ANNFL Solution

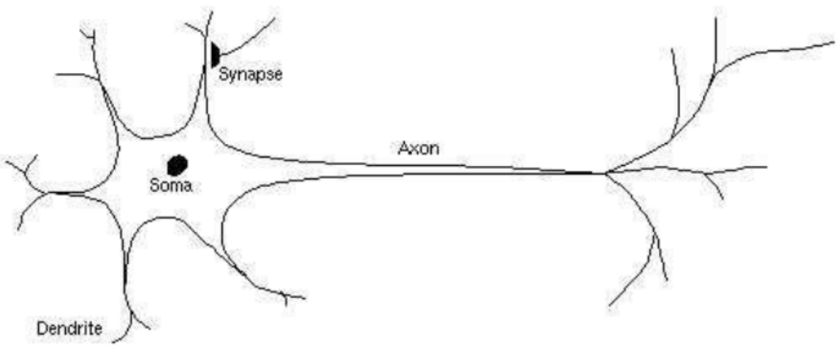
2m

1. Define an artificial neural network.

2. Draw biological neural network and artificial neural network.

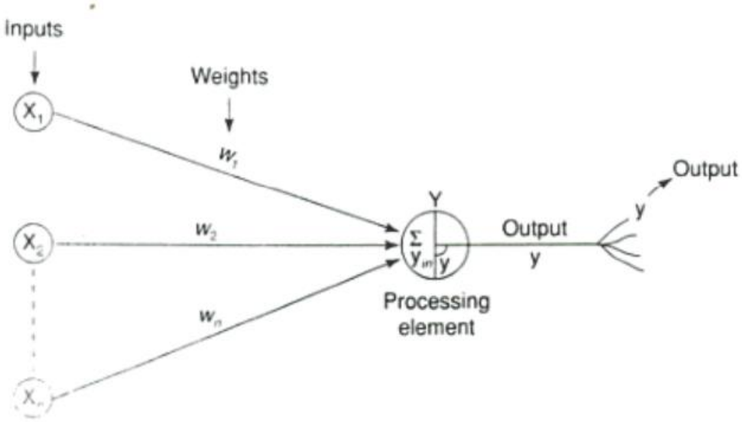
Biological Neural Network

The human brain consists of approximately 10” highly interconnected neurons. A schematic diagram of a biological neuron is shown below:



It consists of 3 main parts:

1) **Soma** or Cell body - where the cell nucleus is located.
2) **Dendrites** - where the nerve is connected to the cell body.
3) **Axon** - which carries the impulses of the neuron.



In this model, the net input

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

where i represents the i^th processing element. The activation function is applied over it to calculate the output. The weight represents the strength of synapse connecting the input and output neurons. A positive weight corresponds to an excitatory synapse, and a negative to an inhibitory synapse.

3. What are the advantages of artificial neural networks?

Benefits of Neural Networks

It is apparent that a neural network derives its computing power through, first, its massively parallel distributed structure and, second, its ability to learn and therefore generalize. *Generalization* refers to the neural network producing reasonable outputs for inputs not encountered during training (learning). These two information-processing capabilities make it possible for neural networks to solve complex (large-scale) problems that are currently intractable. In practice, however, neural networks cannot provide the solution by working individually. Rather, they need to be integrated into a consistent system engineering approach. Specifically, a complex problem of interest is *decomposed* into a number of relatively simple tasks, and neural networks are assigned a subset of the tasks that *match* their inherent capabilities.

a)Nonliearity
b)Input-output mapping
c)Adaptivity
d)Evidential Response
e)Fault Tolerance
f)Vlsi Implementability
g)Contextual Information

4. What are the applications of artificial neural networks?

- Neurons, in one form or another, represent an ingredient *common* to all neural networks.
- This commonality makes it possible to *share* theories and learning algorithms in different applications of neural networks.
- Modular networks can be built through a *seamless integration of modules*.

5. What is the need of activation functions?

Activation Functions

Activation helps in achieving the exact output. The activation function is applied over the net input to calculate the output of an ANN.

The information processing of a processing element can be viewed as consisting of two major parts: input and output. An integration function (say f) is associated with the input of a processing element. This function serves to combine activation, information or evidence from an external source or other processing elements into a net input to the processing element. The nonlinear activation function is used to ensure that a neurons response is bounded, that is the actual response of the neuron is conditioned or dampened as a result of large or small activating stimuli and is thus controllable.

