information. The following problems cannot accommodate hard computing techniques:

1. Recognition problems
2. Mobile robot coordination, forecasting
3. Combinatorial problems

Soft Computing:- Soft computing deals with approximate models. The term “soft computing" was introduced by Professor Lorfi Zadeh with the objective of exploiting the tolerance for imprecision, uncertainty and partial truth to achieve tractability, robustness, low solution cost and better rapport with reality. The ultimate goal is to be able to emulate the human mind as closely as possible. It is a combination of Genetic Algorithm, Neural Network and Fuzzy Logic.

This model is further classified into two

1. Approximate reasoning, and
2. Functional optimization & random search methods.

It handles imprecise and uncertain information of the real world. It can be used in all industries and business sectors to solve problems. Complex systems can be designed with soft computing to deal with the incomplete information, where the system behavior is not completely known or the existence of measures of variables is noisy.

Hard computing vs Soft computing

Hard Computing	Soft Computing
It **uses precisely stated analytical model.**	It is tolerant to imprecision, uncertainty, partial truth and approximation.
It is based on binary logic and crisp systems.	It is based on fuzzy logic and probabilistic reasoning.
It has features such as precision and categoricity.	It has features such as approximation and dispositionality.
It is deterministic in nature.	It is stochastic in nature.
It can work with exact input data.	It can work with ambiguous and noisy data.
It performs sequential computation.	It performs parallel computation.
It produces precise outcome.	It produces approximate outcome.

Introduction to Soft Computing

Soft computing is an emerging approach to computing which parallel the remarkable ability of the human mind to reason and learn in an environment of uncertainty and imprecision.

Soft computing is based on some biological inspired methodologies such as genetics, evolution, ant's behaviors, particles swarming, human nervous systems, etc.

Now, soft computing is the only solution when we don't have any mathematical modeling of problem solving (i.e.algorithm), need a solution to a complex problem in real time, easy to adapt with changed scenarios and can be implemented with parallel computing.

The real-world problems require systems that combine knowledge, techniques, and methodologies from various sources. These systems should possess human-like expertise within a specific domain, adapt themselves and learn to do better in the changing environments and explain how they make decisions or take actions. Natural language is used by humans for reasoning and drawing conclusions. In conventional AI, the human intelligent behavior is expressed in the language form or symbolic

rules. It manipulates the symbols on the assumption that such behavior can be stored in a symbolically structured knowledge base known as the physical symbol system hypothesis.

Soft computing combines different techniques and concepts. Fuzzy logic, neurocomputing, evolutionary and genetic programming, and probabilistic computing are fields of soft computing.
The main computing paradigms of soft computing are: Fuzzy systems, Neural Networks and Genetic Algorithms.

- Fuzzy sets are for knowledge representation via fuzzy If – Then rules.
- Neural network for learning and adaptivity and
- Genetic algorithm for evolutionary computation.

To achieve close resemblance with human-like decision making, soft computing aims to exploit the tolerance for approximation, uncertainty, imprecision, and partial truth.

- Approximation: the model has similar features but not the same.
- Uncertainty: the features of the model may not be the same as that of the entity/belief.
- Imprecision: the model features (quantities) are not the same as the real ones but are close to them.

Premises of Soft Computing

- The real-world problems are imprecise and uncertain.
- Precision and certainty carry a cost.
- There may not be precise solutions for some problems

Guidelines of Soft Computing

The guiding principle of soft computing is to exploit the tolerance for approximation, uncertainty, imprecision and partial truth to achieve tractability, robustness and low solution cost. Human mind is the role model for soft computing

Applications of Soft Computing

The application of soft computing has proved following advantages:

- The application that cannot be modelled mathematically can be solved.
- Non-linear problems can be solved.
- Introducing human knowledge such as cognition, understanding,recognition, learning and others into the field of computing.

It has enormous applications in many application areas such as medical diagnosis, computer vision, handwritten character recondition, pattern recognition, machine intelligence, weather forecasting, network optimization, VLSI design, etc.

Few applications of soft computing are enlisted below:

• **Handwritten Script Recognition using Soft Computing**: It is one of the demanding parts of computer science. It can translate multilingual documents and sort the various scripts accordingly. Block -level technique concept is used by the system to recognize the script from several script documents given. To classify the script according to their features, it uses Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT) together.

• **Image Processing and Data Compression using Soft Computing**: Image analysis is the high-level processing technique which includes recognition and bifurcation of patterns. It is one of the most important parts of the medical field. The problem of computational complexity and efficiency in the classification can be easily be solved using soft computing techniques. Genetic algorithms, genetic programming, classifier systems, evolutionary strategies, etc are the techniques of soft computing that can be used. These algorithms give the fastest solutions to pattern recognition. These help in analysing the medical images obtained from microscopes as well as examine the X-rays.

• **Use of Soft Computing in Automotive Systems and Manufacturing**: Automobile industry has also adapted soft computing to solve some of the major problems.

Classic control methods are built in vehicles using the Fuzzy logic techniques. It takes the example of human behavior, which is described in the forms of rule – "If-Then "statements.

The logic controller then converts the sensor inputs into fuzzy variables that are then defined according to these rules. Fuzzy logic techniques are used in engine control, automatic transmissions, anti skid steering, etc.

• **Soft Computing based Architecture**: An intelligent building takes inputs from the sensors and controls effectors by using them. The construction industry uses the technique of DAI (Distributed Artificial Intelligence) and fuzzy genetic agents to provide the building with capabilities that match human intelligence. The fuzzy logic is used to create behavior-based architecture in intelligent buildings to deal with the unpredictable nature of the environment, and these agents embed sensory information in the buildings.

