Unit-1
Introduction to Neural Network

By: Arjun Saud
CSCSIT, TU

What is ANN?

- Artificial Neural Network (ANN) is commonly referred as Neural Network (NN). It is the computational paradigm that is motivated from the way the computation is performed by human brain or nervous system.
- Brain is a highly complex, non-linear, and parallel computation system that can perform computations like perception, pattern recognition, motor control etc. Neuron or nerve cell is the basic structural unit of brain.

What is ANN?

- Human can perform the task much faster than the fastest digital computers that exists todays. This is possible due to parallel computation of neurons interconnected with each other.
- Thus we can define ANN as *"It is a massively parallel distributed processing system made up of simple processing units that has capability of storing experiential knowledge and making it available for use."*
- ANNs perform useful computations through the process of learning by using some algorithm.

Benefits of ANN

- **Non-Linearity:** An artificial neuron can be linear or non-linear. A neural network made up of an interconnection of nonlinear neurons, is itself nonlinear.
- **Input-Output Mapping**: ANNs can learn from given set of input-output examples (training set) and can apply the experience to predict output for some unknown inputs (test set).
- **Adaptivity:** Neural networks can learn in real-time, and can adapt to changing environments flexibly. This is possible by changing weights of interconnections between neurons.

Benefits of ANN

- **Evidential Response:** A neural network can be designed to provide information not only about which particular pattern to select, but also about the confidence in the decision made.
- **Contextual Information**: Every neuron in a neural network is affected by activity of other neurons in the network. Thus, ANNs have capability of dealing with contextual information.

Benefits of ANN

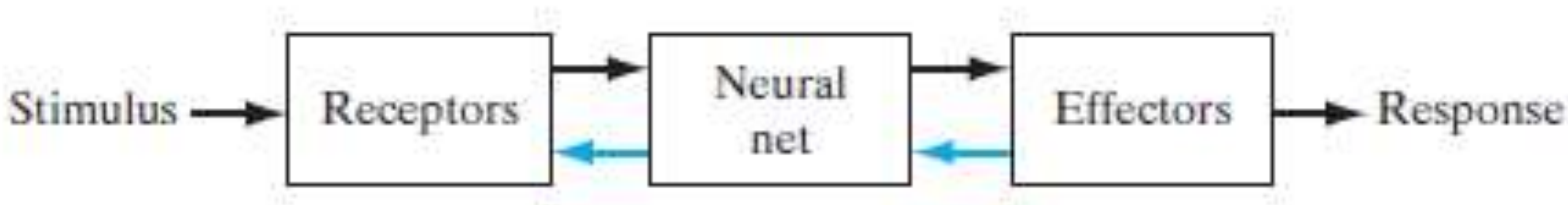
- **Fault Tolerant:** The whole network can be very reliable although each element is weak. When a neuron or its connection link is failed, the entire neural network system still works but with degraded performance.
- **VLSI Implementability**: The massively parallel nature of a neural network makes it potentially fast for the computation. This feature makes a neural network well suited for implementation using very-large-scale-integrated (VLSI) technology.

Benefits of ANN

- **Uniformity of Analysis and Design:** Same notations, theories, and learning algorithms of neural networks can be applied to different applications of neural networks.
- **Neurobiological Analogy:** Analogy of ANNs with human brain is useful to neurobiologists for interpretation of neurobiological phenomenon and it is also useful tool for engineers to solve complex problems on the basis of neurobiological ideas.

Human Brain

- The human nervous system may be viewed as a three-stage system, as depicted in the block diagram of figure given below.



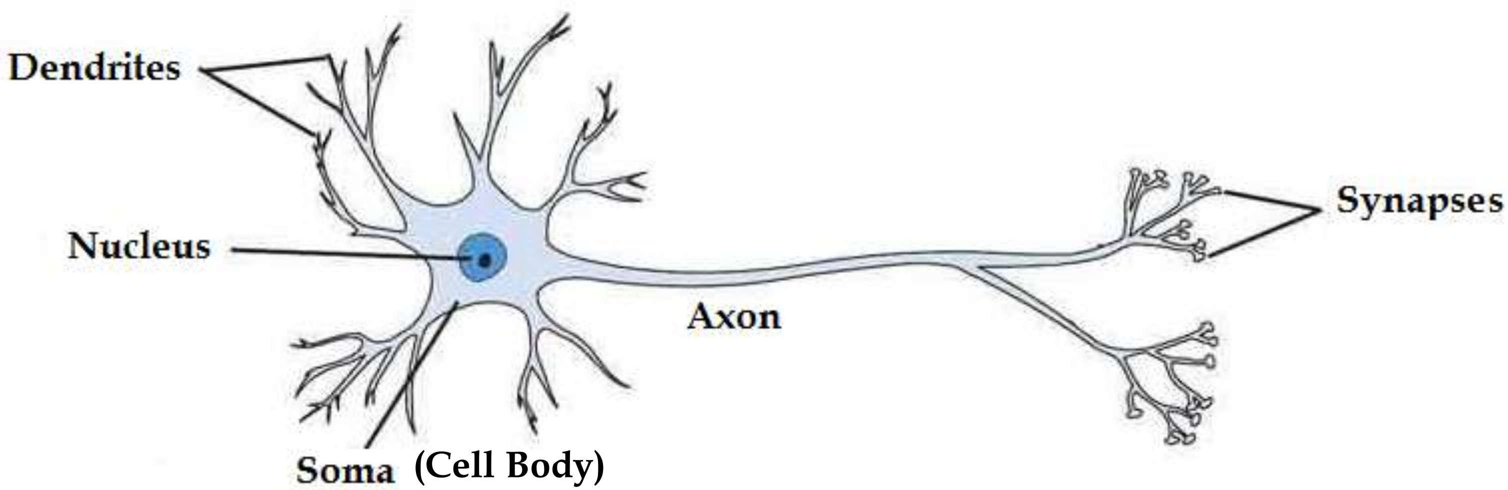
- **Neural Net (Brain):** It is the command center for the human nervous system. It receives signals from the body's sensory organs, make necessary decisions, and outputs information to the muscles. Two sets of arrow represents forward transmission of information and feedback.

Human Brain

- **Receptors**: The *receptors* convert stimuli from the human body or the external environment into electrical impulses that convey information to the neural net (brain).
- **Effectors:** The *effectors* convert electrical impulses generated by the neural net into discernible responses as system outputs.
- It is estimated that there are approximately 10 billion neurons in the human cortex, and 60 trillion synapses or connections

Human Brain

- Neurons come in a wide variety of shapes and sizes in different parts of the brain. Structure of most common type of nerve cell is given the figure below.