Certain nonlinear functions are used to achieve the advantages of a multilayer network from a single layer network. When a signal is fed through a multilayer network with linear activation functions, the output obtained remains same as that could be obtained using a single layer network. So nonlinear functions are widely used in multilayer networks as compared to linear functions.

6. What is learning rate? What is momentum factor?

Learning Rate (α)

It is used to control the amount of weight adjustment at each step of training. The learning rate, ranging from 0 to 1, determines the rate of learning at each step.

Momentum Factor

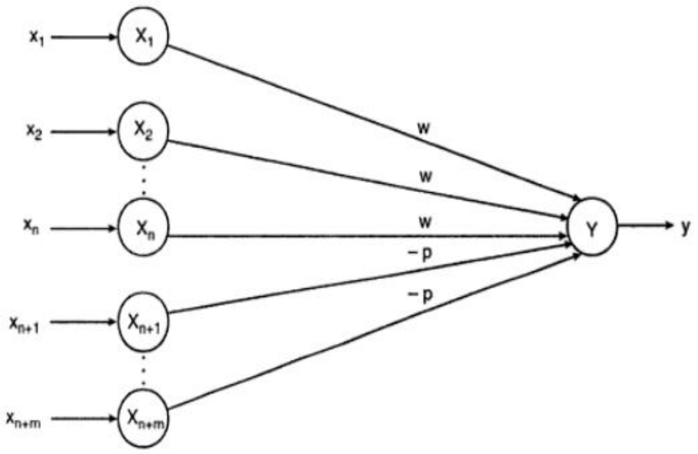
Convergence is made faster if a momentum factor is added to the weight updation process. This is generally done in the back propagation network.

If momentum has to be used, the weights from one or more previous training patterns must be solved.

Momentum helps the net in reasonably large weight adjustments until the corrections are in the same general direction for several patterns.

7. Draw McCulloch-Pitt’s neuron architecture.

Architecture of a simple M-P neuron is shown below:



Inputs from x_1 to x_n possess excitatory weighted connections and inputs from x_{n+1} to x_{n+m} possess inhibitory weighted interconnection. Since the firing of the output neuron is based upon the threshold, the activation function is defined as

$f(y_{in}) = 1$, if $y_{in} \geq \theta$

$= 0$, if $y_{in} < \theta$

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

$\theta > nw\text{-}p$

The output will fire if it receives say 'k' or more excitatory inputs but no inhibitory inputs where

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

The M-P neuron has no particular training algorithm. The M-P neurons are used as building blocks on which we can model any function or phenomenon which can be represented as a logic function.

8. Obtain the output of the neuron Y for the network shown in the figure for given activation functions.

5m

1. Write note on: Architecture of Artificial Neural Network

Connections (Architectures)

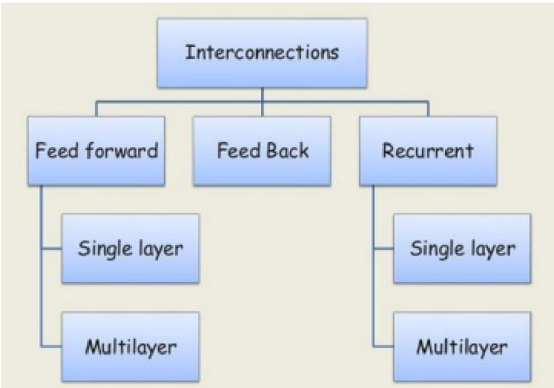
The neurons should be visualized for their arrangements in layers.

An ANN consists of a set of highly interconnected processing elements (neurons) such that each processing element output is found to be connected through weights to other processing elements or to itself.

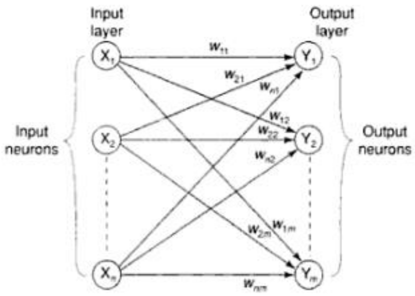
The arrangement of neurons to form layers and the connection pattern formed within and between layers is called the network architecture.