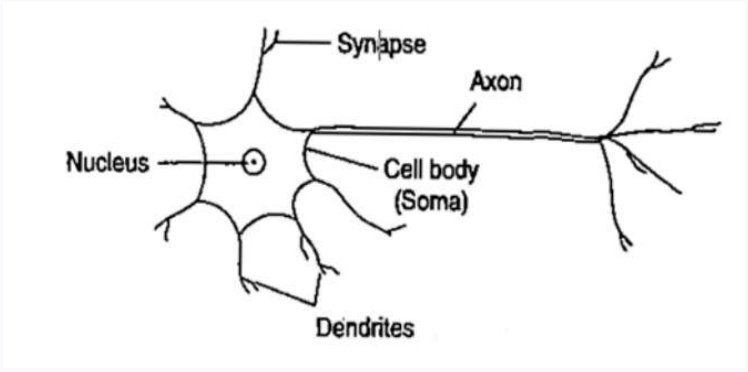
• **Soft Computing and Decision Support System**: Soft computing gives an advantage of reducing the cost of the decision support system. The techniques are used to design, maintain, and maximize the value of the decision process. The first application of fuzzy logic is to create a decision system that can predict any sort of risk. The second application is using fuzzy information that selects the areas which need replacement.

• **Soft Computing Techniques in Power System Analysis**: Soft computing uses the method of Artificial Neural Network (ANN) to predict any instability in the voltage of the power system. Using the ANN, the pending voltage instability can be predicted. The methods which are deployed here are very low in cost.

• **Soft Computing Techniques in Bioinformatics**: The techniques of soft computing help in modifying any uncertainty and indifference that biometrics data may have. Soft computing is a technique that provides distinct low-cost solutions with the help of algorithms, databases, Fuzzy Sets (FSs), and Artificial Neural Networks (ANNs). These techniques are best suited to give quality results in an efficient way.

• **Soft Computing in Investment and Trading**: The data present in the finance field is in opulence and traditional computing is not able to handle and process that kind of data. There are various approaches done through soft computing techniques that help to handle noisy data. Pattern recognition technique is used to analyze the pattern or behavior of the data and time series is used to predict future trading points.

Biological Neuron



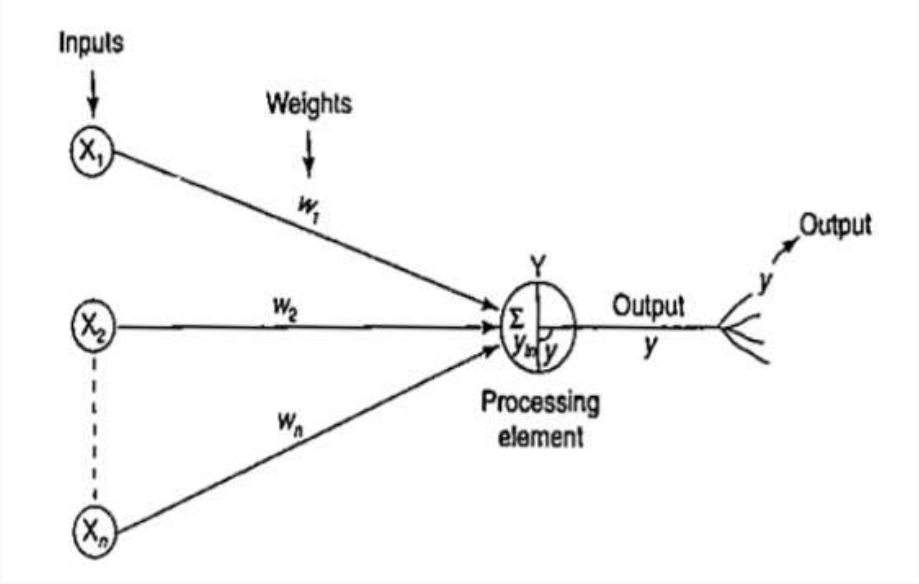
Schematic diagram of a biological neuron

The biological neuron consists of main three parts:

- Soma or cell body-where cell nucleus is located
- Dendrites-where the nerve is connected to the cell body
- Axon-which carries the impulses of the neuron

Dendrites are tree-like networks made of nerve fiber connected to the cell body. An Axon is a single, long connection extending from the cell body and carrying signals from the neuron. The end of the axon splits into fine strands. It is found that each strand terminated into small bulb-like organs called synapse. It is through synapse that the neuron introduces its signals to other nearby neurons. The receiving ends of these synapses on the nearby neurons can be found both on the dendrites and on the cell body. There are approximately 104 synapses per neuron in the human body. Electric impulse is passed between

synapse and dendrites. It is a chemical process which results in increase/decrease in the electric potential inside the body of the receiving cell. If the electric potential reaches a threshold value, receiving cell fires & pulse / action potential of fixed strength and duration is sent through the axon to the synaptic junction of the cell. After that, the cell has to wait for a period called refractory period.



Mathematical model of artificial neuron

Biological neuron	Artificial neuron
Cell	Neuron
Dendrites	Weights or interconnections
Soma	Net input
Axon	Output

In this model net input is calculated as

$$y_{in} = x_1 w_1 + x_2 w_2 + \cdots + x_n w_n = \sum_{i=1}^{n} x_i w_i$$

Where, i representes ith processing element. The activation function applied over it to calculate the output. The weight represents the strength of synapses connecting the input and output.

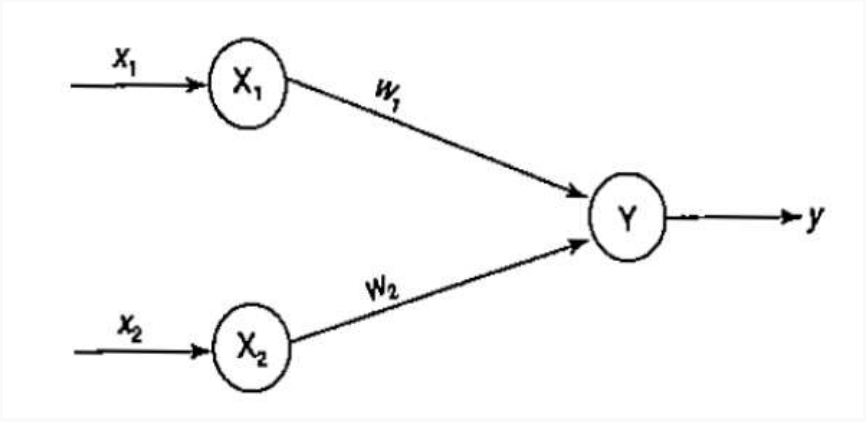
Biological neuron vs Artificial neuron

Term	Brain	Computer
Speed	Execution time is few milliseconds	Execution time is few nano seconds
Processing	Perform massive parallel operations simultaneously	Perform several parallel operations simultaneously. It is faster the biological neuron
Size and complexity	Number of Neuron is 10^{11} and number of interconnections is 10^{15}. So complexity of brain is higher than computer	It depends on the chosen application and network designer.
Storage capacity	• Information is stored in interconnections or in synapse strength. • New information is stored without destroying old one. • Sometimes fails to recollect information	• Stored in continuous memory location. • Overloading may destroy older locations. • Can be easily retrieved
Tolerance	• Fault tolerant • Store and retrieve information even interconnections fails • Accept redundancies	• No fault tolerance • Information corrupted if the network connections disconnected. • No redundancies
Control mechanism	Depends on active chemicals and neuron connections are strong or weak	CPU Control mechanism is very simple

