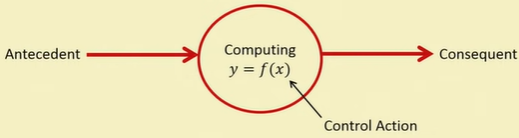
**SOFT COMPTING**

**UNIT-1: Introduction to Soft computing -** Soft computing vs. hard computing, evolution of soft computing, features and, applications of soft computing, Basic concepts of Neural Networks

**Concept of Computing**

Computing is the process of using computer technology to complete a given **goal-oriented task**. It includes the study and experimentation of algorithmic processes, and development of both hardware and software. Computing has scientific, engineering, mathematical, technological and social aspects. The role of computing is to store, access, manipulate, calculate, and analyse data and information. It has three stages of computing ***input, processing and output***.

Computation is a process of converting the input of one form to some other desired output form using certain control actions. According to the concept of computation, the input is called an antecedent and the output is called the consequent. A mapping function converts the input of one form to another form of desired output using certain control actions. The computing concept is mainly applicable to computer science engineering.



*y = f (x),* f is ***a mapping function*** also called a formal method or ***an algorithm*** to solve a problem.

Important Characteristics of Computing are

1. Should provide ***precise solution***.
2. Control action should be ***unambiguous and accurate***.
3. Suitable for problem, this is easy to ***model mathematically***.

There are two types of computing methods namely: ***Hard computing and Soft Computing***

**Hard Computing**: Hard Computing is a conventional approach used in computing and requires an accurately stated analytical model. In 1996, LA Zadeh (LAZ) introduced the term hard computing. Hard computing is a process in which we program the computer to solve certain problems using mathematical algorithms that already exist, which provides a precise output value. According to LAZ: We term a computing as “Hard” computing, if

* **Precise result** is guaranteed
* Control action is **unambiguous**
* Control action is **formally defined** (i.e. with mathematical model or algorithm)

Some of the hard computing examples are:

* Solving **numerical problems** (e.g. Roots of polynomials, Integration etc.)
* **Searching and sorting techniques**
* Solving “**Computational Geometry**” problems (e.g. shortest tour in Graph theory, finding closest pair of points etc.)

The major drawback of hard computing is that it is incapable in solving the real world problems whose behaviour is imprecise and their information being changing continuously. Hard computing is mainly used to perform **sequential computations**.

