KCS056 APPLICATION OF SOFT COMPUTING
UNIT-1

Prepared By- MR. MAN SINGH
ASSISTANT PROFESSOR
DEPARTMENT OF COMPUTER SCIENCE ENGINEERING
UNITED INSTITUTE OF TECHNOLOGY, PRAYAGRAJ

UNIT-1

Neural Networks-I (Introduction & Architecture) : Neuron, Nerve structure and synapse, Artificial Neuron and its model, activation functions, Neural network architecture: single layer and multilayer feed forward networks, recurrent networks. Various learning techniques; perception and convergence rule, Auto-associative and hetro-associative memory.

LECTURE-1

What is Soft Computing ?

•Zadeh, defined Soft Computing into one multidisciplinary system as the fusion of the fields of Fuzzy Logic, Neuro-Computing, Evolutionary and Genetic Computing, and Probabilistic Computing.

•The idea of soft computing was initiated in 1981 when Lotfi A. *Zadeh* published his first paper on soft data analysis "What is Soft Computing", Soft Computing. Springer-Verlag Germany/USA 1997.].

•Soft Computing is the fusion of methodologies designed to model and enable solutions to real world problems, which are not modeled or too difficult to model mathematically.

•The aim of Soft Computing is to exploit the tolerance for imprecision, uncertainty, approximate reasoning, and partial truth in order to achieve close resemblance with human like decision making.

•The Soft Computing – development history

SC	=	***EC***	+	***NN***	+	***FL***
Soft Computing		**Evolutionary Computing**		**Neural Network**		**Fuzzy Logic**
Zadeh 1981		**Rechenberg 1960**		**McCulloch 1943**		**Zadeh 1965**

EC	=	***GP***	+	***ES***	+	***EP***	+	***GA***
Evolutionary Computing		***Genetic Programming***		***Evolution Strategies***		***Evolutionary Programming***		***Genetic Algorithms***
Rechenberg 1960		**Koza 1992**		**Rechenberg 1965**		**Fogel 1962**		**Holland 1970**

Soft computing vs. hard computing:

Following points clearly differentiate the both:
•Soft Computing is tolerant of imprecision, uncertainty, partial truth and approximation whereas Hard Computing requires a precisely state analytic model.

•Soft Computing is based on fuzzy logic, neural sets, and probabilistic reasoning whereas hard Computing is based on binary logic, crisp system, numerical analysis and crisp software.

•Soft computing has the characteristics of approximation and dispositional whereas hard computing has the characteristics of precision and categorical.

•Soft computing can evolve its own programs whereas hard computing requires programs to be written.

•Soft computing can use multi valued or fuzzy logic whereas hard computing uses two-valued logic.

Introduction

To begin, first explained, the definitions, the goals, and the importance of the soft computing. Later, presented its different fields, that is, Fuzzy Computing, Neural Computing, Genetic Algorithms, and more.

Definitions of Soft Computing (SC)

Lotfi A. Zadeh, 1992 : “Soft Computing is an emerging approach to computing which parallel the remarkable ability of the human mind to reason and learn in a environment of uncertainty and imprecision”.

The Soft Computing consists of several computing paradigms mainly :

Fuzzy Systems, Neural Networks, and Genetic Algorithms.

- Fuzzy set : for knowledge representation via fuzzy If – Then rules.
- Neural Networks : for learning and adaptation
- Genetic Algorithms : for evolutionary computation These methodologies form the core of SC.

Hybridization of these three creates a successful synergic effect;

that is, hybridization creates a situation where different entities cooperate advantageously for a final outcome.

Soft Computing is still growing and developing.

Hence, a clear definite agreement on what comprises Soft Computing has not yet been reached. More new sciences are still merging into Soft Computing.

Goals of Soft Computing

Soft Computing is a new multidisciplinary field, to construct new generation of Artificial Intelligence, known as **Computational Intelligence**.

- The main goal of Soft Computing is to develop intelligent machines to provide solutions to real world problems, which are not modeled, or too difficult to model mathematically.

- Its aim is to exploit the tolerance for **Approximation, Uncertainty**, **Imprecision**, and **Partial Truth** in order to achieve close resemblance with human like decision making.

Approximation : here the model features are similar to the real ones, but not the same.

Uncertainty : here we are not sure that the features of the model are the same as that of the entity (belief).

Imprecision : here the model features (quantities) are not the same as that of the real ones, but close to them.

LECTURE-2

Importance of Soft Computing

soft computing differs from hard (conventional) computing. Unlike hard computing, the soft computing is **tolerant of imprecision, uncertainty, partial truth, and approximation**. The guiding principle of soft computing is to exploit these tolerance to achieve tractability, robustness and low

solution cost. In effect, the role model for soft computing is the human mind.

The four fields that constitute Soft Computing (SC) are : **Fuzzy Computing** (FC), **Evolutionary Computing**(EC), **Neural computing** (NC), and **Probabilistic**

Computing (PC), with the latter subsuming belief networks, chaos theory and parts of learning theory.

Soft computing is not a concoction, mixture, or combination, rather, **Soft computing is a partnership** in which each of the partners contributes a distinct methodology for addressing problems in its domain. In principal the constituent methodologies in Soft computing are complementary rather than competitive.