Artificial Neural Network (ANN)

An artificial neural network (ANN) is an efficient information processing system which resembles the characteristics of biological neural network. ANNs contain large number of highly interconnected processing elements called nodes or neurons or artificial neurons or units. Each neuron is connected with other by connection link and each connection link is associated with weights which contain information about the input signal. This information is used by neuron net to solve a particular problem. ANNs have ability to learn, recall and

generalize training pattern or data similar to that of human brain.



Architecture of a simple artificial neuron net

Each neuron has an internal state of its own, called activation or activity level of neuron which is the function of the inputs the neuron receives. The activation signal of a neuron is transmitted to other neurons. A neuron can send only one signal at a time which can be transmitted to several neurons.

Consider the above figure, here X1 and X2 are input neurons, Y is the output neuron W1 and W2 are the weights net input is calculated as

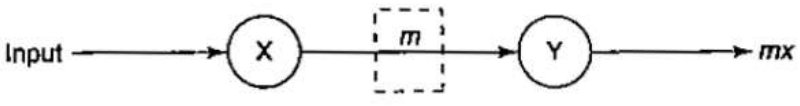
$$y_{in} = x_1 w_1 + x_2 w_2$$

where x1 and x2 are the activation of the input neurons X1 and X2, i.e., is the output of the input signals. The output y of the output neuron Y can be obtained by applying activations over the net input.

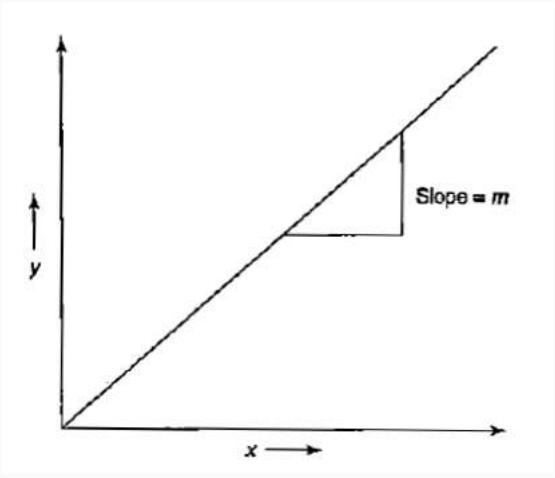
$$y = f(y_{in})$$

Output = Function (net input calculated)

The function to be applied over the net input is called activation function. The net input calculation is similar to the calculation of output of a pure linear straight line equation y=mx



Neural net of pure linear equation



Graph for y = mx

Characteristics of ANN:

- It is a neurally implemented mathematical model
- Large number of processing elements called neurons exists here.
- Interconnections with weighted linkage hold informative knowledge.
- Input signals arrive at processing elements through connections and connecting weights.
- Processing elements can learn, recall and generalize from the given data.
- Computational power is determined by the collective behavior of neurons.
    - ANN is a connection models, parallel distributed processing models, self-organizing systems, neuro-computing systems and neuro - morphic system.

Evolution of neural networks

Year	Neural network	Designer	Description
1943	McCulloch and Pitts neuron	McCulloch and Pitts	Arrangement of neurons is combination of logic gate. Unique feature is thresh hold
1949	Hebb network	Hebb	If two neurons are active, then their connection strengths should be increased.
1958, 1959, 1962, 1988,	Perceptron	Frank Rosenblatt, Block, Minsky and Papert	Here the weights on the connection path can be adjusted.

1960	Adaline	Widrow and Hoff	Here the weights are adjusted to reduce the difference between the net input to the output unit and the desired output.
1972	Kohonen self-organizing feature map	Kohonen	Inputs are clustered to obtain a fired output neuron.
1982, 1984, 1985, 1986, 1987	Hopfield network	John Hopfield and Tank	Based on fixed weights. Can act as associative memory nets
1986	Back propagation network	Rumelhart, Hinton and Williams	• Multilayered • Error propagated backward from output to the hidden units
1988	Counter propagation network	Grossberg	Similar to kohonen network.
1987-1990	Adaptive resonance Theory(ART)	Carpenter and Grossberg	Designed for binary and analog inputs.
1988	Radial basis function network	Broomhead and Lowe	Resemble back propagation network, but activation function used is Gaussian function.
1988	Neo cognitron	Fukushima	For character recognition.

Basic models of artificial neural networks

Models are based on three entities

- The model's synaptic interconnections.
- The training or learning rules adopted for updating and adjusting the connection weights.
- Their activation functions

1. **Connections**

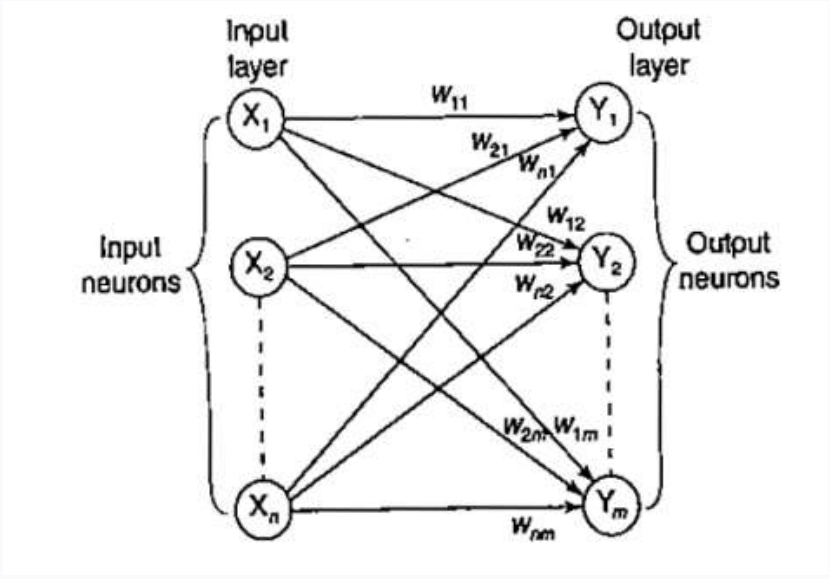
   The arrangement of neurons to form layers and the connection pattern formed within and between layers is called the network architecture. There exist five basic types of connection architecture. They are:

   1. Single layer feed forward network
   2. Multilayer feed-forward network

3. Single node with its own feedback
4. Single-layer recurrent network
5. Multilayer recurrent network

Feed forward network: If no neuron in the output layer is an input to a node in the same layer / proceeding layer.
Feedback network: If outputs are directed back as input to the processing elements in the same layer/proceeding layer.
Lateral feedback: If the output is directed back to the input of the same layer.
Recurrent networks: Are networks with feedback networks with closed loop.

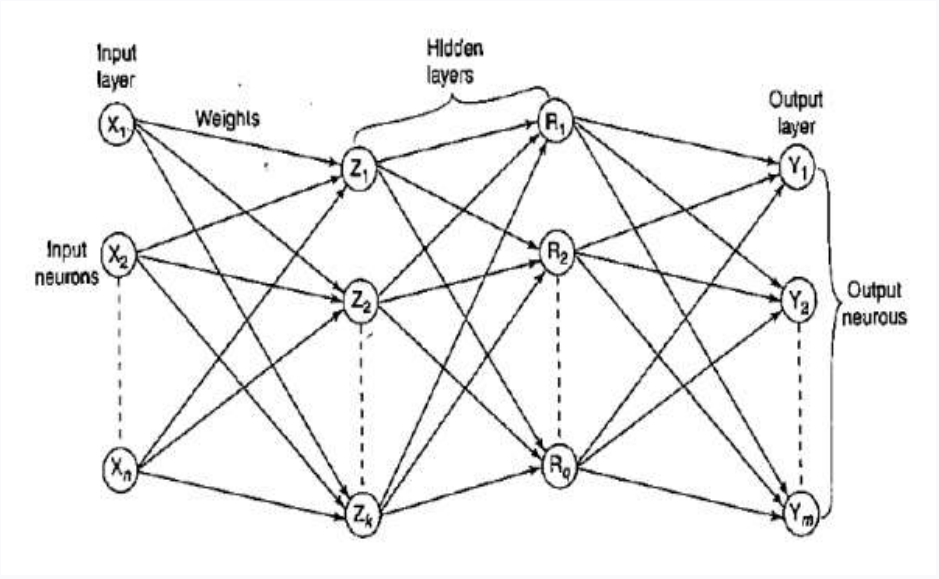
1. **Single layer feed forward network**
   Layer is formed by taking processing elements and combining it with other processing elements. Input and output are linked with each other Inputs are connected to the processing nodes with various weights, resulting in series of outputs one per node.



When a layer of processing nodes is formed the inputs can be connected to these nodes with various weights, resulting in a serious of outputs, one per node. This is called single layer feedforward network.

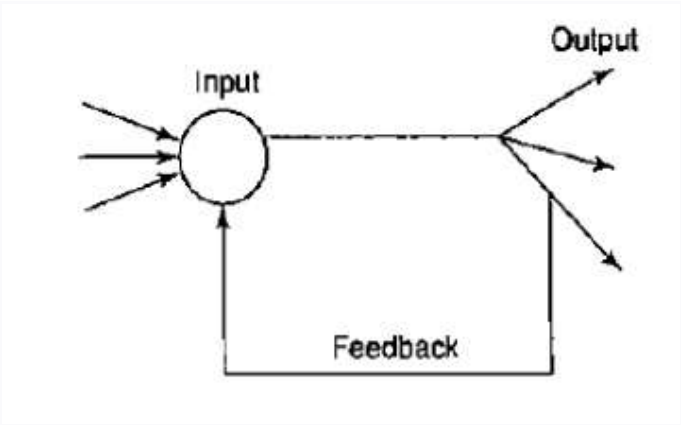
2. **Multilayer feed-forward network**
   This network is formed by the interconnection of several layers. Input layer receives input and buffers input signal. Output layer generated output. Layer between input and output is called hidden layer. Hidden layer is internal to the network. There are Zero to several hidden layers

in a network. More the hidden layer more is the complexity of the network, but efficient output is produced.



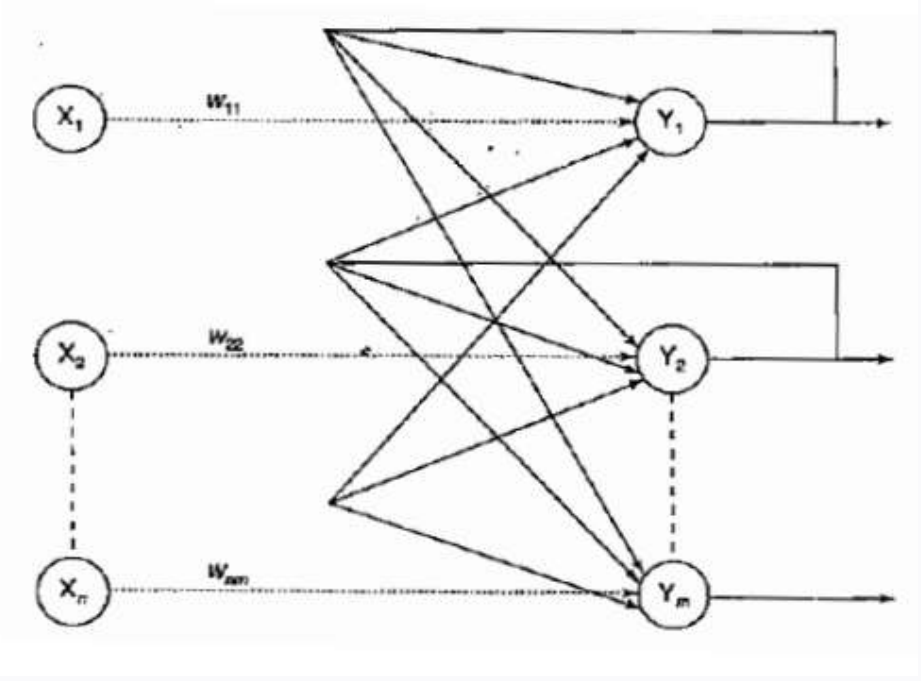
3. **Single node with its own feedback**

   It is a simple recurrent neural network having a single neuron with feedback to itself.