Human Brain

- A human neuron contains:
  - a cell body for signal processing,
  - many dendrites to receive signals,
  - an axon for outputting the result, and
  - synapses between the axon and other dendrites

Human Brain

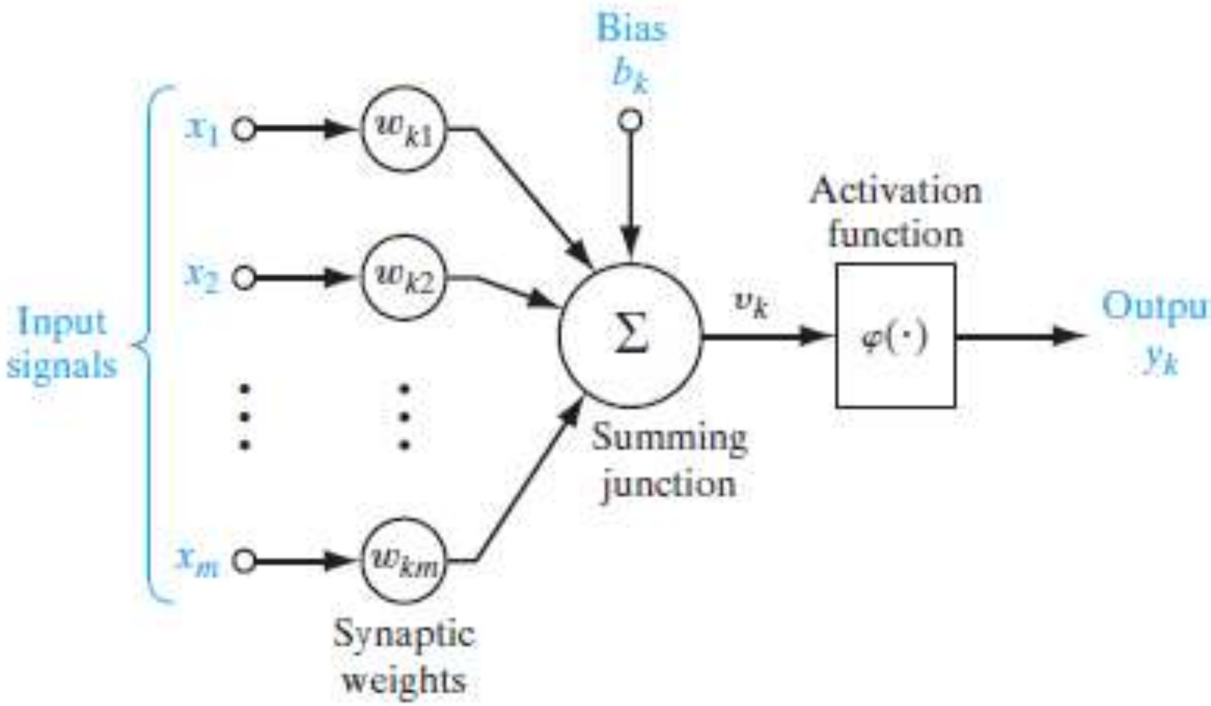
- **Working of human nerve cell (neuron)**
  - Signals (impulses) come into the dendrites through the synapses.
  - All signals from all dendrites are summed up in the cell body
  - When the sum is larger than a threshold, the neuron fires, and sends out an impulse signal to other neurons through the axon.

Models of Neuron

- A *neuron* is an information-processing unit that is fundamental to the operation of a neural network.
- Basically, Models of neuron can be divided into two categories:
  - **Deterministic model of Neuron**
  - **Stochastic model of Neuron**

Models of Neuron

Deterministic Model of Neuron



Models of Neuron

Deterministic Model of Neuron

- The block diagram this *model* of a neuron is presented in previous slide. Three basic elements of this neural model are: *Synapses or connecting links, Adder, and Activation Function.*
  - **Synapses:** These are the connecting links that are used to collect input for the neuron. Each link is characterized by weight that defines strength of the link.
  - **Adder:** It is responsible for finding weighted sum of inputs to the neuron.
  - **Activation Function:** It is responsible for finding output of the neuron. It is also referred as squashing function.

Models of Neuron

Deterministic Model of Neuron

- The neural model also includes an externally applied *bias*, denoted by b_k. The bias b_k has the effect of increasing or lowering the net input. In mathematical terms, we may describe the neural model by writing the set of equations:

$$u_k = \sum_{j=1}^{n} x_j * w_{kj} \qquad v_k = u_k + b_k$$

$$y_k = \varphi(u_k + b_k) = \varphi(v_k)$$

Where, $x_1, x_2, \ldots, x_n$ *are input signals,* $w_{k1}, w_{k2}, \ldots, w_{kn}$ *are weights,* u_k *is weighted sum of inputs,* φ *is activation function, and* y_k *is output signal*

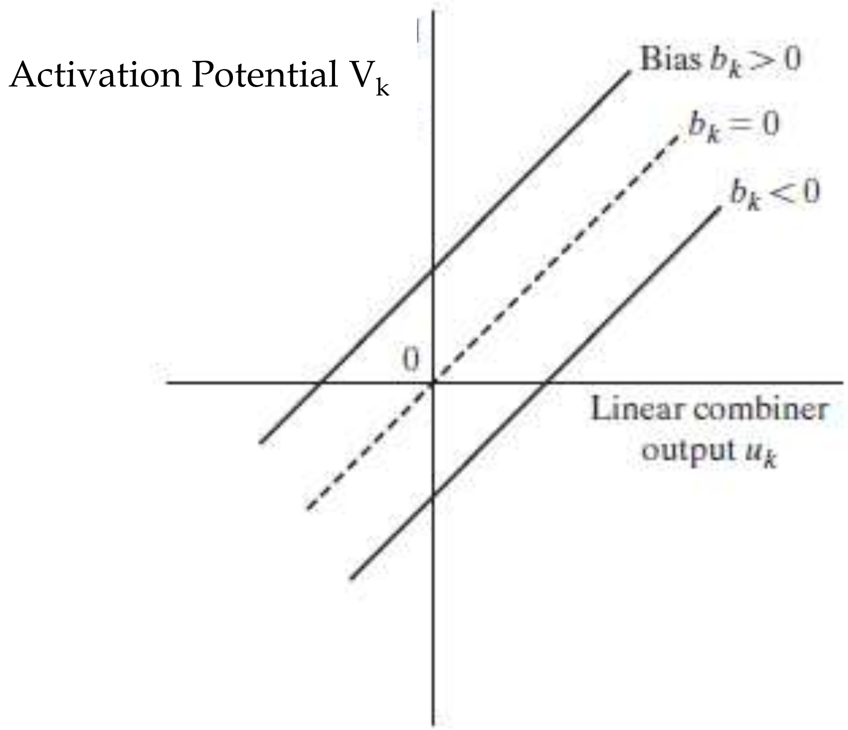
Models of Neuron

Deterministic Model of Neuron

- The use of bias b_k has the effect of applying an affine transformation to the output u_k of the linear combiner in the neural model.
- Depending on whether the bias b_k is positive or negative, the relationship between the *activation potential* (v_k) and the linear combiner output (u_k) is modified as below.

