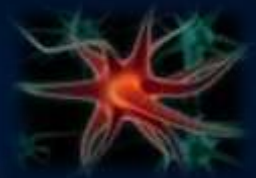



INTRODUCTION TO ARTIFICIAL NEURAL NETWORKS

CONTENTS

INTRODUCTION

- “Neural“ is an adjective for neuron, and “network” denotes a graph like structure.
- Artificial Neural Networks are also referred to as “neural nets” , “artificial neural systems”, “parallel distributed processing systems”, “connectionist systems”.
- For a computing systems to be called by these pretty names, it is necessary for the system to have a labeled directed graph structure where nodes performs some simple computations.
- “Directed Graph” consists of set of “nodes”(vertices) and a set of “connections”(edges/links/arcs) connecting pair of nodes.
- A graph is said to be “labeled graph” if each connection is associated with a label to identify some property of the connection

CONTD…

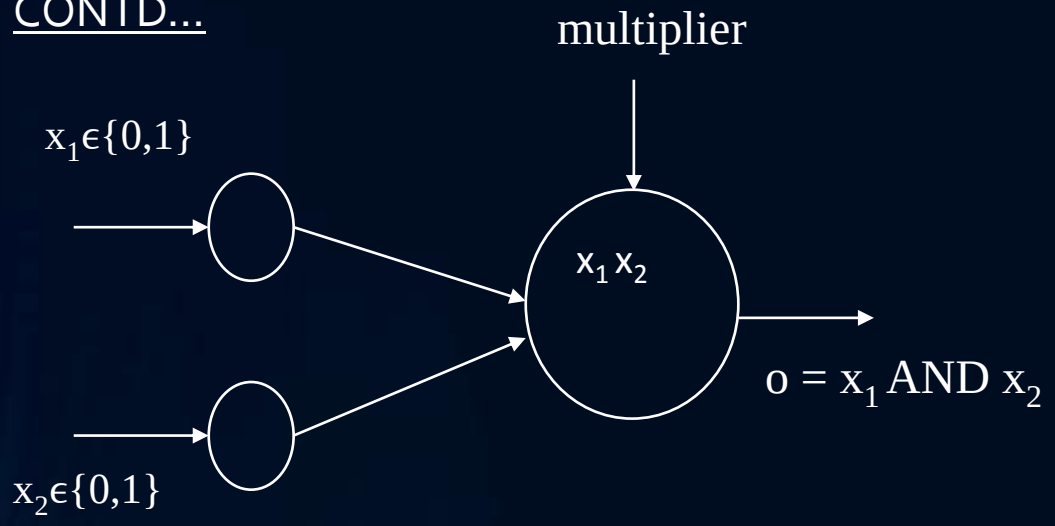


Fig 1: AND gate graph

This graph cannot be considered a neural network since the connections between the nodes are fixed and appear to play no other role than carrying the inputs to the node that computed their conjunction.