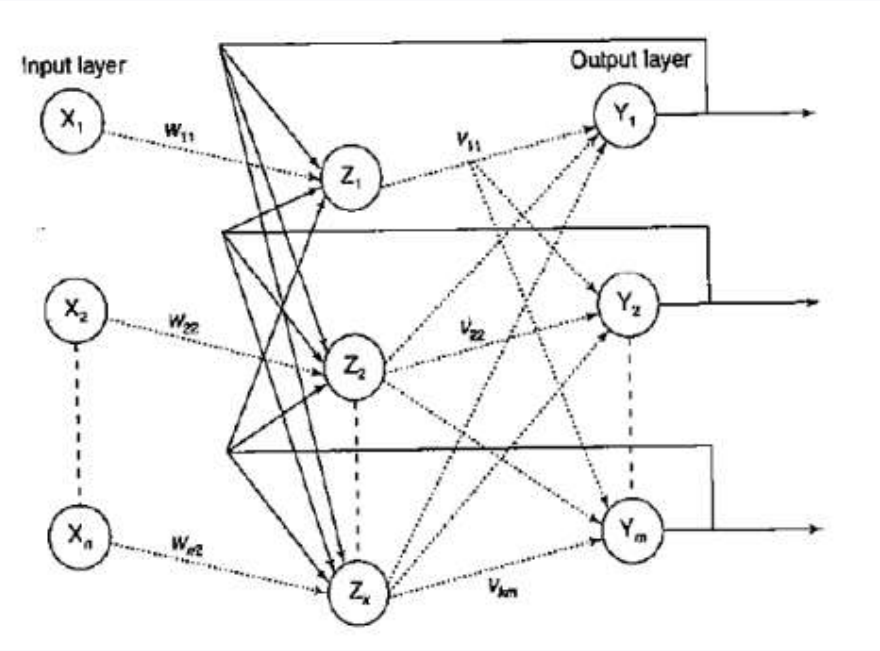
4. **Single layer recurrent network**

   A single layer network with feedback from output can be directed to processing element itself or to other processing element/both.

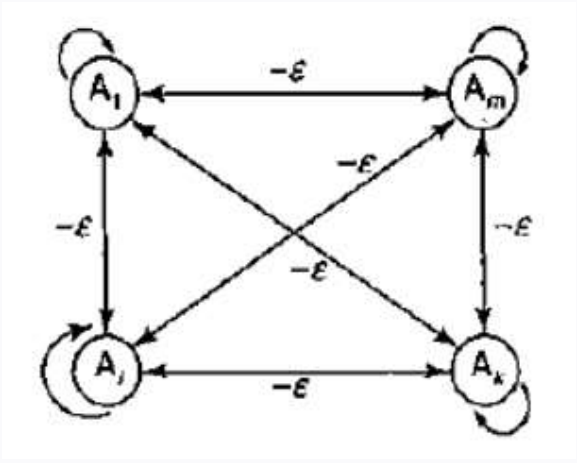


5. Multilayer recurrent network

Processing element output can be directed back to the nodes in the preceding layer, forming a multilayer recurrent network.

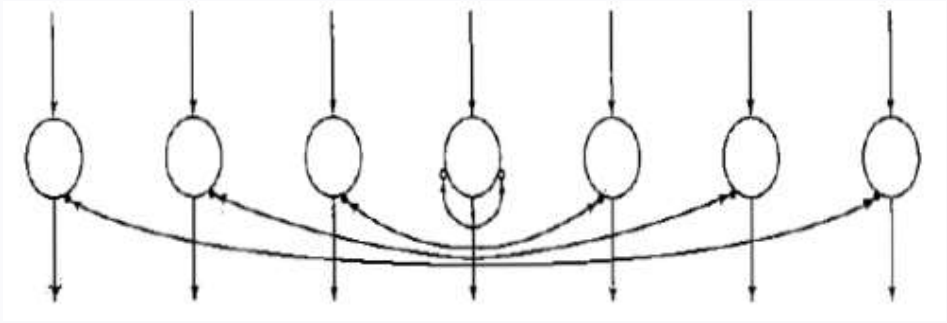


Maxnet –competitive interconnections having fixed weights



Competitive networks

On-center-off-surround/lateral inhibition structure – each processing neuron receives two different classes of inputs- "excitatory" input from nearby processing elements & "inhibitory" elements from more distantly located processing elements. This type of interconnection is shown below



Lateral inhibition structure

2. **Learning**

   Learning or Training is the process by means of which a neural network adapts itself to a stimulus by making proper parameter adjustments, resulting in the production of desired response.

   Two broad kinds of learning in ANNs is:

   i) Parameter learning – updates connecting weights in a neural net.

   ii) Structure learning – focus on change in the network.