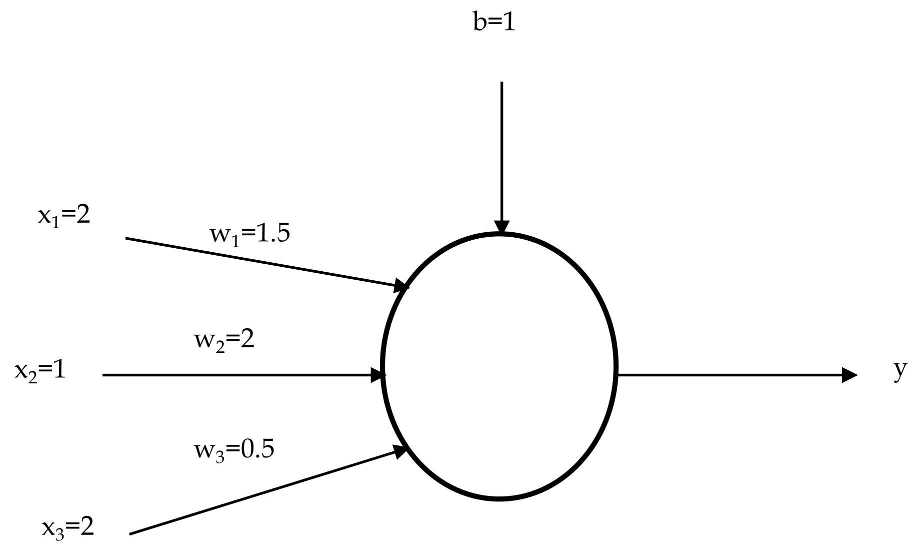
Models of Neuron

Deterministic Model of Neuron



Models of Neuron

Example: *Consider following neuron and compute its output by assume activation function F(x)=1 if x>5 and F(x)=0, otherwise*



$u=x_1*w_1+x_2*w_2+x_3*w_3$

$=2*1.5+1*2+2*0.5=6$

$v=u+b=6+1=7$

Now,

$y=f(v)=1$

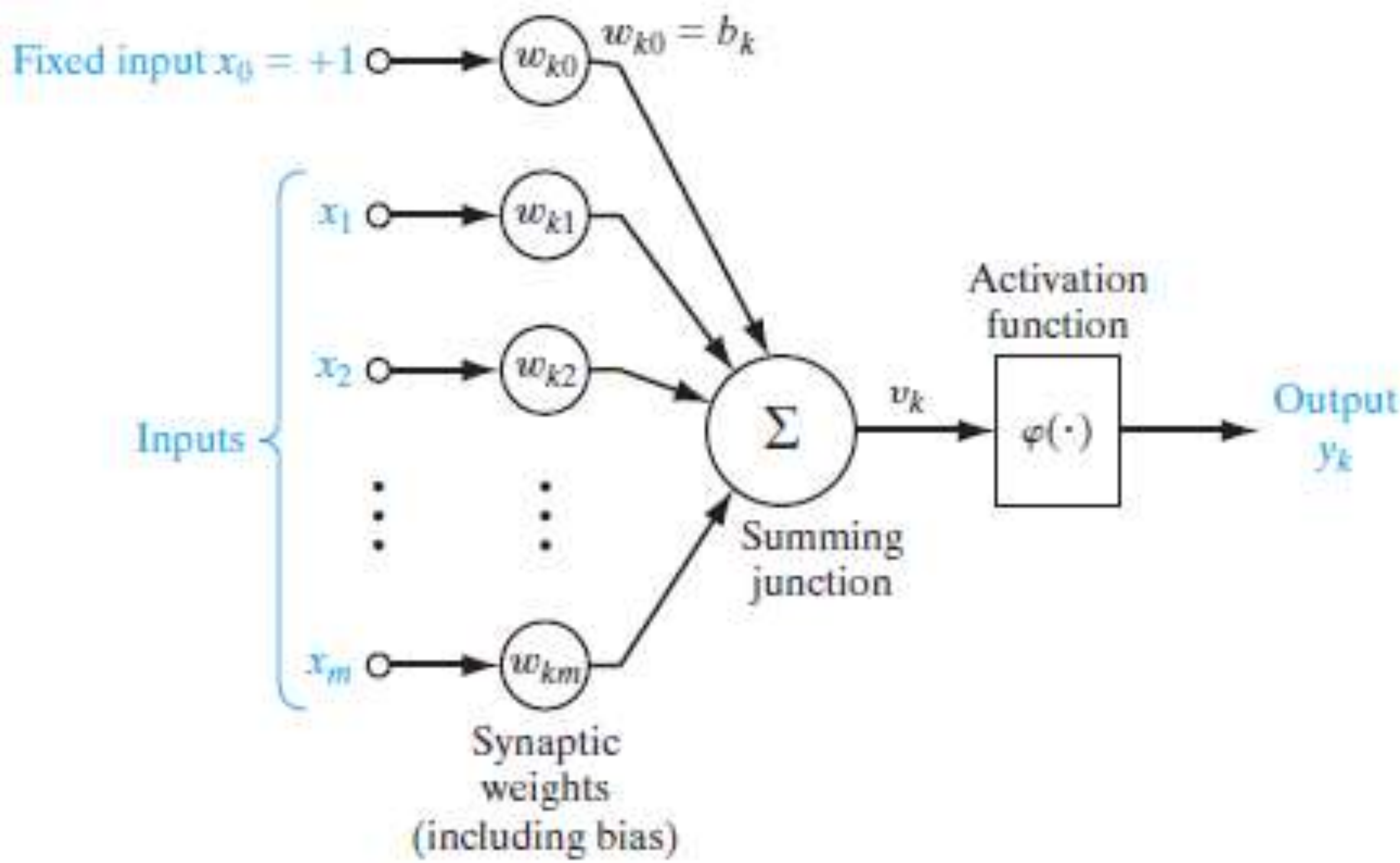
Models of Neuron

Deterministic Model of Neuron

- We can reformulate the model of neuron by doing two things:
  1. Adding a new input signal fixed at 1, and
  2. Adding a new synaptic weight equal to the bias *bk.*
- Although the two models are different in appearance, they are mathematically equivalent.

Models of Neuron

Deterministic Model of Neuron



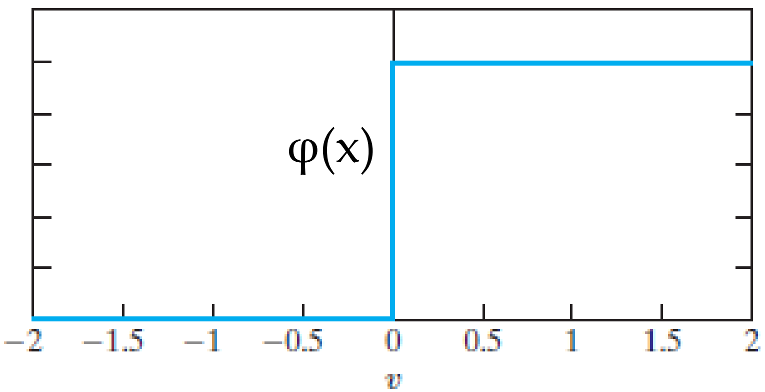
Activation Functions

- Activation functions are the functions responsible to convert a input signal of a node in a ANN to an output signal.
- The activation function is the non linear transformation that we do over the input signal. This transformed output is then sent to the next layer of neurons as input.
- Some widely used activation functions are: *Threshold, linear, sigmoid, tanh,* etc.