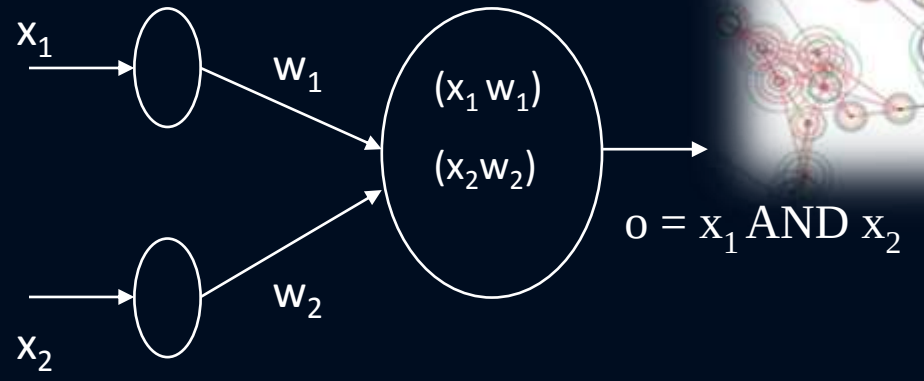


Fig 2: AND gate network

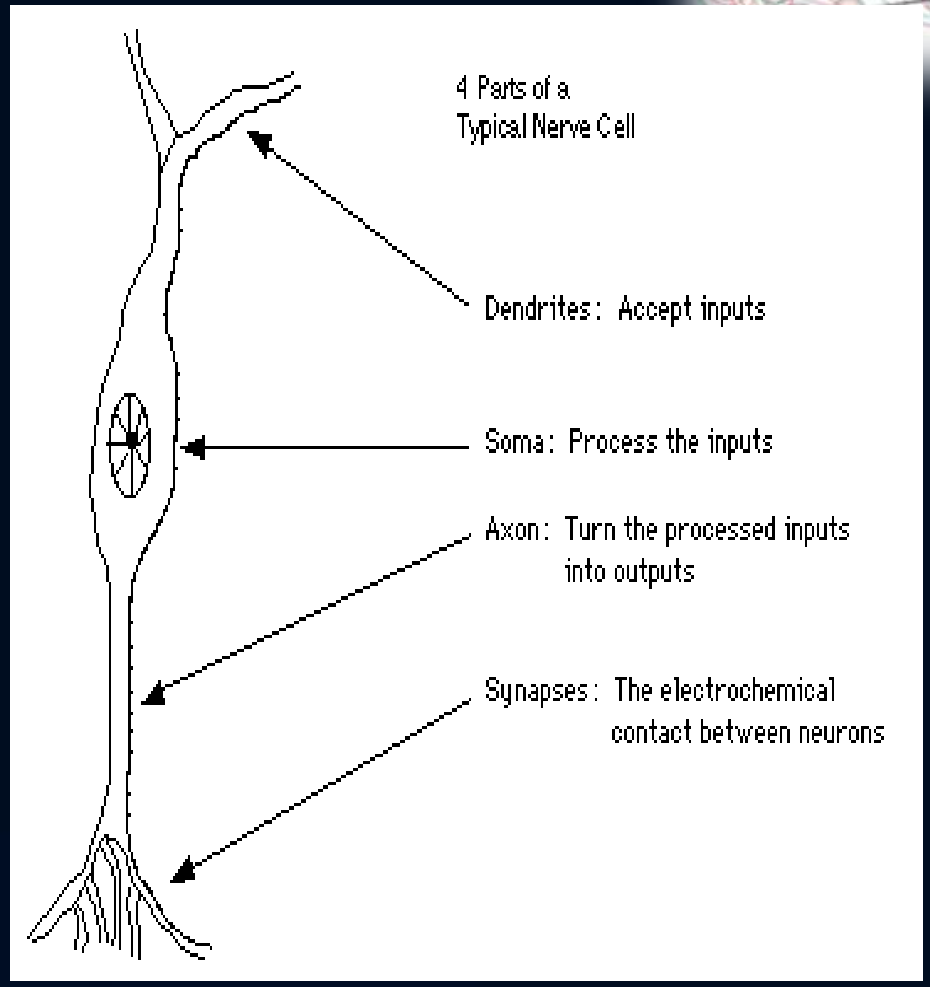
The graph structure which connects the weights modifiable using a learning algorithm, qualifies the computing system to be called an artificial neural networks.

- **The field of neural network was pioneered by BERNARD WIDROW of Stanford University in 1950's.**

BIOLOGICAL NEURON MODEL

Four parts of a typical nerve cell : -

- DENDRITES: Accepts the inputs
- SOMA : Process the inputs
- AXON : Turns the processed inputs into outputs.
- SYNAPSES : The electrochemical contact between the neurons.

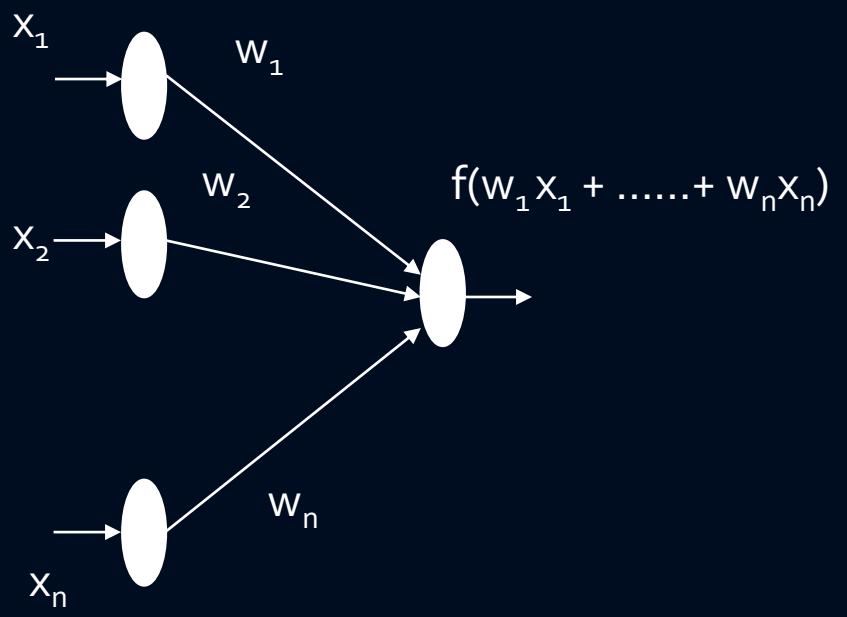


ARTIFICIAL NEURON MODEL

- Inputs to the network are represented by the mathematical symbol, x_n
- Each of these inputs are multiplied by a connection weight , w_n

$$sum = w_1 x_1 + \ldots\ldots + w_n x_n$$

- These products are simply summed, fed through the transfer function, f( ) to generate a result and then output.