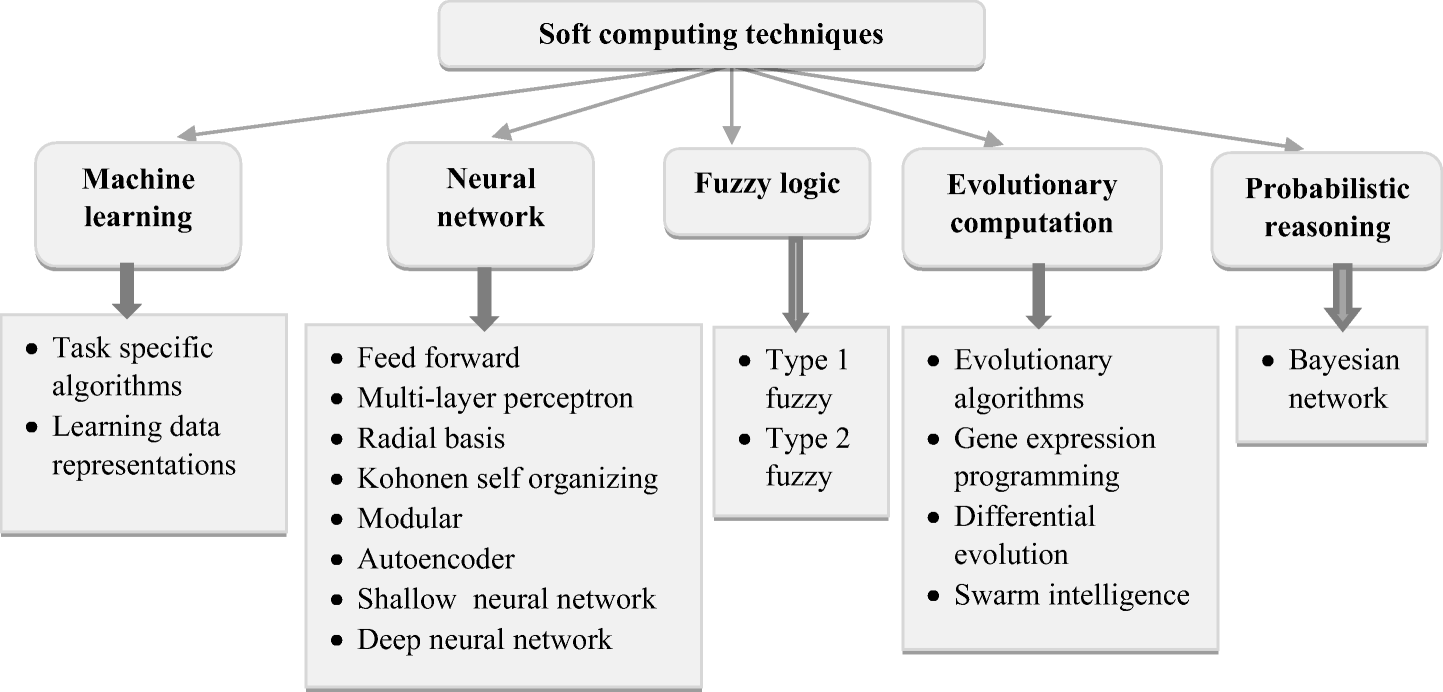
**Soft-Computing**

Branch of Artificial Intelligence that deals with systems and methodologies that can perform approximate, qualitative, human-like reasoning

* Humans can make intelligent decisions using incomplete and imprecise information **“Soft” reasoning**
* Computer algorithms require complete and precise information **“Hard” reasoning**
* Soft computing aims to **bridge this gap**

The idea of soft computing was given by Lotfi A. Zadeh in 1981. Soft Computing is a new multidisciplinary field, to construct new generation of Artificial Intelligence, known as **Computational Intelligence**. Therefore, soft computing is also known as computational intelligence. “It is the fusion of the fields of Fuzzy Logic, Neuro Computing, Evolutionary, Genetic computing and Probabilistic Computing.

* Fuzzy Logic (FL) - Human knowledge
* Artificial Neural Networks (ANN) -Human brain
* Probabilistic Reasoning - Human uncertainty and randomness
* Evolutionary Computing- Genetic algorithms (GA), Biological evolution



***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.***

Soft computing is an approach where we compute solutions to the existing complex problems, where output results are imprecise or fuzzy in nature, one of the most important features of soft computing is it should be adaptive so that any change in environment does not affect the present process.

**According to Lotfi A. Zadeh:**

***The aim of soft computing is to exploit the tolerance for imprecision, uncertainty, approximate reasoning and partial truth in order to achieve tractability, robustness, low-cost solutions, and close resemblance with human-like decision-making.” The role model of soft computing is human mind.***

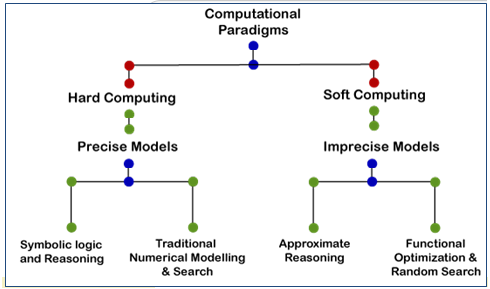
**Imprecision** - model features (quantities) are not the same as that of the real ones, but close to them.

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

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

Characteristics of Soft Computing are:

1. It does not require any mathematical modeling of problem solving.
2. It may not yield the precise solution.
3. Algorithms are adaptive (i.e., it can adjust to the change of dynamic environment).
4. Use some biological inspired methodologies such as genetics, evolution, ant’s behaviors, particles swarming, human nervous system, etc.).