Soft computing may be viewed as a foundation component for the emerging field of Conceptual Intelligence.

Neural Computing

Neural Computers mimic certain processing capabilities of the human brain.

- Neural Computing is an information processing paradigm, inspired by biological system, composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems.

- A neural net is an artificial representation of the human brain that tries to simulate its learning process. The term "artificial" means that neural nets are implemented in computer programs that are able to handle the large number of necessary calculations during the learning process.

- Artificial Neural Networks (ANNs), like people, learn by example.

- An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process.

- Learning in biological systems involves adjustments to the synaptic connections that exist between the neurons. This is true for ANNs as well.

Biological Model:

The human brain consists of a large number (more than a billion) of neural cells that process information. Each cell works like a simple processor. The massive interaction between all cells and their parallel processing, makes the brain's abilities possible. The structure of neuron is shown below.

Dendrites are the *branching* fibers extending from the cell body or soma.

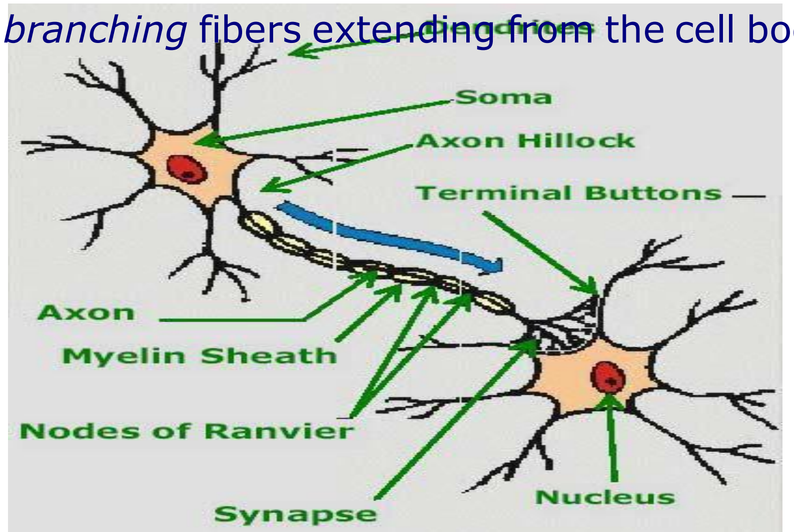


Fig. Structure of Neuron

Soma or cell body of a neuron contains the nucleus and other structures, support chemical processing and production of neurotransmitters.

Axon is a singular fiber carries information away from the soma to the synaptic sites of other neurons (dendrites and somas), muscles, or glands.

Axon hillock is the site of summation for incoming information.

Myelin Sheath consists of fat-containing cells that insulate the axon from electrical activity.

Nodes of Ranvier are the gaps between myelin sheath cells long axons.

Synapse is the point of connection between two neurons or a neuron and a muscle or a gland. Electrochemical communication between neurons takes place at these junctions.

Terminal Buttons of a neuron are the small knobs at the end of an axon.

LECTURE-3

•Information flow in a Neural Cell

The input /output and the propagation of information are shown below.