There are 5 basic types of neuron connection architectures:

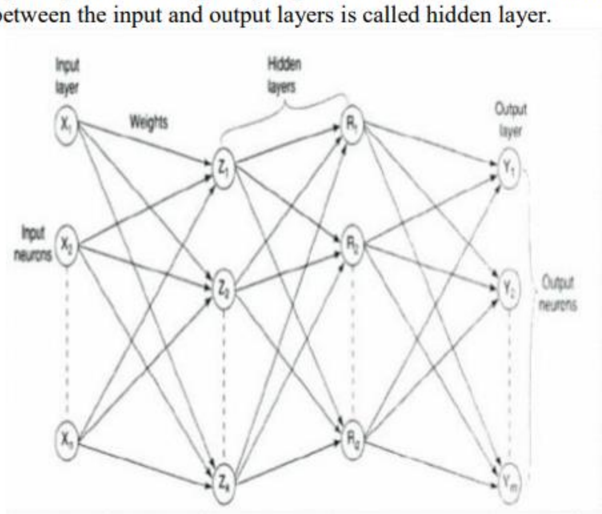
1) **Single layer feed forward network.**
2) **Multilayer layer feed forward network.**
3) **Single node with its own feedback**
4) **Single layer recurrent network**
5) **Multilayer recurrent network**



Basically, neural nets are classified into single layer or multiplayer neural nets. A layer is formed by taking a processing element and combining it with other processing elements. The linked interconnections lead to the formation of various network architectures. When a layer of the processing nodes is formed, the inputs can be connected to these nodes with various weights, resulting in a series weights, resulting in a series of outputs, one per node. Thus a single layer feed-forward network is formed Write note on: Types of Learning of ANNs.

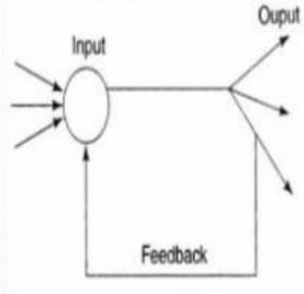


A multilayer feed forward network is formed by interconnection of several layers.
The input layer is that which receives the input and this layer has no function except buffering the input signal. The output layer generates the output of the network. Any layer that is formed between the input and output layers is called hidden layer.

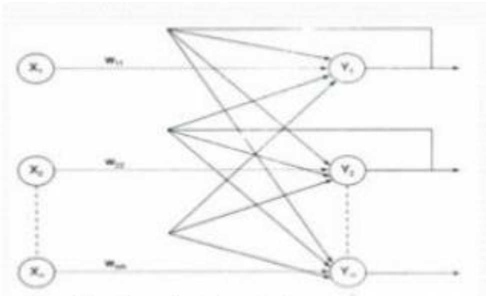


The hidden layer is internal to the network and has no direct contact with the external environment. There may be zero to several hidden layers in an ANN. More the number of hidden layers more is the complexity of the network. In case of a fully connected network, every output from one layer is connected to each and every node in the next layer.

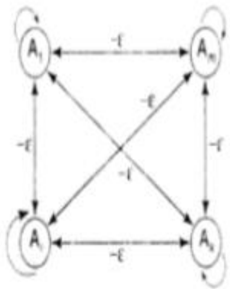
A network is said to be a feed forward network if no neuron in the output layer is an input to the node in the same layer or in the preceding layer. Feedback networks have outputs which can be directed back as inputs to the same or preceding layer nodes. If the feedback is directed back as input to the processing elements in the same layer then it is called lateral feedback. Recurrent networks are feedback networks with closed loops. The following figure shows a simple recurrent neural network having a single neuron with feedback to itself



The figure below shows a single layer network with a feedback connection in which a processing element's output can be directed back to the processing element itself or to other processing elements or both.

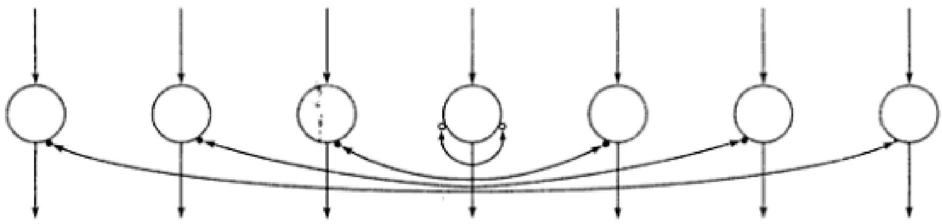


The architecture of a competitive layer is shown below:

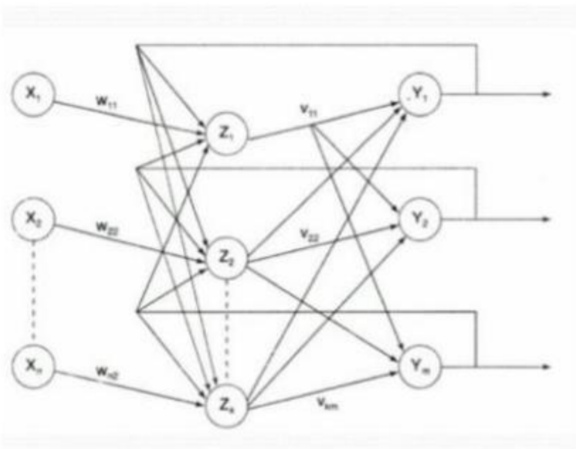


The competitive interconnections have fixed weights – ε. This net is called Maxnet.