Activation Functions

Threshold Function

$$\varphi(x)=\begin{cases}1 & if \quad x\geq c\\ 0 & if \quad x<c\end{cases}$$

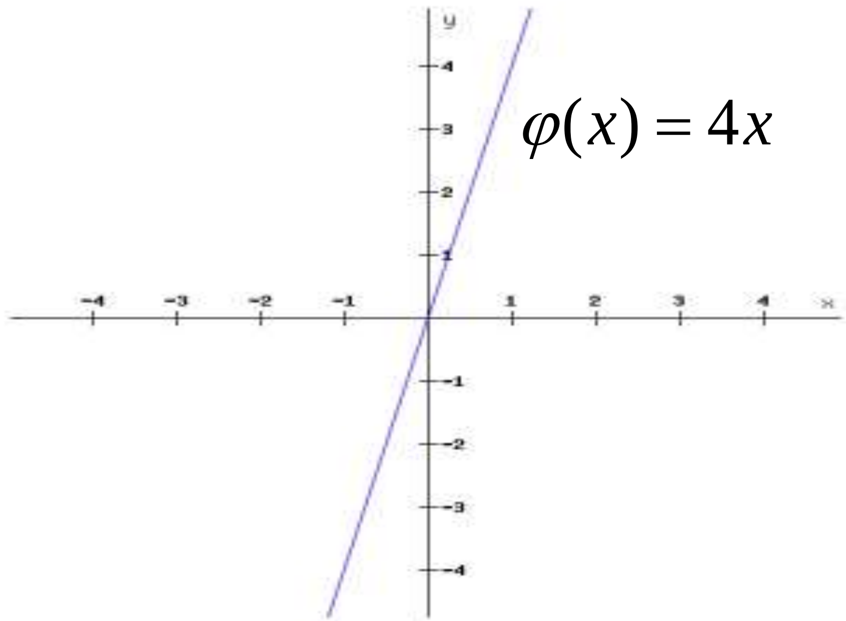


- It is also refered as Healviside function. The non-linear neural model that use threshold function as activation function is referred as the *McCulloch–Pitts model.*

Activation Functions

<u>Linear Function</u>

$$f(x) = ax + b$$

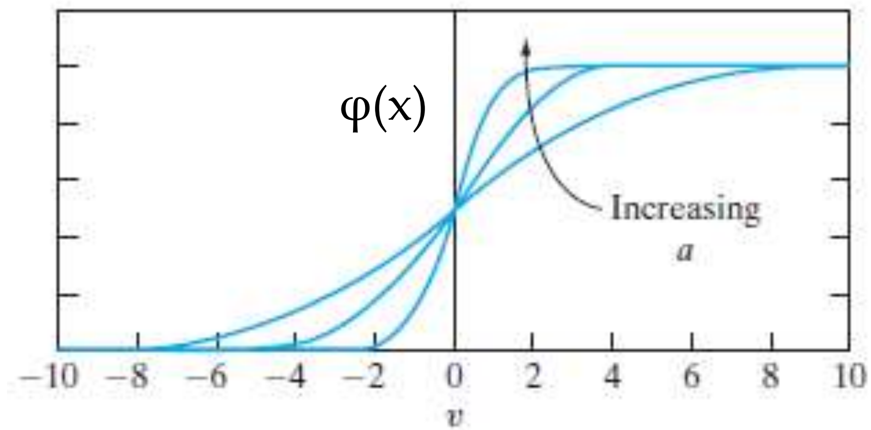


$$\varphi(x) = 4x$$

Activation Functions

Sigmoid Function

$$\varphi(x) = \frac{1}{1+e^{-ax}}$$

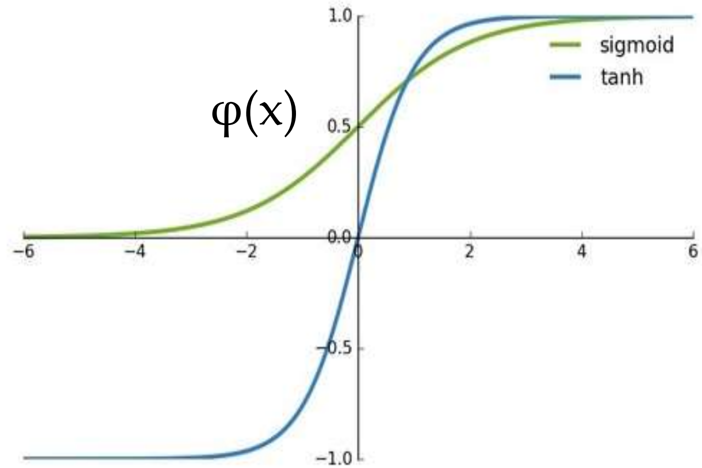


- By varying the parameter *a (slope)*, we obtain sigmoid functions of different slopes. The sigmoid function is the class of functions whose graph is S-shaped curve. It is the most common form of activation function used in the construction of neural networks. An example of the sigmoid function is the *logistic function, where a=1*. It squashes the output in the range(0,1).

Activation Functions

Tanh Function

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



- It has characteristics similar to sigmoid that we discussed above. But, it squashes the output between (-1,1). Tanh is also a very popular and widely used activation function. It is special case of sigmoid function.

Models of Neuron Contd…

Stochastic Model of Neuron

- The deterministic neural model defines input–output behavior precisely for all inputs. However, stochastic model of neuron makes input-output behavior non-deterministic.
- Stochastic neural model achieves this by giving probabilistic interpretation to the activation function used in deterministic neural model.
- Specifically, a neuron is permitted to reside in only one of two states:+1 (ON) or -1 (OFF). The decision for a neuron to *fire* is probabilistic. Let x denote the state of the neuron and P(v) denote the *probability* of firing, where v is the activation potential of the neuron.

Models of Neuron Contd…

Stochastic Model of Neuron

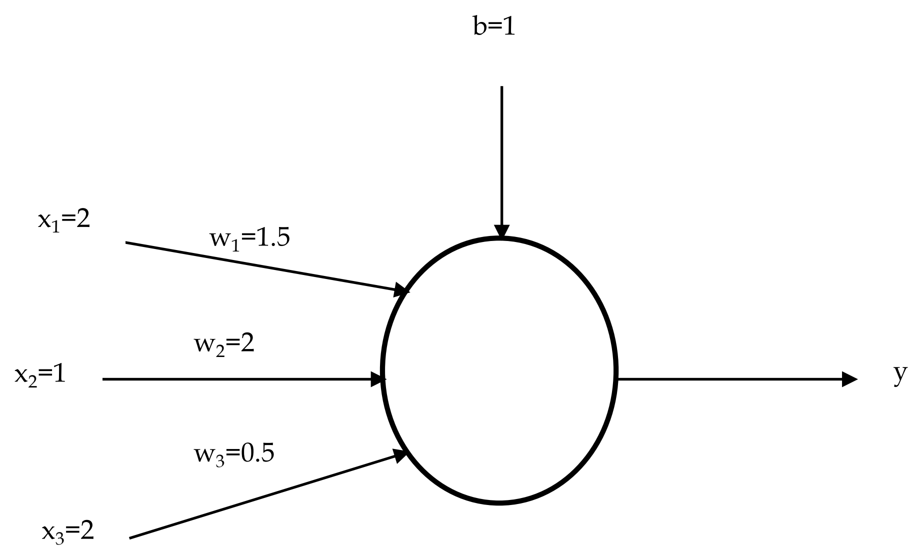
$$x = \begin{cases} +1 & \text{with probabilit y p(v)} \\ -1 & \text{with probabilit y 1 - p(v)} \end{cases}$$

$$p(v) = \frac{1}{1+e^{-v/T}}, \text{where T is parameter to control noise level}$$

- This adds uncertainty to firing of neuron and hence makes the input-output behavior stochastic. Rest of things in stochastic model of neuron is similar to the deterministic model.

Models of Neuron Contd..