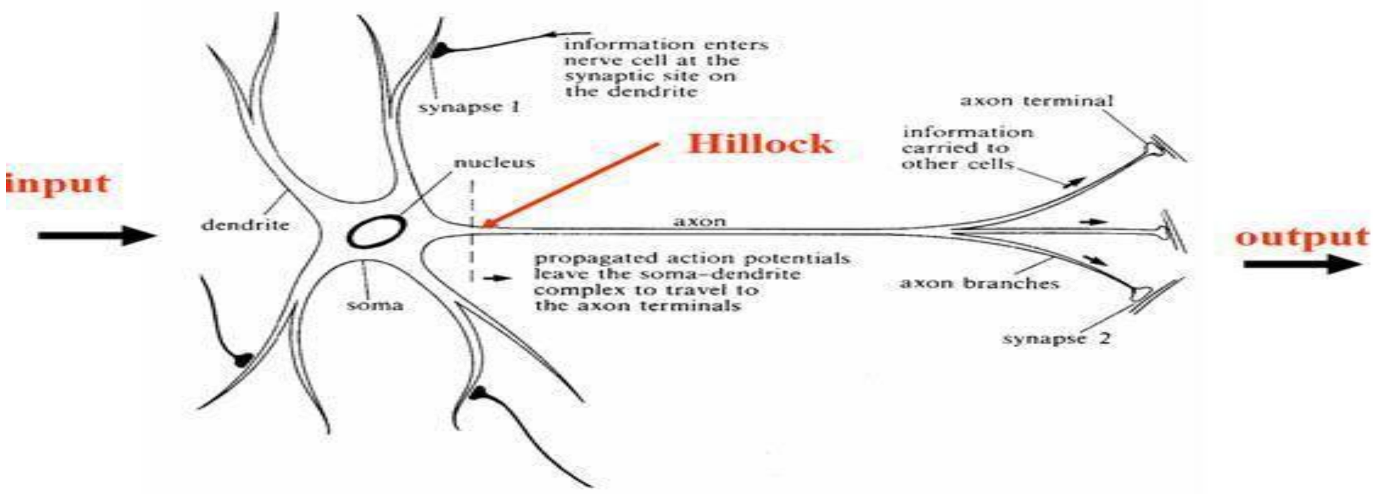


Fig. Structure of a neural cell in the human brain

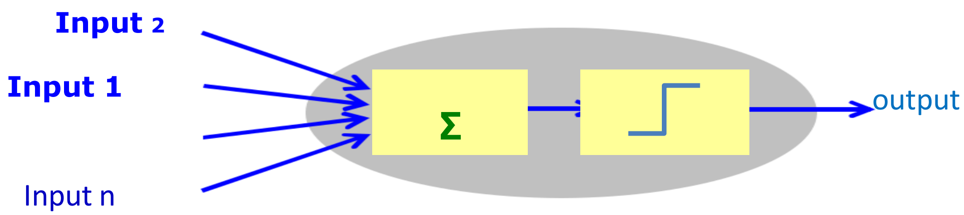
- Dendrites receive activation from other neurons.
- Soma processes the incoming activations and converts them into output activations.
- Axons act as transmission lines to send activation to other neurons.
- Synapses the junctions allow signal transmission between the axons and dendrites.
- The process of transmission is by diffusion of chemicals called neuro- transmitters.

McCulloch-Pitts introduced a simplified model of this real neurons.

Artificial Neuron

- **The McCulloch-Pitts Neuron**

This is a simplified model of real neurons, known as a Threshold Logic Unit.



▪A set of synapses (ie connections) brings in activations from other neurons.

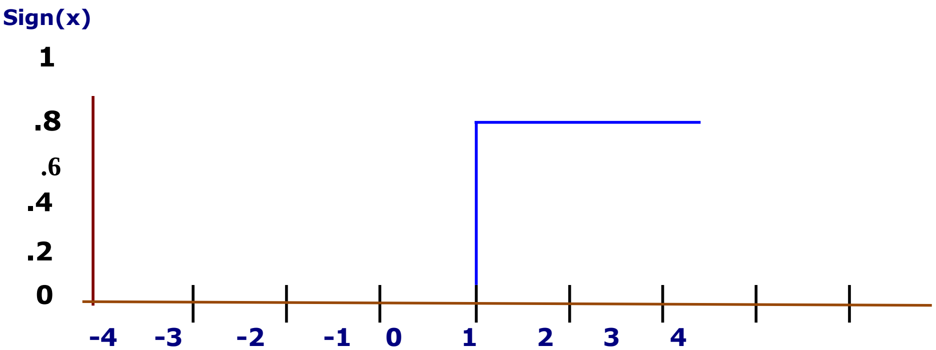
▪A processing unit sums the inputs, and then applies a non-linear activation function (i.e. transfer / threshold function).

▪An output line transmits the result to other neurons.

• **Functions :** The function **y = f(x)** describes a relationship, an input-output mapping, from **x** to **y**.

■ **Threshold or Sign function sgn(x)** : defined as

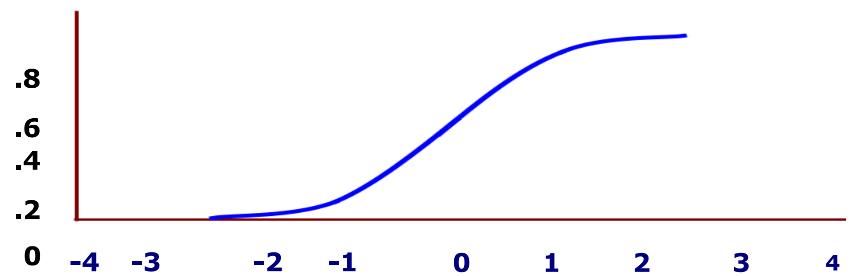
$$\text{sgn}(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$



■ **Threshold or Sign function sigmoid (x)** : defined as a smoothed (differentiable) form of the threshold function

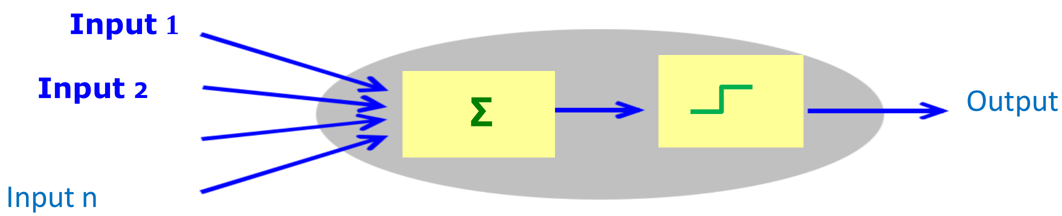
Sign(x)

$$\text{sigmoid}(x) = \frac{1}{1 + e^{-x}}$$



•McCulloch-Pitts (M-P) Neuron Equation

Fig below is the same previously shown simplified model of a real neuron, as a threshold Logic Unit.



The equation for the output of a McCulloch-Pitts neuron as a function of **1** to **n** inputs is written as :

$$\textbf{Output} = \operatorname{sgn}\left(\sum_{i=1}^{n} \text{Input}_i - \Phi \right)$$

where **Φ** is the neuron's activation threshold.

$$\text{If} \sum_{i=1}^{n} \text{Input } i \;\varepsilon\; \Phi \quad \text{then Output} = 1$$

$$\text{If} \sum_{i=1}^{n} \text{Input } i < \Phi \quad \text{then Output} = 0$$

Note : The McCulloch-Pitts neuron is an extremely simplified model of real biological neurons. Some of its missing features include: non-binary input and output, non-linear summation, smooth thresholding, stochastic (non- deterministic), and temporal information processing.