TERMINOLOGY

Biological Terminology	Artificial Neural Network Terminology
Neuron	Node/Unit/Cell/Neurode
Synapse	Connection/Edge/Link
Synaptic Efficiency	Connection Strength/Weight
Firing frequency	Node output

ARTIFICIAL NEURAL NETWORK

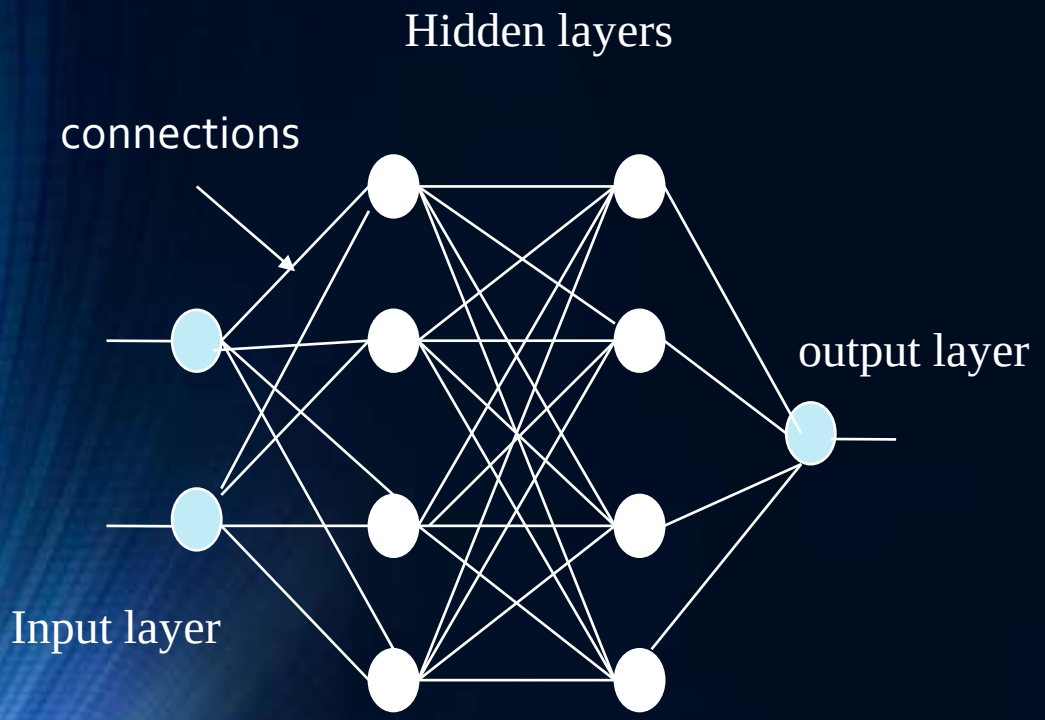
- **Artificial Neural Network (ANNs)** are programs designed to solve any problem by trying to mimic the structure and the function of our nervous system.
- Neural networks are based on simulated neurons, Which are joined together in a variety of ways to form networks.
- Neural network resembles the human brain in the following two ways: -

  * A neural network acquires knowledge through learning.

  *A neural network's knowledge is stored within the interconnection strengths known as synaptic weight.

CONTD...