Example: *Consider following stochastic neuron and compute its probability of firing by assuming T=5*



$u=x_1*w_1+x_2*w_2+x_3*w_3$

$=2*1.5+1*2+2*0.5=6$

$v=u+b=6+1=7$

Now,

$P(v)=1/(1+e^{-v/T})$

$=1/(1+e^{-7/5})=0.802$

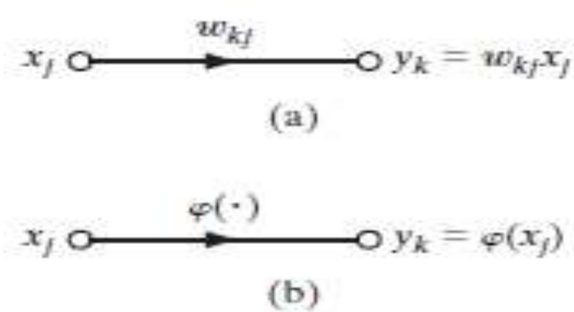
Thus, the probability of firing the neuron is 0.802

Neural Network as Directed Graph

- Signal flow graph provides neat method of representing the flow of signals in neural network. Signal flow graph is a network of directed links that are connected at some nodes.
- A node j in the signal flow graph has associated node signal x_j. A directed link originates at node j and terminates at node k and the link has associated transfer function that defines how the signal y_k at node k depends upon signal x_j at node j.

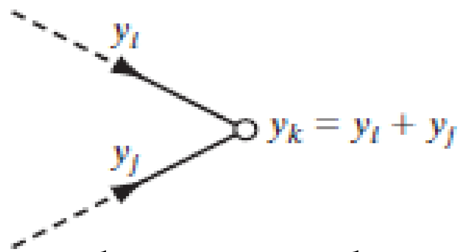
Neural Network as Directed Graph

- Flowing three rules defines the flow of signals in signal flow graph.
  - **Rule 1.** A signal flows along a link only in the direction defined by the arrow on the link. Links can be divided into two categories: *Synaptic Links, Activation Links*.

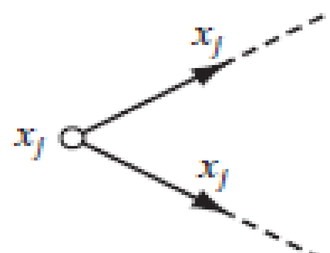


Neural Network as Directed Graph

- **Rule 2.** A node signal equals the algebraic sum of all signals entering the pertinent node via the incoming links.



- **Rule 3.** The signal at a node is transmitted to each outgoing link originating from that node, with the transmission being entirely independent of the transfer functions of the outgoing links.

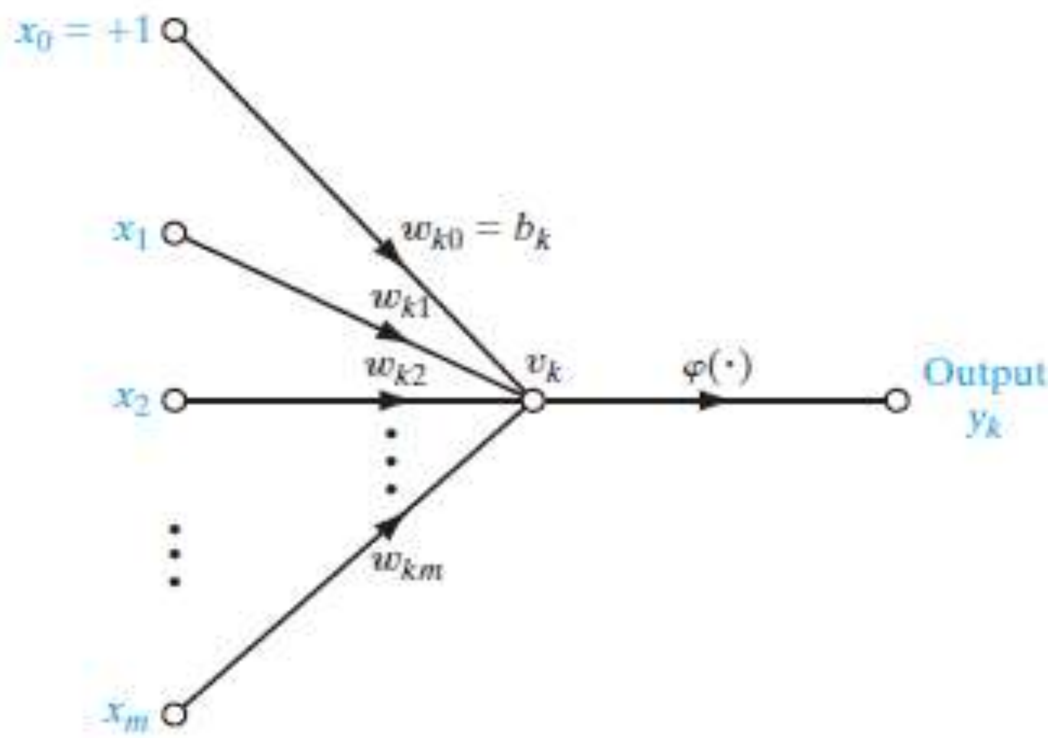


Neural Network as Directed Graph

- Now we may offer the following mathematical definition of a neural network: *A neural network is a directed graph consisting of nodes with interconnecting synaptic and activation links and is characterized by four properties*:
  1. Each neuron is represented by a set of linear synaptic links, bias, and a possibly nonlinear activation link.
  2. The synaptic links of a neuron weight their respective input signals.
  3. The weighted sum of the input signals defines the activation potential of the neuron.
  4. The activation link squashes the activation potential of the neuron to produce an output.

Neural Network as Directed Graph

Example



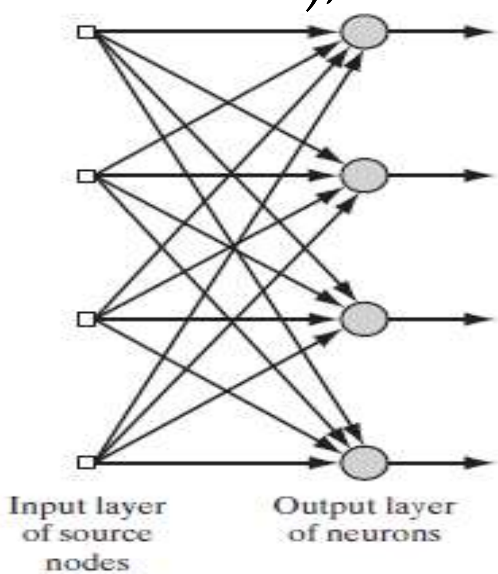
Neural Network Architectures

- The manner in which neurons of a neural network are structured is called neural network architecture. Broadly, we can divide neural network architectures into three categories.
  - Single-Layer Feedforward Networks
  - Multi-Layer Feedforward Networks
  - Recurrent Networks

Neural Network Architectures

Single-Layer Feedforward Networks

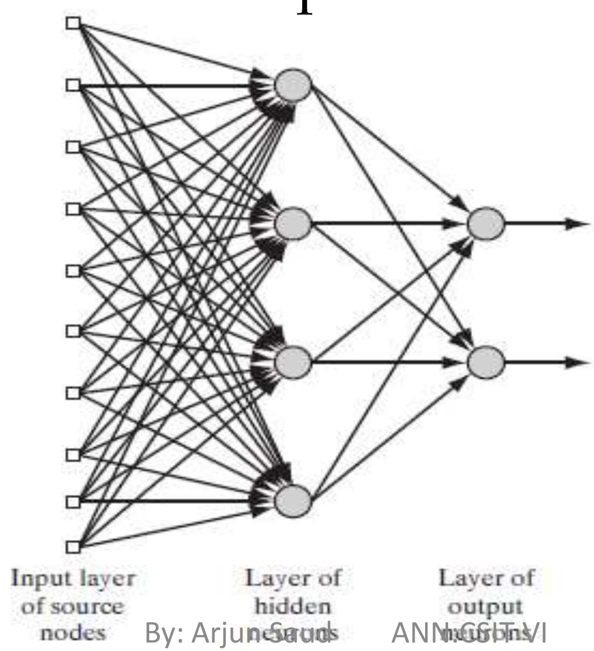
- It is the simplest form of a network architecture. In this architecture we have an *input layer* of source nodes that are connected directly with an *output layer* of neurons (computation nodes), but not vice versa.