LECTURE-4

•Basic Elements of an Artificial Neuron

It consists of three basic components - weights, thresholds, and a single activation function.

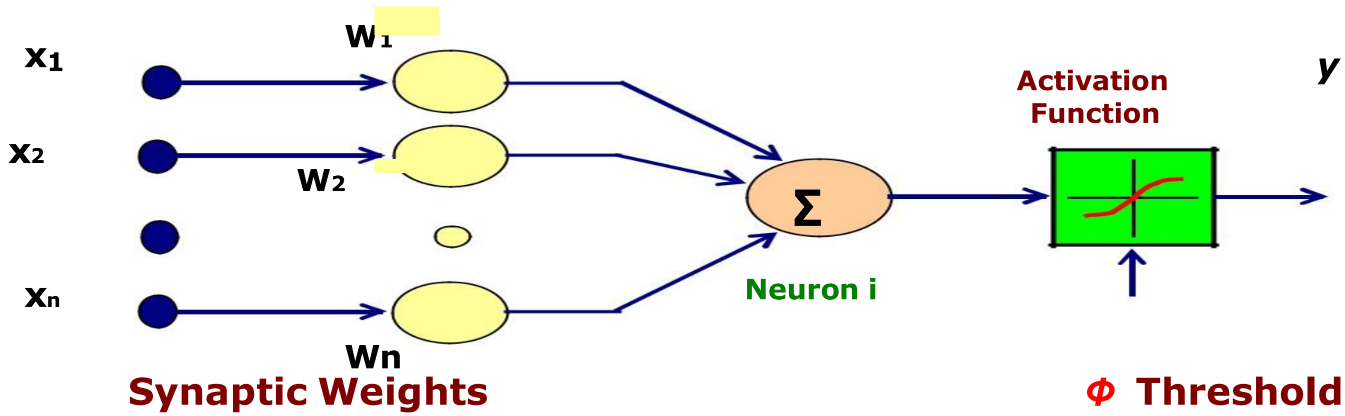


Fig Basic Elements of an Artificial Neuron

Weighting Factors:

The values **W1 , W2 , . . . Wn** are weights to determine the strength of input row vector **X = [x1 , x2 , . . . , xn]^T** . Each input is multiplied by the associated weight of the neuron connection **X^T W**. The **+ve** weight excites and the **-ve** weight inhibits the node output.

Threshold:

The node's internal threshold **Φ** is the magnitude offset. It affects the activation of the node output **y** as:

$$y = \sum_{i=1}^{n} X_i \; W_i - \Phi_k$$

Activation Function: An activation function performs a mathematical operation on the signal output. The most common activation functions are, Linear Function, Threshold Function, Piecewise Linear Function, Sigmoidal (S shaped) function, Tangent hyperbolic function and are chose depending upon the type of problem to be solved by the network.

• **Example :** A neural network consists four inputs with the weights as shown.

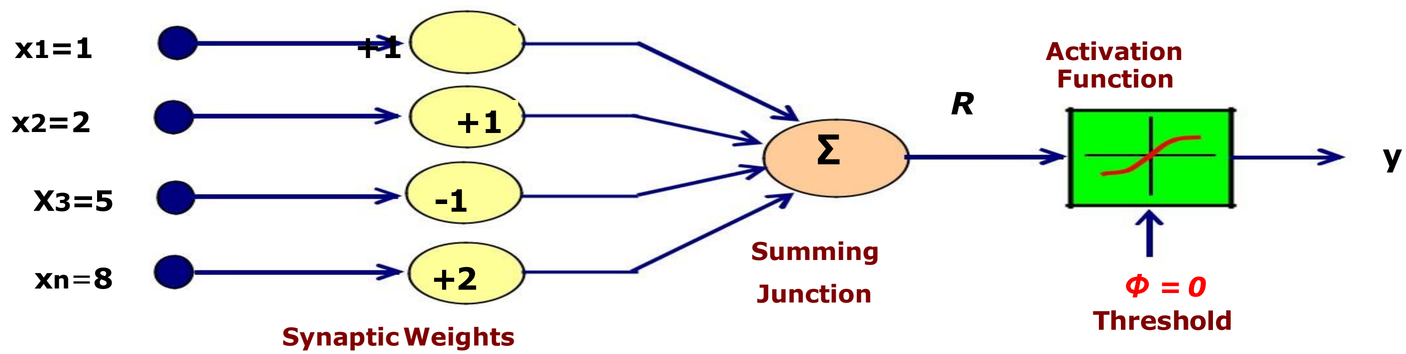


Fig Neuron Structure of Example

The output **R** of the network, prior to the activation function stage, is

$$R = W^T \cdot X = \begin{bmatrix} 1 & 1 & -1 & 2 \end{bmatrix} \bullet \begin{bmatrix} 1 \\ 2 \\ 5 \\ 8 \end{bmatrix} = 14$$

$$= (1 \times 1) + (1 \times 2) + (-1 \times 5) + (2 \times 8) = 14$$

With a binary activation function, the outputs of the neuron is:

y (threshold) = 1

•Networks of McCulloch-Pitts Neurons

One neuron can not do much on its own. Usually we will have many neurons labeled by indices **k, i, j** and activation flows between them via synapses with strengths **wki, wij** :