ARTIFICIAL NEURAL NETWORK MODEL



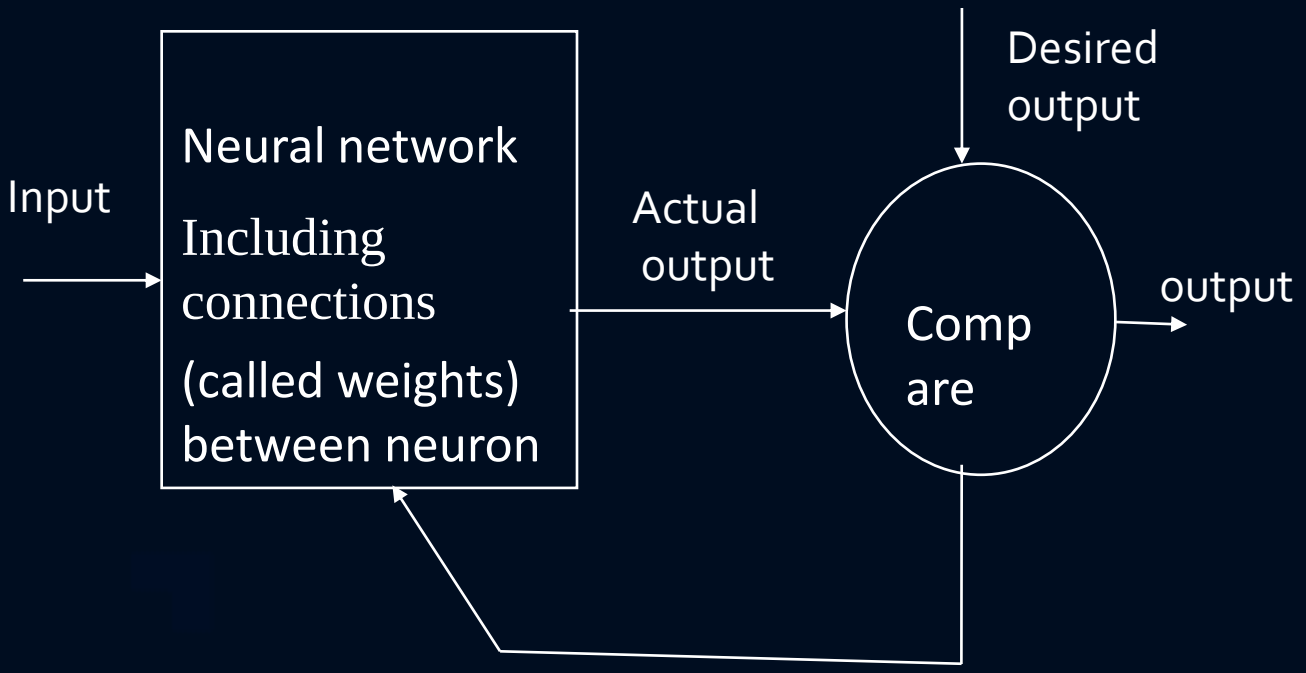


Fig 1 : artificial neural network model

Figure showing adjust of neural network

NEURAL NETWORK ARCHITECTURES

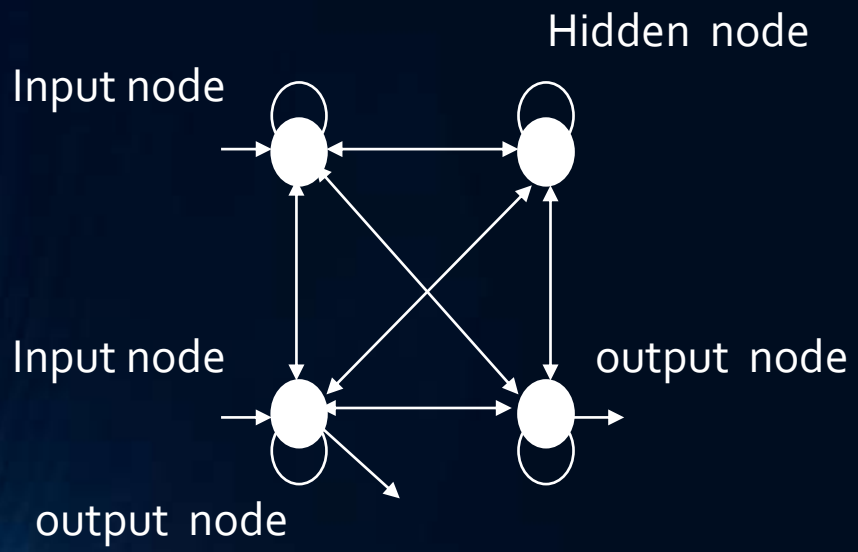


Fig: fully connected network

The neural network in which every node is connected to every other nodes, and these connections may be either excitatory (positive weights), inhibitory (negative weights), or irrelevant (almost zero weights).

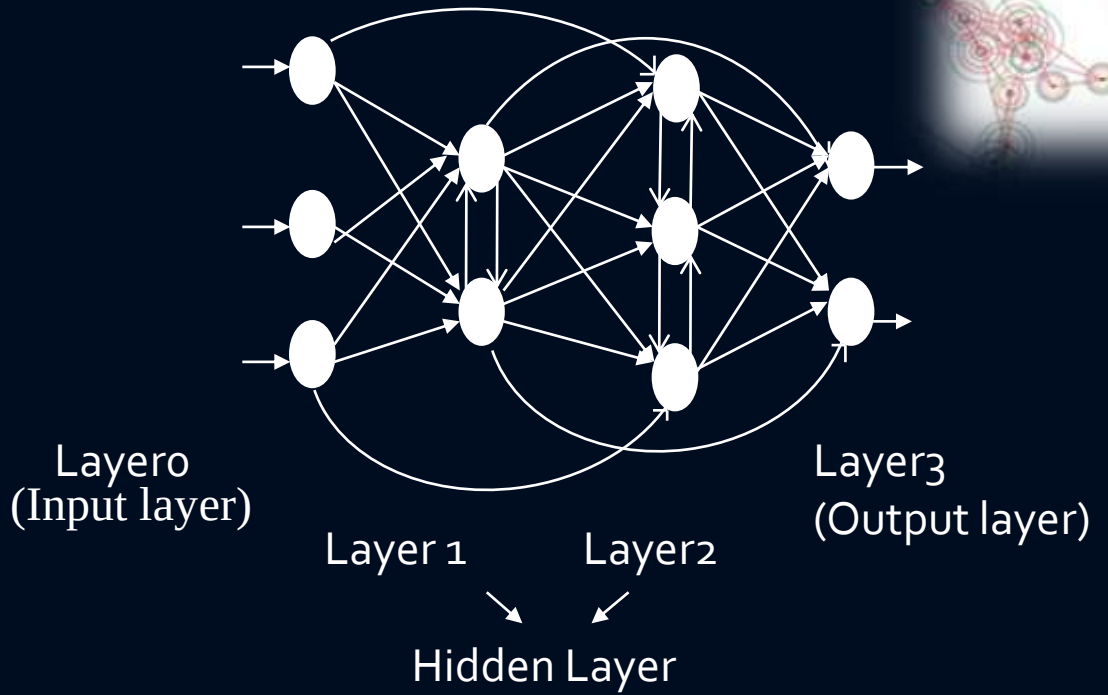


fig: layered network

These are networks in which nodes are partitioned into subsets called layers, with no connections from layer j to k if j > k.