   Apart from these, learning in ANN is classified into three categories as

   i) Supervised learning

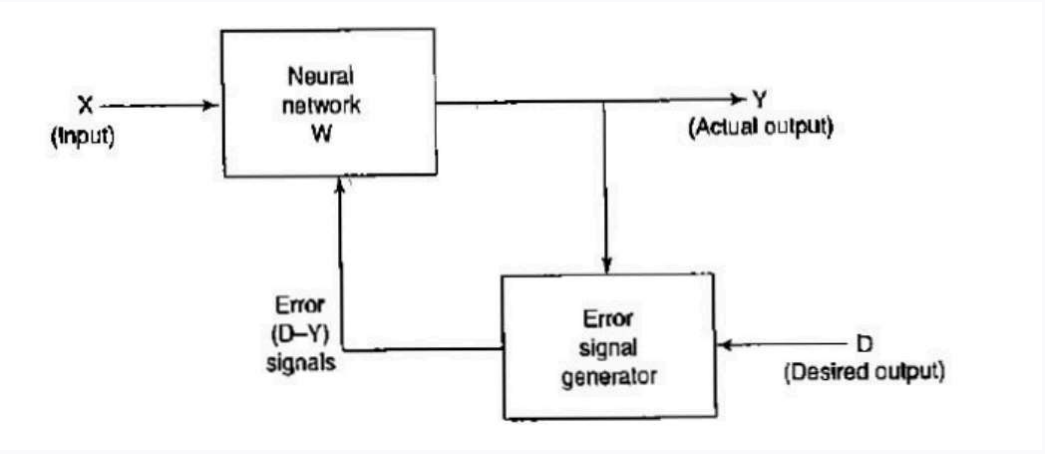
   ii) Unsupervised learning

   iii) Reinforcement learning

   i) Supervised learning

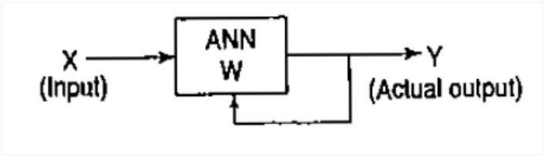
   The Learning here is performed with the help of a teacher. Example: Consider the learning process of a small child. Child doesn't know how to read/write. Their each and every action is supervised by a teacher. Actually a child works on the basis of the output that he/she has to produce. In ANN, each input vector requires a corresponding target vector, which represents the desired output. The input vector along with target vector is called training pair. Input

vector results in output vector. The actual output vector is compared with desired output vector. If there is a difference means an error signal is generated by the network. It is used for adjustment of weights until actual output matches desired output.



Unsupervised learning

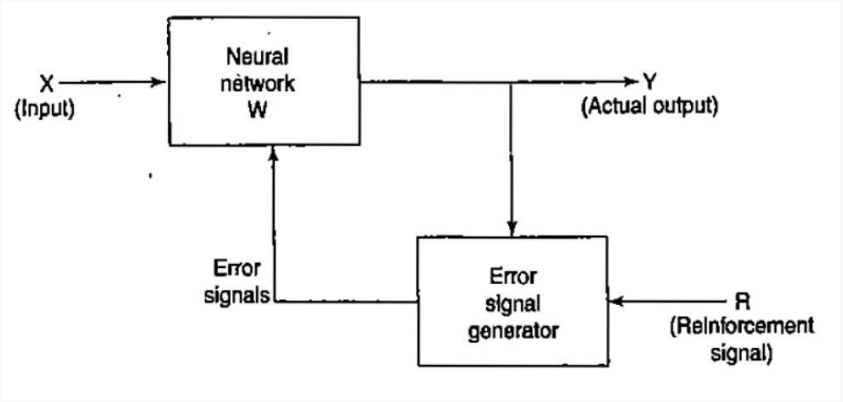
Learning is performed without the help of a teacher. Example: tadpole – learn to swim by itself. In ANN, during training process, network receives input patterns and organize it to form clusters.



From the above figure it is observed that no feedback is applied from environment to inform what output should be or whether they are correct. The network itself discover patterns, regularities, features/ categories from the input data and relations for the input data over the output. Exact clusters are formed by discovering similarities & dissimilarities so called as self – organizing.

Reinforcement learning

It is similar to supervised learning. Learning based on critic information is called reinforcement learning & the feedback sent is called reinforcement signal. The network receives some feedback from the environment. Feedback is only evaluative.



The external reinforcement signals are processed in the critic signal generator, and the obtained critic signals are sent to the ANN for adjustment of weights properly to get critic feedback in future.

Activation Functions

To make work more efficient and for exact output, some force or activation is given. Like that, activation function is applied over the net input to calculate the output of an ANN. Information processing of processing element has two major parts: input and output. An integration function (f) is associated with input of processing element.

Several activation functions are there.

1. Identity function: It is a linear function which is defined as

$$f(x) = x \text{ for all } x$$

The output is same as the input.

2. Binary step function: This function can be defined as

$$f(x) = \begin{cases} 1 \text{ if } x \geq \theta \\ 0 \text{ if } x < \theta \end{cases}$$

Where, θ represents thresh hold value. It is used in single layer nets to convert the net input to an output that is binary (0 or 1).

3. Bipolar step function: This function can be defined as

$$f(x) = \begin{cases} 1 \text{ if } x \geq \theta \\ -1 \text{ if } x < \theta \end{cases}$$

Where, θ represents threshold value. It is used in single layer nets to convert the net input to an output that is bipolar (+1 or -1).

4. Sigmoid function: It is used in Back propagation nets.

Two types:

a) Binary sigmoid function: It is also termed as logistic sigmoid function or unipolar sigmoid function. It is defined as

$$f(x) = \frac{1}{1 + e^{-\lambda x}}$$

where, λ represents steepness parameter. The derivative of this function is

$$f'(x) = \lambda f(x)[1 - f(x)]$$

The range of sigmoid function is 0 to 1.

b) Bipolar sigmoid function: This function is defined as

$$f(x) = \frac{2}{1 + e^{-\lambda x}} - 1 = \frac{1 - e^{-\lambda x}}{1 + e^{-\lambda x}}$$