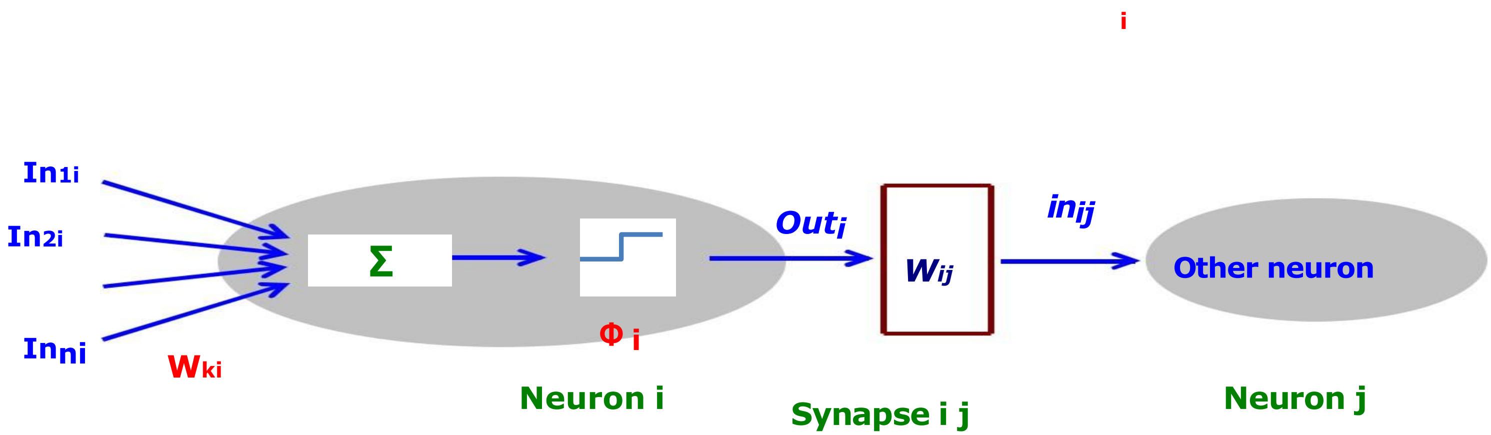


Fig McCulloch-Pitts Neuron

$in_{ki} \; W_{ki} = Out_k$, $\quad Out_i = sgn\left(\sum_{k=1}^{n} In_{ki} - \Phi\right)$, $\quad In_{ij} = Out_i \, W_{ij}$

LECTURE-5

•Single and Multi - Layer Perceptrons

A perceptron is a name for simulated neuron in the computer program. The usually way to represent a neuron model is described below.

The neurons are shown as circles in the diagram. It has several inputs and a single output. The neurons have gone under various names.

- Each individual cell is called either a **node** or a **perceptron**.
- A neural network consisting of a layer of nodes or perceptrons between the input and the output is called a **single layer perceptron**.
- A network consisting of several layers of single layer perceptron stacked on top of other, between input and output , is called a **multi-layer perceptron**

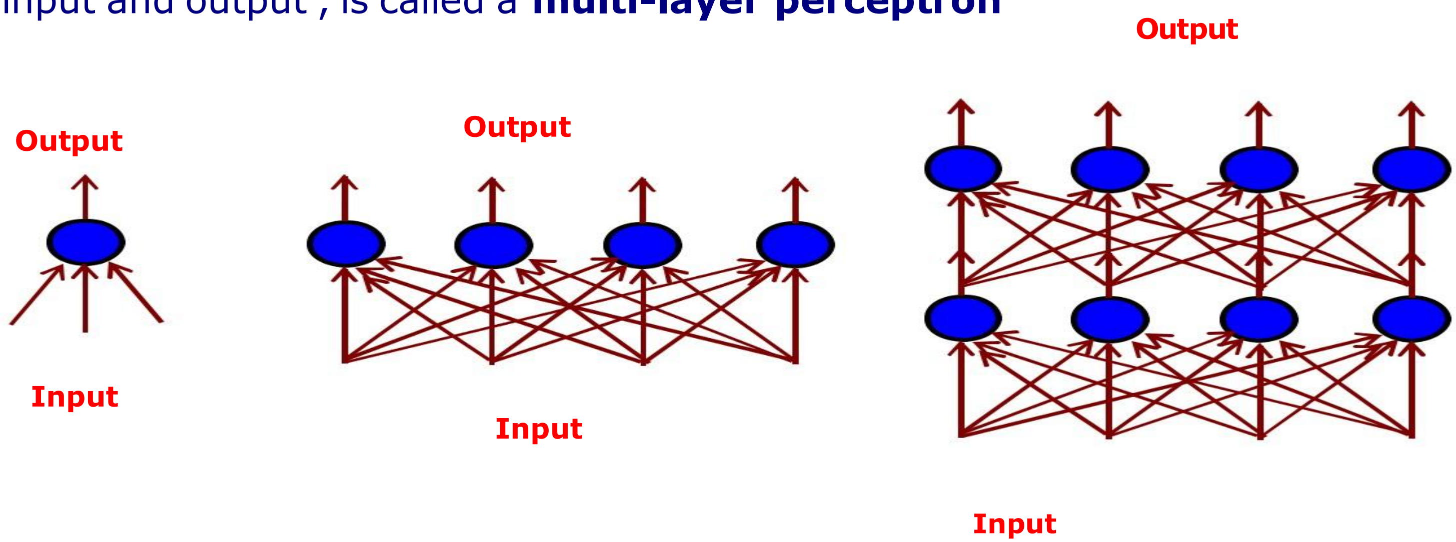


Fig Single and Multi - Layer Perceptrons

Multi-layer perceptrons are more powerful than single-layer perceptrons.