CONTD...

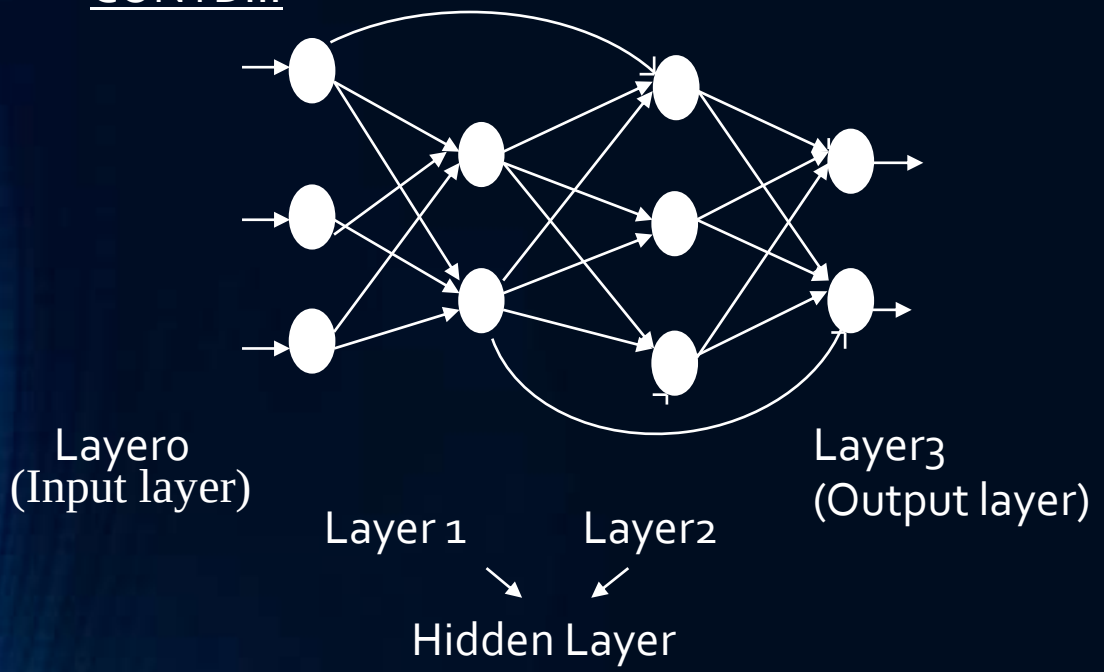


Fig : Acyclic network

This is the subclass of the layered networks in which there is no intra-layer connections. In other words, a connection may exist between any node in layer i and any node in layer j for i < j, but a connection is not allowed for i=j.