Neural Network Architectures

Multi-Layer Feedforward Networks

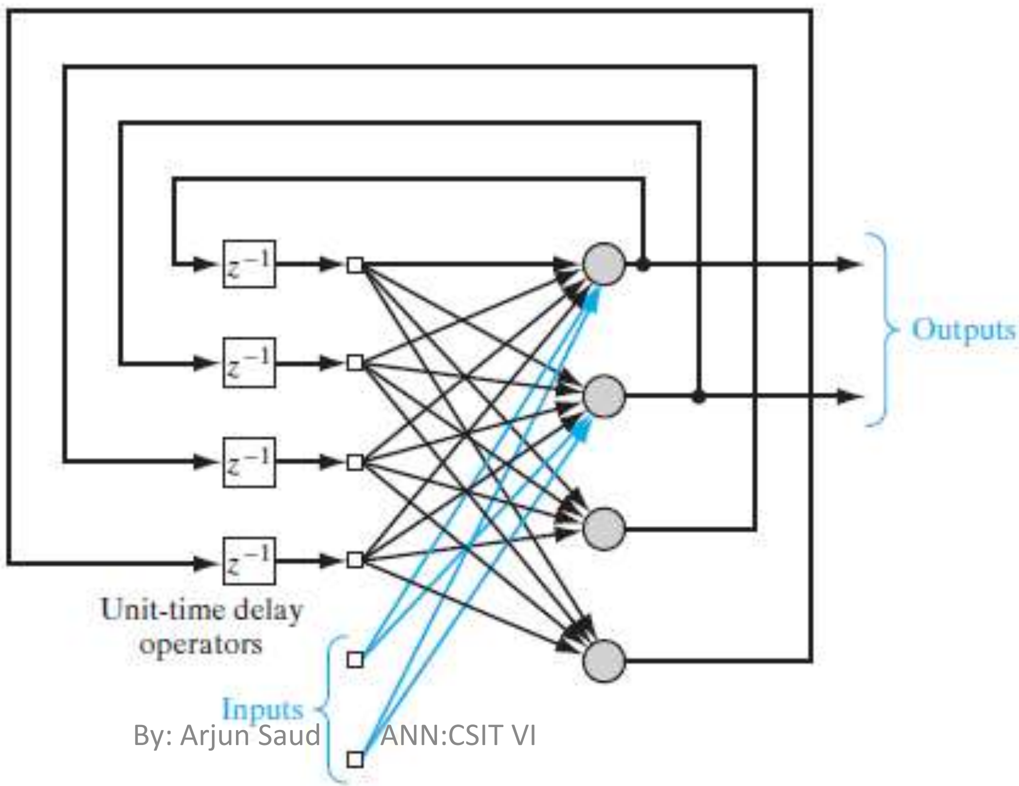
- In this type of network architecture, one or more hidden layers are present between input and output layers. These layers are not directly visible and information only flows in the direction of input to output layer. This type of network is designed to extract higher order statistics from input.



Neural Network Architectures

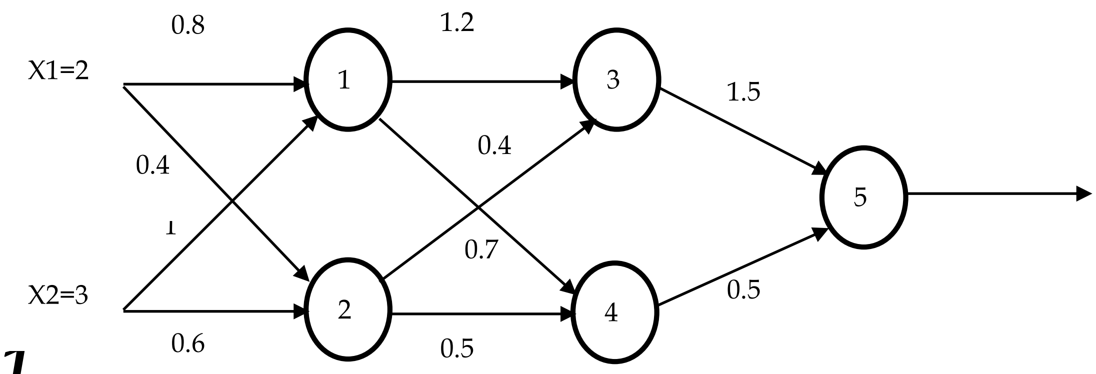
Recurrent Networks

- This class of networks are also referred ad feedback neural networks. This type of network architecture may contain feedback loop that can feed output of neuron as input to neurons of previous layers.



Neural Network Architectures

<u>Example:</u> *Consider following Neural Network and compute its output using activation function F(x)=2x-1. Weights of synaptic links are provided above each link.*



<u>For Node 1</u>

u1=2*0.8+3*1=4.6 => y1=f(u1)=2*4.6-1=8.2

<u>For Node 2</u>

u2=2*0.4+3*0.6=2.6 => y2=f(u2)=4.2

Neural Network Architectures

For Node 3

*u3=8.2*1.2+4.2*0.4=11.51* =>*y3=f(u3)=22.04*

For Node 4

*u4=8.2*0.7+4.2*0.5=7.84* =>*y4=f(u4)=14.68*

For Node 5

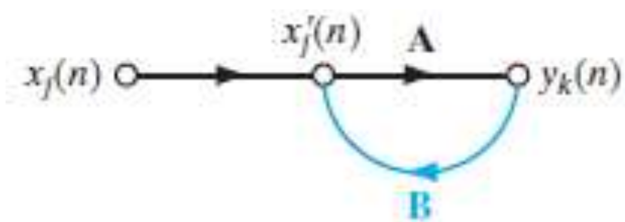
*u5=22.04*1.5+14.68*0.5=40.4* => *y5=f(u5)=79.8*

Thus,

Final output of the neural network (y)=79.8

Feedback

- When output of a system is again applied to the system as input, it is called feedback. Thus, there exists closed path for the transmission of signals in the systems with feedback.
- Feedbacks present in special class of neural network called recurrent neural network.
- Let us consider a signal flow graph of single loop feedback system as below, where operators A and B represents feed-forward and feedback paths respectively.