•Perceptron

Any number of McCulloch-Pitts neurons can be connected together in any way.

Definition : An arrangement of one input layer of McCulloch-Pitts neurons, that is feeding forward to one output layer of McCulloch-Pitts neurons is known as a Perceptron.

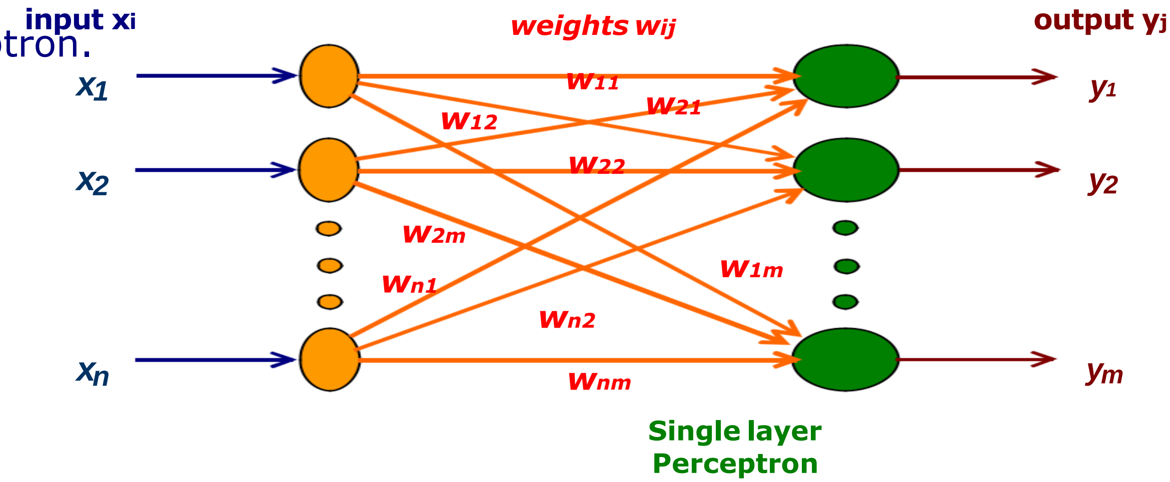


Fig. Simple Perceptron Model

$$y_j = f(net_j) = \begin{cases} 1 \text{ if } net_j \; \varepsilon \; 0 \\ 0 \text{ if } net_j < 0 \end{cases} \quad \text{where } net_j = \sum_{i=1}^{n} x_i \, w_{ij}$$

A Perceptron is a powerful computational device.

•Mechanics of Biological Evolution

Genetic Algorithms are a way of solving problems by mimicking processes the nature uses - Selection, Crosses over, Mutation and Accepting to evolve a solution to a problem.

- Every **organism** has a set of **rules**, describing how that organism is built, and encoded in the **genes** of an organism.

- The genes are connected together into long strings called **chromosomes**.

- Each gene represents a specific **trait** (feature) of the organism and has several different settings, e.g. setting for a hair color gene may be black or brown.

- The genes and their settings are referred as an organism's **genotype**.

- When two organisms mate they share their genes. The resultant offspring may end up having half the genes from one parent and half the genes from the other parent. This process is called **crossover** (recombination).

- The newly created offspring can then be mutated. A gene may be **mutated** and expressed in the organism as a completely new trait. Mutation means, that the elements of DNA are a bit changed. This change is mainly caused by errors in copying genes from parents.

- The **fitness** of an organism is measured by success of the organism in its life.

Artificial Evolution and Search Optimization

The problem of finding solutions to problems is itself a problem with no general solution. Solving problems usually mean looking for solutions, which will be the best among others.

- In engineering and mathematics finding the solution to a problem is often thought as a process of optimization.

- Here the process is : first formulate the problems as mathematical models expressed in terms of functions; then to find a solution, discover the parameters that optimize the model or the function components that provide optimal system performance.

The well-established search / optimization techniques are usually classified in to three broad categories : Enumerative, Calculus-based, and Guided random search techniques. A taxonomy of Evolution & Search Optimization classes is illustrated in the next slide.