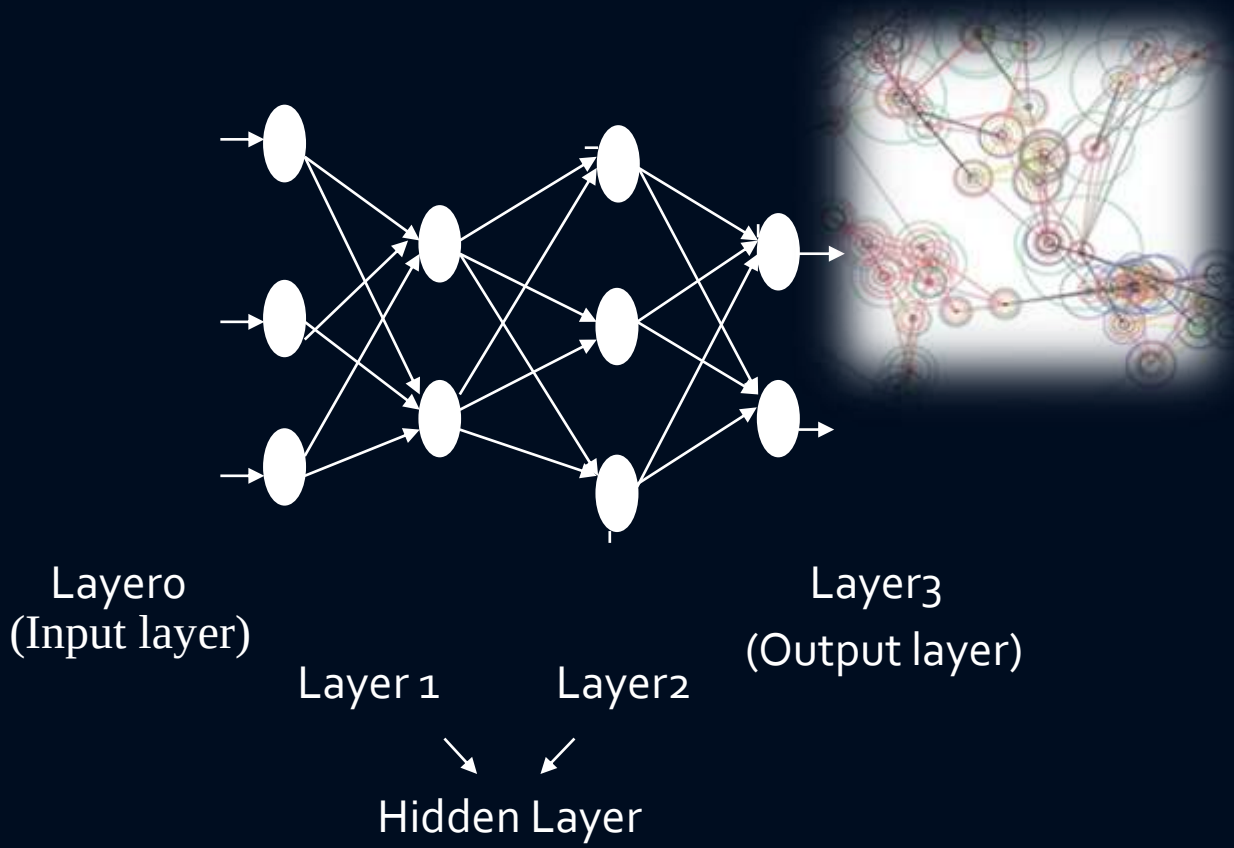


fig : Feedforward network

This is a subclass of acyclic networks in which a connection is allowed from a node in layer i only to nodes in layer i+1

CONTD...

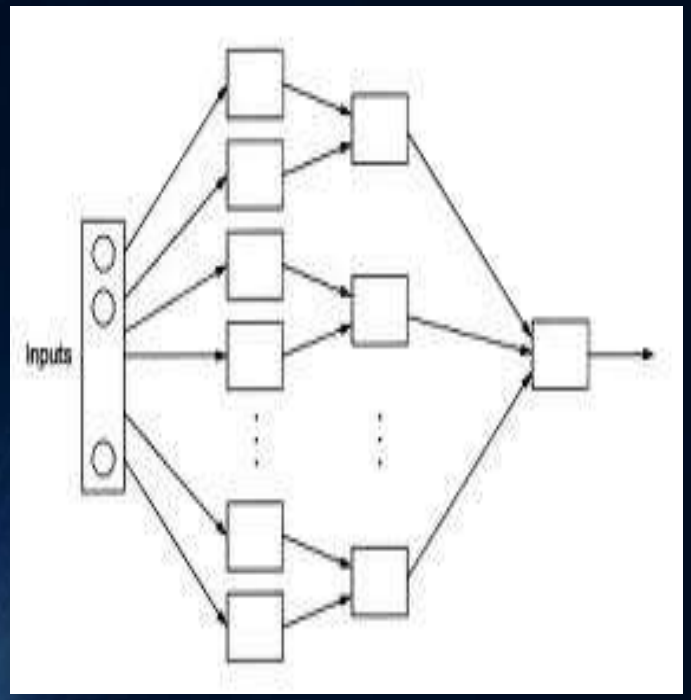


Fig : Modular neural network

Many problems are best solved using neural networks whose architecture consists of several modules, with sparse interconnections between them. Modules can be organized in several different ways as Hierarchial organization, Successive refinement, Input modularity

LEARNING

- Neurons in an animal's brain are "hard wired". It is equally obvious that animals, especially higher order animals, learn as they grow.
- How does this learning occur?
- What are possible mathematical models of learning?
- In artificial neural networks, learning refers to the method of modifying the weights of connections between the nodes of a specified network.
- The learning ability of a neural network is determined by its architecture and by the algorithmic method chosen for training.

CONTD…

SUPERVISED LEARNING

- A teacher is available to indicate whether a system is performing correctly, or to indicate the amount of error in system performance. Here a teacher is a set of training data.
- The training data consist of pairs of input and desired output values that are traditionally represented in data vectors.
- Supervised learning can also be referred as classification, where we have a wide range of classifiers, (Multilayer perceptron, k nearest neighbor..etc)

UNSUPERVISED LEARNING

- This is learning by doing.
- In this approach no sample outputs are provided to the network against which it can measure its predictive performance for a given vector of inputs.
- One common form of unsupervised learning is clustering where we try to categorize data in different clusters by their similarity.