Feedback

- We can express input-output relationships of above signal flow graph as below.

$$y_k(n) = A[x'_j(n)] \qquad (1)$$

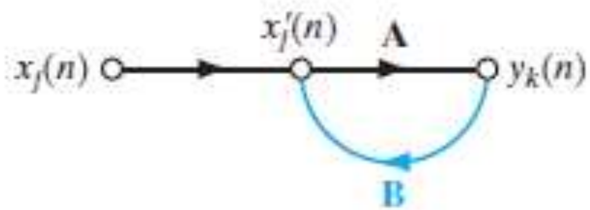
$$x'_j(n) = x_j(n) + B[y_k(n)] \qquad (2)$$

- Putting Eq.(2) in Eq.(1) and solving, we can get

$$y_k(n) = \frac{A}{1-AB}[x_j(n)] \qquad (3)$$

Feedback

- Let us replace operators A and B with fixed weight w and unit delay operator z^{-1} respectively as below.



- Now, $$\frac{A}{1-AB} = \frac{w}{1-wz^{-1}} = w(1-wz^{-1})^{-1} \qquad (4)$$

- Using binomial expansion

$$(1-wz^{-1})^{-1} = \sum_{l=1}^{\infty} w^l z^{-l} \qquad (5)$$

Feedback

- Now, Eq(3) becomes

$$y_k(n) = w\sum_{l=1}^{\infty} w^l z^{-l}[x_j(n)] \qquad (6)$$

- From the definition of z^{-1}, we have

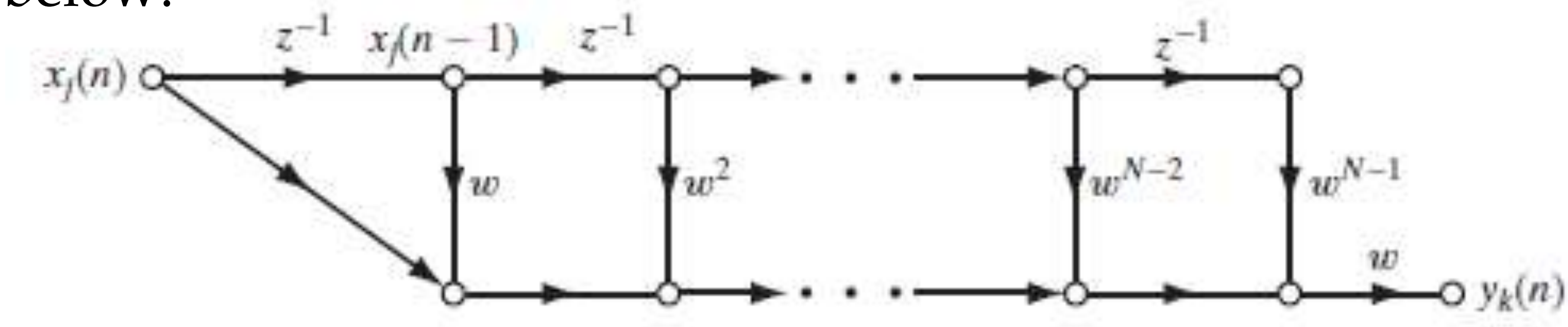
$$z^{-l}[x_j(n)] = x_j(n-l) \qquad (7)$$

- Thus,

$$y_k(n) = \sum_{l=1}^{\infty} w^{l+1} x_j(n-l) \qquad (8)$$

Feedback

- This clearly shows that every feedback system can be represented by a feed forward system. The equivalent feedforward system for the give feedback system is given below.



(b)

- There are two cases for this feed forward system
  - For $|w|<1$, the output signal $y_k(n)$ converges
  - For $|w|\geq 1$, the output signal $y_k(n)$ diverges

Knowledge Representation

- Knowledge refers to the information or model used by a machine to interpret, predict, and respond to the outside world. Two primary characteristics of knowledge representation are:
  - What knowledge is represented
  - How knowledge is encoded so that it can be used subsequently.
- Major tasks of neural networks is to learn from environment or world and keep the model consistent with the real world. Knowledge of the world consists two kinds of information.
  - The known world state represented by known facts. This is called prior information.
  - Observation of the environment obtained through sensors.

Knowledge Representation

- Representing knowledge in ANNs is very complex task. However, There are four commonsense rules used for knowledge representation in ANNs.
  1. Similar inputs from similar classes should usually produce similar representations inside the network, and should therefore be classified as belonging to the same class.
  2. Items to be categorized as separate classes should be given widely different representations in the network.
  3. If a particular feature is important, then there should be a large number of neurons involved in the representation of that item in the network.
  4. Prior information and invariances should be built into the design of a neural network whenever they are available, so as to simplify the network design by its not having to learn them.

Learning Processes

- Neural networks learns from its environment. Broadly, we cam categorize learning processes in neural networks into two categories.
  - Learning with Teacher
  - Learning without Teacher

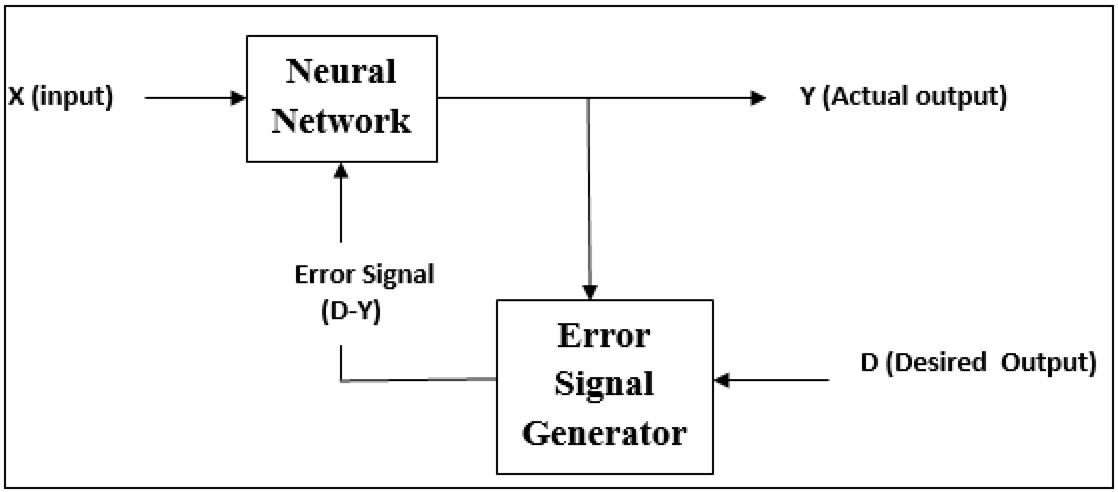
Learning Processes

Learning with Teacher

- It is also called supervised learning. In this learning paradigm, we present examples of correct input-output pairs to the neural network during the training phase.
- This training set of examples is equivalent to the teacher for the neural network. During the training of ANN under supervised learning, the ANN takes input vector and computes output vector.
- An error signal is generated, if there is a difference between the computed output and the desired output vector. On the basis of this error signal, the weights are adjusted until the actual output is matched with the desired output.
- This form of learning is called error correction learning.

Learning Processes

Learning with Teacher



- Supervised machine learning is used for performing tasks like: *Regression and Classification*.