LECTURE-6

- **Taxonomy of Evolution & Search Optimization Classes**

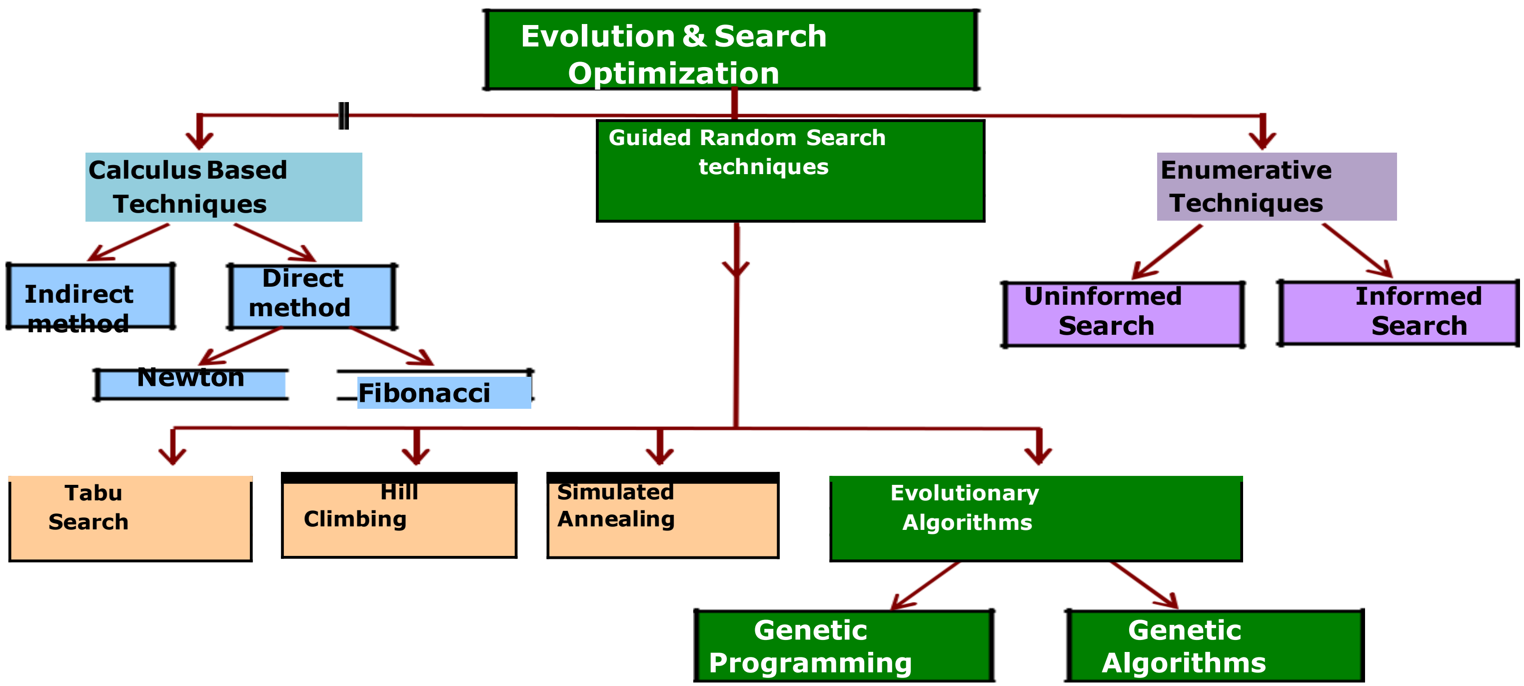


Fig Evolution & Search Optimization Techniques

Associative Memory

An associative memory is a content-addressable structure that maps a set of input patterns to a set of output patterns. The associative memory are of two types : auto-associative and hetero-associative.

- An **auto-associative memory** retrieves a previously stored pattern that most closely resembles the current pattern.

- In a **hetero-associative memory**, the retrieved pattern is, in general, different from the input pattern not only in content but possibly also in type and format.

•Example : Associative Memory

The figure below shows a memory containing names of several people.

If the given memory is content-addressable,

Then using the erroneous string "Crhistpher Columbos" as key is sufficient to retrieve the correct name "Christopher Colombus."

In this sense, this type of memory is robust and fault-tolerant, because this type of memory exhibits some form of error-correction capability.

Fig. A content-addressable memory, Input and Output

Description of Associative Memory

An associative memory is a content-addressable structure that maps specific input representations to specific output representations.

- A content-addressable memory is a type of memory that allows, the recall of data based on the degree of similarity between the input pattern and the patterns stored in memory.
- It refers to a memory organization in which the memory is accessed by its content and not or opposed to an explicit address in the traditional computer memory system.
- This type of memory allows the recall of information based on partial knowledge of its contents.

LECTURE-7

▪It is a system that "associates" two patterns **(X, Y)** such that when one is encountered, the other can be recalled.

- Let **X** and **Y** be two vectors of length **m** and **n** respectively.
- Typically, **XÎ $\{-1, +1\}^m$, Y Î $\{-1, +1\}^n$**
- The components of the vectors can be thought of as pixels when the two patterns are considered as bitmap images.

▪There are two classes of associative memory:

- auto-associative and hetero-associative.

An **auto-associative** memory is used to retrieve a previously stored pattern that most closely resembles the current pattern.

In a **hetero-associative** memory, the retrieved pattern is, in general, different from the input pattern not only in content but possibly also different in type and format.

▪Artificial neural networks can be used as associative memories.

The simplest artificial neural associative memory is the linear associater**.** The other popular ANN models used as associative memories are Hopfield model and Bidirectional Associative Memory (BAM) models

Adaptive Resonance Theory (ART)

ART stands for "*Adaptive Resonance Theory*", invented by Stephen Grossberg in 1976. ART encompasses a wide variety of neural networks, based explicitly on neurophysiology. The word "*Resonance*" is a concept, just a matter of being within a certain threshold of a second similarity measure.