Another type of architecture with lateral feedback which is called On-Centre-Off-Surround or Lateral Inhibition Structure is shown below:



The output of a processing element can be directed back to the nodes in a preceding layer' forming a multilayer recurrent network. In these networks, a processing element output can be directed to the processing element itself and to others in the same layer.



2. Write note on: Types of Learning of ANNs

Learning

The main property of an ANN is its capability to learn.
Learning or training is a 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.
Broadly, there are two kinds of learning:

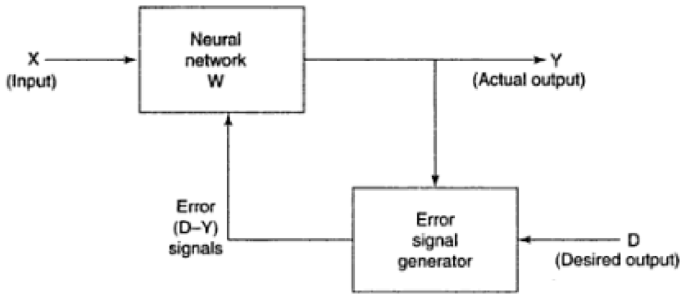
1) Parameter Learning : It updates the connecting weights
2) Structure Learning: It focuses on the change in network structure which includes the number of processing elements as well as their connections.

These two types of learning can be performed simultaneously or separately.
Apart from these two categories, the learning in an ANN can be generally classified into 3 categories:

1) Supervised learning
2) Unsupervised learning
3) Reinforcement learning

1) **Supervised Learning**:

The learning here is performed with the help of a teacher e.g. learning process of a child. The child does not know how to read/write. Training is done by a teacher and each and every action is supervised. The child works on the basis of the output that is to be produced.
Similarly, in an ANNs following the supervised learning, each input vector requires a corresponding target vector, which represents the desired output. The input vector along with the target vector is called training pair. The network here is informed precisely about what should be emitted as output.



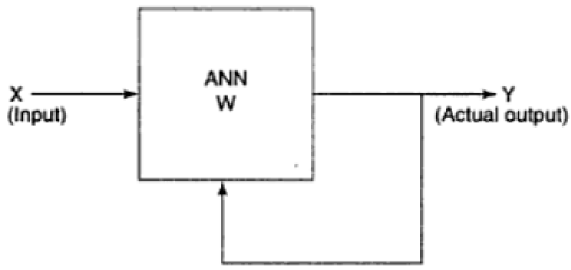
During training, the input vector is presented to the network, which results in an output vector. This output vector is the actual output vector. Then the actual output vector is compared with the desired (target) output vector. If there exists a difference between the two output vectors then an error signal is generated by the network. This error signal is used for adjustment of weights until the actual output matches the desired (target) output. In this type of training, a supervisor is required for error minimization. Hence the network trained by this method is said to be using supervised training methodology. Here it is assumed that correct “target” output values are know for each input pattern.

2) Unsupervised Learning

The learning performed here is without the help of a teacher.

In ANNs following unsupervised learning, the input vectors of similar type are grouped without the use of training data to specify how a member of each group looks or to which group a number belongs. In the training process, the network receives the input patterns and organizes these patterns to form clusters.

When a new pattern is applied as input, the NN gives an output response indicating the class to which the input pattern belongs. If for an input, a pattern class cannot be found, then a new class is generated. The block diagram below shows unsupervised learning.

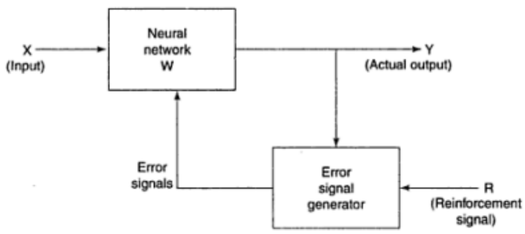


It is clear that there is no feedback from the environment to inform what the outputs should be or whether the outputs are correct. In this case, the network must itself discover patterns, regularities, features or categories from the input data and relations for the input data over the output. While discovering all these features, the network undergoes change in its parameters. This process is called "self-organizing" in which exact clusters will be formed by discovering similarities and dissimilarities among the objects.