Learning Processes

<u>Learning without Teacher</u>

- In this learning paradigm, we do not provide training set to the neural network to teach it about mapping between input and output.
- There are two types of learning processes under this learning paradigm.
  - Unsupervised Learning
  - Reinforcement Learning

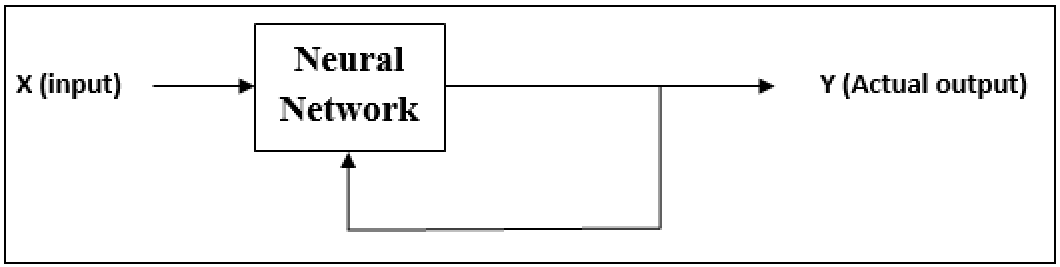
Learning Processes

Unsupervised Learning

- In unsupervised learning neural network is provided with dataset without desired output.
- The neural network then attempts to find structure in the data by extracting useful features and analyzing its structure.
- To perform this type of learning, we use competitive learning rule.

Learning Processes

Unsupervised Learning



- Unsupervised learning algorithms are widely used for tasks like: *clustering, dimensionality reduction, association mining etc.*

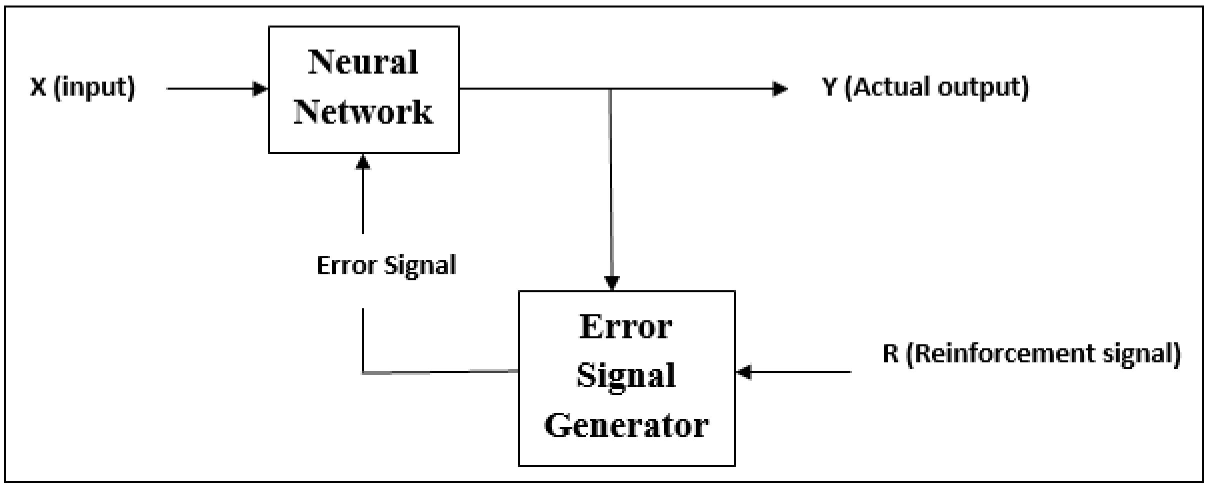
Learning Processes

Reinforcement Learning

- In reinforcement learning, we do not provide the machine with examples of correct input-output pairs, but we do provide a method for the machine to quantify its performance in the form of a reward signal.
- Reinforcement learning methods resemble how humans and animals learn: the machine tries a bunch of different things and is rewarded with performance signal.

Learning Processes

Reinforcement Learning



- Reinforcement learning algorithms are widely used for training agents interacting with its environment.

Learning Tasks

- As already mentioned, ANN can learn from environment using various learning rules. Selection of particular learning rule depends upon the learning task to be performed. Some of the major learning tasks are listed below.
  - Pattern Association
  - Pattern Recognition
  - Function Approximation
  - Control

Learning Tasks

Pattern Association

- Associative memory is defined as the ability to learn the relationship between unrelated items, such as face-name pairs. Association takes one of two forms: *auto-association* and *hetero-association*.
- In auto-association, a neural network is required to *store* a set of patterns. The network is subsequently presented with distorted (noisy) version of an original pattern stored in it, and the task is to *retrieve* that particular pattern. Auto-association involves the use of unsupervised learning.

Learning Tasks

Pattern Association

- In auto-association dimension of input and output vector is same. Character recognition is an example of auto-association.
- In hetero-association an arbitrary set of input patterns is *paired* with another arbitrary set of output patterns. The type of learning involved in hetero-association is supervised. Here, dimension of input and output vector may be different. Mapping English Sentence to Nepali Sentence is an example of hetero-association.

Learning Tasks

Pattern Recognition

- *Pattern recognition* is formally defined as *the process whereby a received pattern/signal is assigned to one of a prescribed number of classes.*
- A neural network performs pattern recognition by first undergoing a training session during which the network is repeatedly presented with a set of input patterns along with the category to which each particular pattern belongs.
- Later, the network is presented with a new pattern that has not been seen before. The network is able to identify the class of that particular pattern.

Learning Tasks

Pattern Recognition

- Pattern recognition is always supervised learning. pattern-recognition machines using neural networks may take one of two forms: The machine is split into two parts, an unsupervised network for *feature extraction* and a supervised network for *classification*. Image classification is an example of pattern recognition.

Learning Tasks

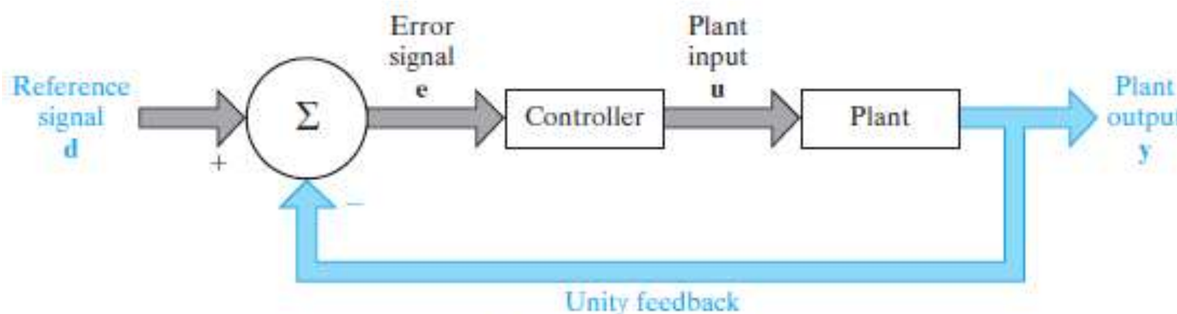
Function Approximation

- Function approximation is the process of approximating function d=f(x) from the given set of training example. The function must map input x to the output d.
- The requirement is to design a neural network that approximates the unknown function **f**(x) such that the function **F**(**x**) realized by the network, is close enough to **f**(x).
- i.e. $\left| F(x) - f(x) \right| < \in$

Learning Tasks

Control

- ANNs can be used for designing feedback control system, where output is fed back directly to the input.
- In such systems output **y** is subtracted from a *reference signal* **d** supplied from an external source. The error signal **e** so produced is applied to a neural *controller* for the purpose of adjusting its free parameters.



Block diagram of feedback control system.