Where λ represents steepness parameter and the sigmoid range is between -1 and +1. The derivative of this function can be

$$f'(x) = \frac{\lambda}{2}[1 + f(x)][1 - f(x)]$$

It is closely related to hyperbolic tangent function, which is written as

$$h(x) = = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$$

$$h(x) = = \frac{1 - e^{-2x}}{1 + e^{-2x}}$$

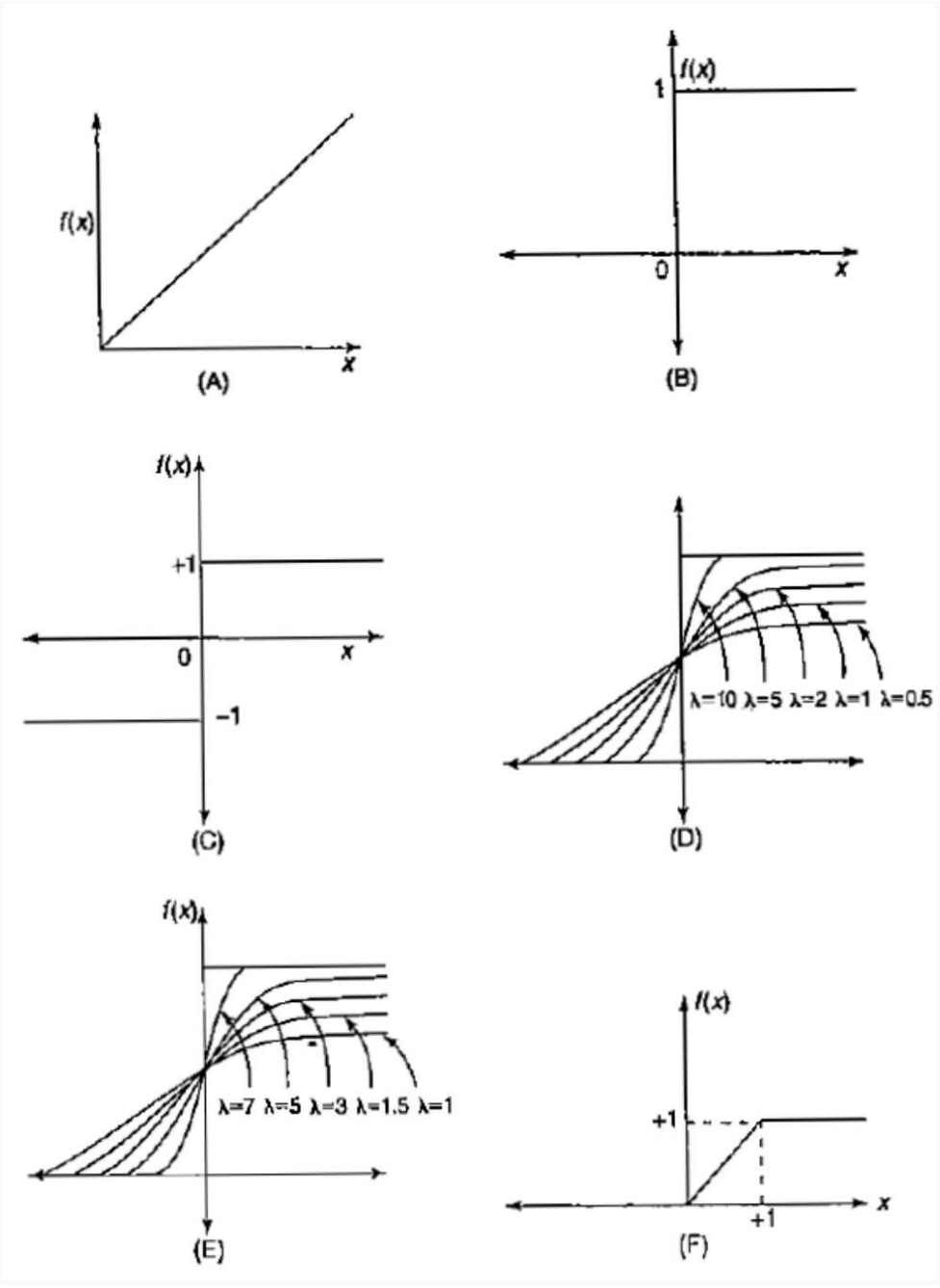
The derivative of the hyperbolic tangent function is

$$h'(x) = [1 + h(x)][1 - h(x)]$$

5. Ramp function: The ramp function is defined as

$$f(x) = \begin{cases} 1 \; if x > 1 \\ x \; if \; 0 \leq x \leq 1 \\ 0 \; if \; x < 0 \end{cases}$$

The graphical representation of all these function is given in the upcoming figure

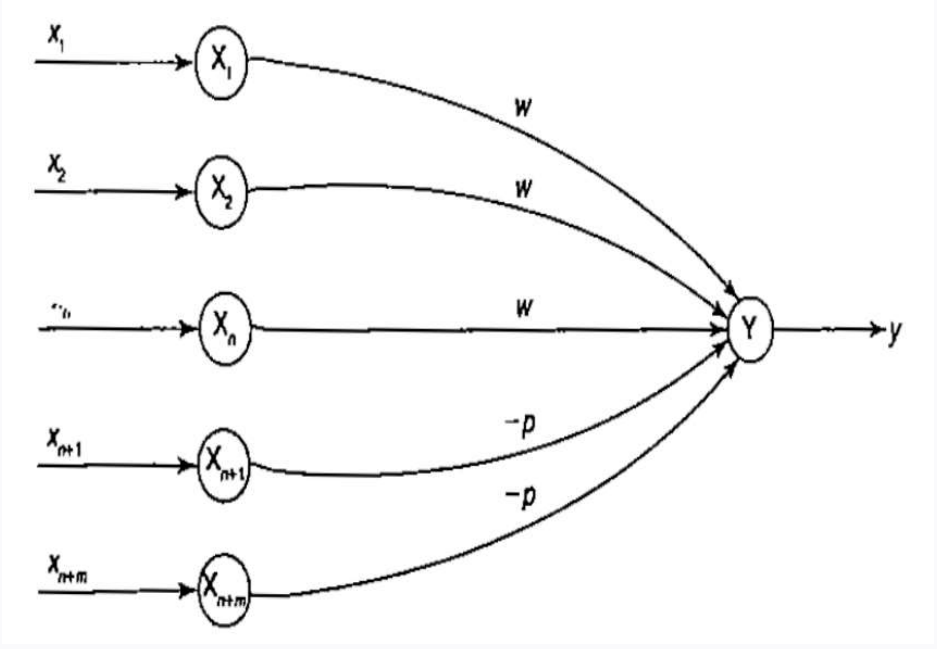


Depiction of activation functions: (A) identity function; (B) binary step function; (C) bipolar step function; (D) binary sigmoidal function; (E) bipolar sigmoidal function; (F) ramp function.