The basic ART system is an **unsupervised learning model**, similar to many iterative clustering algorithm where each case is processed by finding the nearest" cluster seed that resonate with the case and update the cluster seed to be "closer" to the case. If no seed resonate with the case then a new cluster is created.

Note : The terms *nearest* and *closer* are defined in many ways in clustering algorithm. In ART, these two terms are defined in slightly different way by introducing the concept of "resonance".

Definitions of ART and other types of Learning

ART is a neural network topology whose dynamics are based on Adaptive Resonance Theory (ART). Grossberg developed ART as a theory of human cognitive information processing. The emphasis of ART neural networks lies at unsupervised learning and self-organization to mimic biological behavior. Self-organization means that the system must be able to build stable recognition categories in real-time.

The unsupervised learning means that the network learns the significant patterns on the basis of the inputs only. There is no feedback. There is no external teacher that instructs the network or tells to which category a certain input belongs. Learning in biological systems always starts as unsupervised learning; Example : For the newly born, hardly any pre- existing categories exist.

The other two types of learning are reinforcement learning and supervised learning. In reinforcement learning the net only limited feedback, like *"on this input you performed well or "on this input you have made an error"*. In supervised mode of learning a net receives for each input the correct response.

Note: A system that can learn in unsupervised mode can always be adjusted to learn in the other modes, like reinforcement mode pr supervised mode. But a system specifically designed to learn in supervised mode can never perform in unsupervised mode

•Description of Adaptive Resonance Theory

The basic ART system is an **unsupervised learning model**. The model typically consists of :

- a comparison field and a recognition field composed of neurons,
- a vigilance parameter, and
- a reset module.

The functions of each of these constituents are explained below.

▪Comparison field and Recognition field

- The Comparison field takes an input vector (a 1-D array of values) and transfers it to its best match in the Recognition field; the best match is, the single neuron whose set of weights (weight vector) matches most closely the input vector.
- Each Recognition Field neuron outputs a negative signal (proportional to that neuron's quality of match to the input vector) to each of the other Recognition field neurons and inhibits their output accordingly.
- Recognition field thus exhibits lateral inhibition, allowing each neuron in it to represent a category to which input vectors are classified.

▪Vigilance parameter: It has considerable influence on the system memories:

- higher vigilance produces highly detailed memories,
- lower vigilance results in more general memories

▪Reset module: After the input vector is classified, the Reset module compares the strength of the recognition match with the vigilance parameter.

- If the vigilance threshold is met, Then training commences.
- Else, the firing recognition neuron is inhibited until a new input vector is applied;

LECTURE-8

Training ART-based Neural Networks

Training commences only upon completion of a search procedure. What happens in this search procedure :

- The Recognition neurons are disabled one by one by the reset function until the vigilance parameter is satisfied by a recognition match.
- If no committed recognition neuron's match meets the vigilance threshold, then an uncommitted neuron is committed and adjusted towards matching the input vector.

Methods of training ART-based Neural Networks:

There are two basic methods, the slow and fast learning.

- Slow learning method : here the degree of training of the recognition neuron's weights towards the input vector is calculated using differential equations and is thus dependent on the length of time the input vector is presented.
- Fast learning method : here the algebraic equations are used to calculate degree of weight adjustments to be made, and binary values are used.

Note : While fast learning is effective and efficient for a variety of tasks, the slow learning method is more biologically plausible and can be used with continuous-time networks (i.e. when the input vector can vary continuously).

•Types of ART Systems :

The ART Systems have many variations : ART 1, ART 2, Fuzzy ART, ARTMAP

- **ART 1:** The simplest variety of ART networks, accept only binary inputs.
- **ART 2 :** It extends network capabilities to support continuous inputs.
- **Fuzzy ART :** It Implements fuzzy logic into ART's pattern recognition, thus enhances generalizing ability. One very useful feature of fuzzy ART is complement coding, a means of incorporating the absence of features into pattern classifications, which goes a long way towards preventing inefficient and unnecessary category proliferation.
- **ARTMAP :** Also known as Predictive ART, combines two slightly modified ARTs , may be two ART-1 or two ART-2 units into a supervised learning structure where the first unit takes the input data and the second unit takes the correct output data, then used to make the minimum possible adjustment of the vigilance parameter in the first unit in order to make the correct classification.

Applications of Soft Computing

The applications of Soft Computing have proved two main advantages.

- First, in solving nonlinear problems, where mathematical models are not available, or not possible.
- Second, introducing the human knowledge such as cognition, recognition, understanding, learning, and others into the fields of computing.

This resulted in the possibility of constructing intelligent systems such as autonomous self-tuning systems, and automated designed systems.

The relevance of soft computing for pattern recognition and image processing is already established during the last few years. The subject has recently gained importance because of its potential applications in problems like :

- Remotely Sensed Data Analysis,
- Data Mining, Web Mining,
- Global Positioning Systems,
- Medical Imaging,
- Forensic Applications,
- Optical Character Recognition,
- Signature Verification,
- Multimedia,
- Target Recognition,
- Face Recognition and
- Man Machine Communication.

END