3) Reinforcement Learning (Learning with a Critic)

This learning process is similar to supervised learning.

In case of supervised learning, the correct target output values are known for each input pattern. But in some cases, less information might be available. For example, the network might be told that its actual output is only 50% correct or so. Thus, here only critic information is available, not the exact information. The learning based on this critic information is called reinforcement learning and the feedback sent is called reinforcement signal.



The feedback obtained here is only evaluative and not instrument. The external reinforcement signals are processed in the critic signal generator and the obtained critic signals are sent to the ANN for adjustments of weights so as to get a better critic feedback in future.

3. Using appropriate neural network, realize the given logic function (AND or OR). [Hint: use Perceptron]
4. Using appropriate neural network, realize ANDNOT function [Hint: use McCulloch Pitt's neuron model]

5. What is linear separability concept?

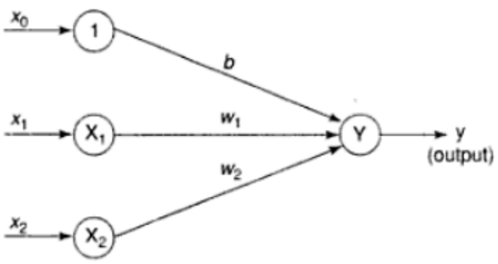
An ANN does not give exact solution for a nonlinear problem; however it provides possible approximation solutions.
Linear Separability is the concept where in the separation of the input space into regions is based on whether the network response is positive or negative.
A decision line is drawn to separate positive and negative responses. The decision line may also be called as decision support or linear-separable line.
This is required to classify the patterns based upon their output responses.
Generally, the net input calculated to the output unit is given as

$$y_{in} = b + \sum_{i=1}^{n} x_i w_i$$

There exists a boundary between the regions where $y_{in} > 0$ and $y_{in} < 0$. This region may be called as **decision boundary** and can be determined by the relation

$$b + \sum_{i=1}^{n} x_i w_i = 0$$

On the basis of the number of input units in the network, the above equation may represent a line, a plane or a hyperplane.
The Linear Separability of the network is based on the decision boundary line. If there exist weights(with bias) for which the training input vectors having positive response, +1, lie on one side of the decision boundary and all the other vectors having negative response, -1, lie on the other side of the decision boundary then we can conclude the problem is "linearly separable"
Consider a single layer network with bias included:



The net input for the network is
$Y_{in} = b + x_1w_1 + x_2w_2$
The separating line for which the boundary lies between values x_1 and x_2, so that the net gives a positive response on one side and negative response on other side is
$b + x_1w_1 + x_2w_2 = 0$
if $w_2 \neq 0$,

$$x_2 = -\frac{w_1}{w_2}x_1 - \frac{b}{w_2}$$

Thus the requirement for the positive response of the net is

$b + x_1w_1 + x_2w_2 > 0$

During training process, the values of w_1, w_2, b are determined so that the net will produce a positive (correct) response for the training data.

If threshold value is being used, then

Net input received > θ (threshold)

$y_{in} > \theta$

$x_1w1 + x_2w_2 > \theta$

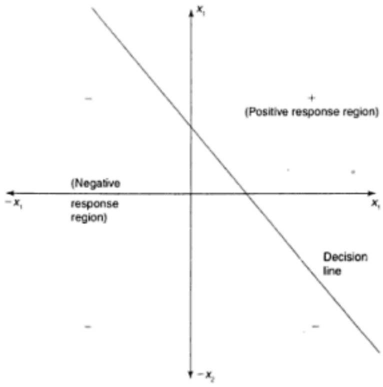
Then separating line equation will be

$x_1w_1 + x_2w_2 = \theta$

$$x_2 = -\frac{w_1}{w_2}x_1 + \frac{\theta}{w_2} \quad (\text{with } w_2 \neq 0)$$

During training process, the values of w1, w2, b have to be determined, so that the net will have a correct response to the training data.

The figure below shows a network having positive response in first quadrant and negative response in all other quadrants (AND function)



[In all networks, representation of data plays a major role. Bipolar representation is better than binary representation. Suing bipolar data representation, missing data can be distinguished from mistaken data. Missing values are represented by 0 and mistakes can be represented by reversing the input value from +1 to -1 or vice-versa]