McCulloch and Pitts Neuron

It is discovered in 1943 and usually called as M-P neuron. M-P neurons are connected by directed weighted paths. Activation of M-P neurons is binary (i.e) at any time step the neuron may fire or may not fire. Weights associated with communication links may be excitatory (wgts are positive)/inhibitory (wgts are negative). Threshold plays major role here. There is a fixed threshold for each neuron and if the net input to the neuron is greater than the threshold then the neuron fires. They are widely used in logic functions. A

simple M-P neuron is shown in the figure. It is excitatory with weight w (w>0) / inhibitory with weight –p (p<0). In the Fig., inputs from x_1 to x_n possess excitatory weighted connection and X_{n+1} to x_{n+m} has inhibitory weighted interconnections.



McCulloch-Pins neuron model

Since the firing of neuron is based on threshold, activation function is defined as

$$f(x) = \begin{cases} 1 \ if \ y_{in} \geq \theta \\ 0 \ if \ y_{in} < \theta \end{cases}$$

For inhibition to be absolute, the threshold with the activation function should satisfy the following condition:

$$\theta > nw - p$$

Output will fire if it receives "k" or more excitatory inputs but no inhibitory inputs where

$$kw \geq \theta > (k-1)\ w$$

The M-P neuron has no particular training algorithm. An analysis is performed to determine the weights and the threshold. It is used as a building block where any function or phenomenon is modeled based on a logic function.

Hebb network

Donald Hebb stated in 1949 that "In brain, the learning is performed by the change in the synaptic gap". When an axon of cell A is near enough to excite cell B, and repeatedly or permanently takes place in firing it, some growth process or metabolic change takes place in one or both the cells such that A's efficiency, as one of the cells firing B, is increased. According to Hebb rule, the weight vector is found to increase proportionately to the product of

the input and the learning signal. In Hebb learning, two interconnected neurons are 'on' simultaneously. The weight update in Hebb rule is given by

$$w_i(new) = w_i(old) + x_i y$$

Hebbs network is suited more for bipolar data. If binary data is used, the weight updation formula cannot distinguish two conditions namely:

1. A training pair in which an input unit is "on" and the target value is "off".
2. A training pair in which both the input unit and the target value is "off".

Training algorithm

The training algorithm is used for the calculation and adjustment of weights. The flowchart for the training algorithm of Hebb network is given below

Step 0: First initialize the weights. Basically in this network they may be set to zero, i.e., $w_i = 0$, for i= 1 to n where $"n"$ may be the total number of input neurons.

Step 1: Steps 2-4 have to be performed for each input training vector and target output pair, s: t.

Step 2: Input units activations are set. Generally, the activation function of input layer is identity function: $x_i = s_i$ *for* $i=1$ *to* n

Step 3: Output units activations are set: $y = t$.

Step 4: Weight adjustments and bias adjustments are performed:

$$w_i(new) = w_i(old) + x_i y$$

$$b(new) = b(old) + y$$

In step 4, the weight updation formula can be written in vector form as

$$w(new) = w(old) + y$$

Hence, Change in weight is expressed as

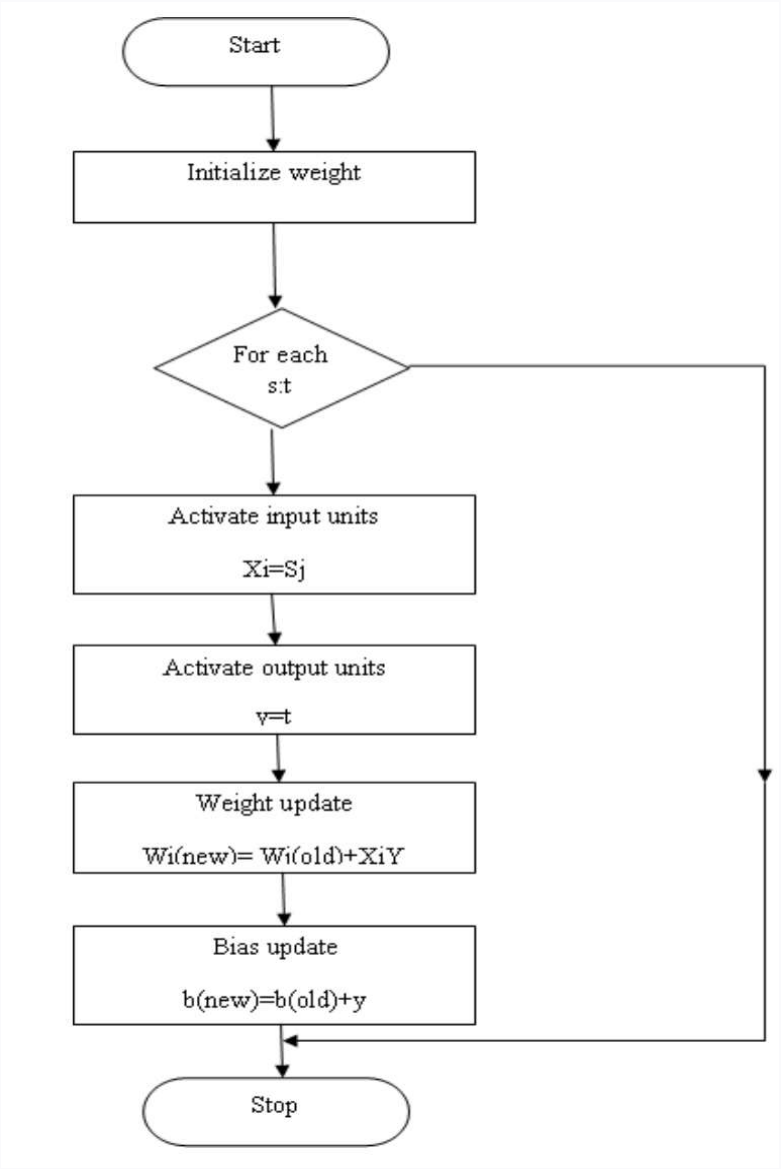
$$\Delta w = xy$$

As a result,

$$w(new) = w(old) + \Delta w$$

Hebb rule is used for pattern association, pattern categorization, pattern classification and over a range of other areas.

Flowchart of Training algorithm



Flowchart of Hebb training algorithm