THE BACKPROPAGATION ALGORITHM

- The backpropagation algorithm (Rumelhart and McClelland, 1986) is used in layered feed-forward Artificial Neural Networks.
- Back propagation is a multi-layer feed forward, supervised learning network based on gradient descent learning rule.
- we provide the algorithm with examples of the inputs and outputs we want the network to compute, and then the error (difference between actual and expected results) is calculated.
- The idea of the backpropagation algorithm is to reduce this error, until the Artificial Neural Network *learns* the training data.

- The activation function of the artificial neurons in ANNs implementing the backpropagation algorithm is a weighted sum (the sum of the inputs x_i multiplied by their respective weights w_{ji})

$$A_j(\bar{x}, \bar{w}) = \sum_{i=0}^{n} x_i w_{ji}$$

- The most common output function is the sigmoidal function:

$$O_j(\bar{x}, \bar{w}) = \frac{1}{1 + e^{A_j(\bar{x}, \bar{w})}}$$

- Since the error is the difference between the actual and the desired output, the error depends on the weights, and we need to adjust the weights in order to minimize the error. We can define the error function for the output of each neuron:

$$E_j(\bar{x}, \bar{w}, d) = \left(O_j(\bar{x}, \bar{w}) - d_j\right)^2$$

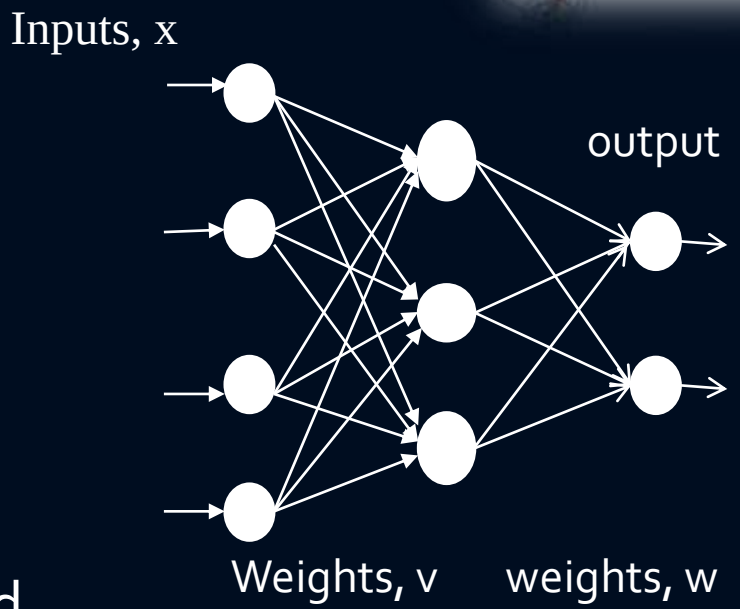


Fig: Basic Block of
Back propagation neural network

CONTD…

- The backpropagation algorithm now calculates how the error depends on th output, inputs, and weights.

$$\Delta w_{ji} = -\eta \frac{\partial E}{\partial w_{ji}}$$

the adjustment of each weight (Δwji) will be the negative of a constant eta (η) multiplied by the dependance of the "wji" previous weight on the error of the network.

- First, we need to calculate how much the error depends on the output

$$\frac{\partial E}{\partial O_j} = 2(O_j - d_j)$$

- Next, how much the output depends on the activation, which in turn depends on the weights

$$\frac{\partial O_j}{\partial w_{ji}} = \frac{\partial O_j}{\partial A_j}\frac{\partial A_j}{\partial w_{ji}} = O_j(1-O_j)x_i$$

- And so, the adjustment to each weight will be

$$\Delta w_{ji} = -2\eta(O_j - d_j)O_j(1-O_j)x_i$$

CONTD…

- If we want to adjust v_{ik}, the weights (let's call them v_{ik}) of a previous layer, we need first to calculate how the error depends not on the weight, but in the input from the previous layer i.e. replacing w by x as shown in below equation.

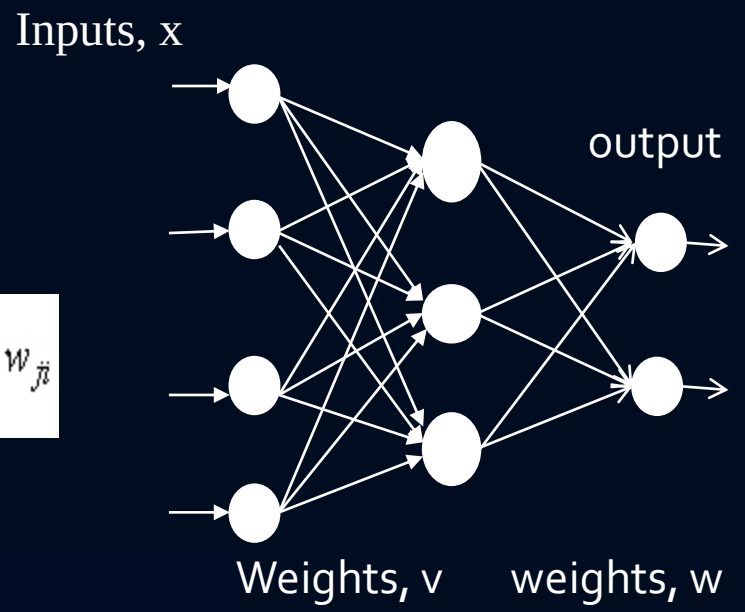
$$\Delta v_{ik} = -\eta \frac{\partial E}{\partial v_{ik}} = -\eta \frac{\partial E}{\partial x_i} \frac{\partial x_i}{\partial v_{ik}}$$

where

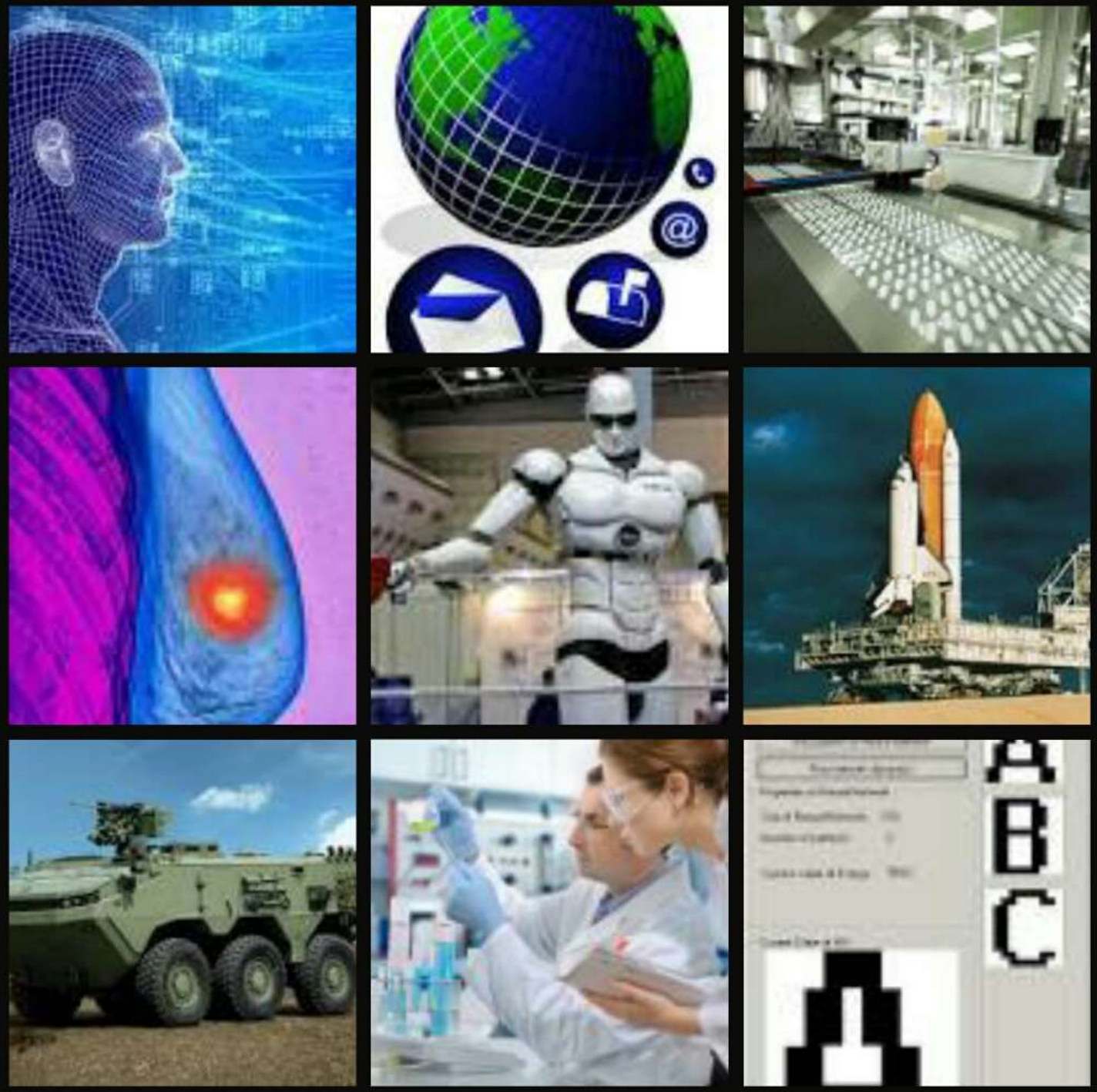
$$\frac{\partial E}{\partial w_{ji}} = 2\left(O_j - d_j\right) O_j (1 - O_j) w_{ji}$$

- and

$$\frac{\partial x_i}{\partial v_{ik}} = x_i (1 - x_i) v_{ik}$$



NEURAL NETWORK APPLICATIONS



ADVANTAGES

- It involves human like thinking.
- They handle noisy or missing data.
- They can work with large number of variables or parameters.
- They provide general solutions with good predictive accuracy.
- System has got property of continuous learning.
- They deal with the non-linearity in the world in which we live.

CONCLUSION


Thank you!