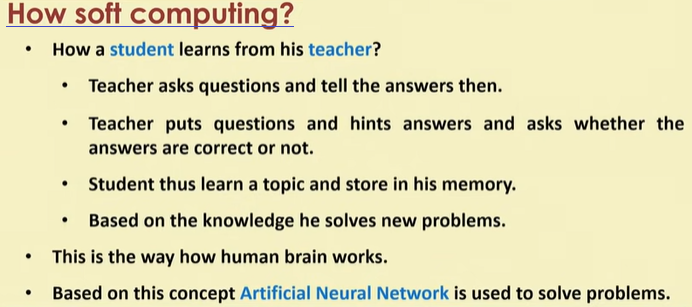
|  |  |  |
| --- | --- | --- |
|  | **Hard Computing** | **Soft Computing** |
| **Input** | Requires exact data | It can deal with ambiguous and noisy data i.e., Approximate and incomplete data |
| **Constraints** | Real time constraints | Artificial constraints |
| **Reasoning** | Rational | Human-like |
| **Model** | It requires a precisely stated analytical model. | It is tolerant of imprecision, uncertainty, partial truth and approximation |
| **Computation time** | Takes more computation time. | Takes less computation time. |
| **Dependency** | It is based on binary logic, crisp systems, numerical analysis and crisp software. | It is based on fuzzy logic, neural nets and probabilistic reasoning. |
| **Characteristic** | Precision and categoricity | approximation and dis-positionality |
| **Computation type** | Sequential computation | Parallel computation |
| **Result/Output** | Exact and precise result | It can yield approximate results |
|  | It is deterministic | It incorporates stochasticity |
| **Example** | Any numerical problem or traditional methods of solving using personal computers. | Neural Networks, such as Madaline, Adaline, Art Networks. |

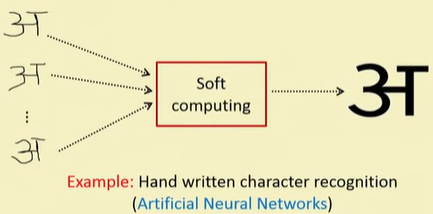
It extends its application to various disciplines of Engineering. & Science. Typically human can:

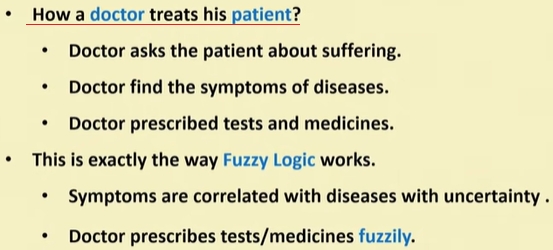
1. Take decisions
2. Inference from previous situations experienced
3. Expertise in an area
4. Adapt to changing environment
5. Learn to do better
6. Social behavior of collective intelligence

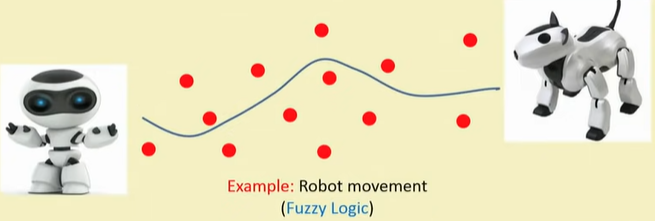
Intelligent control strategies have emerged from the above mentioned characteristics of human/ animals.

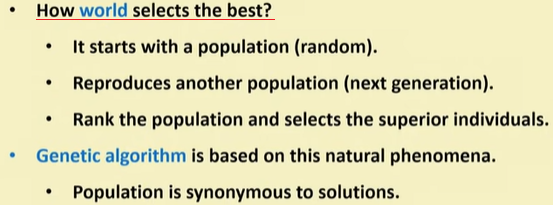
* The first two characteristics have given rise to Fuzzy logic;
* 2nd , 3rd and 4th have led to Neural Networks;
* 4th, 5th and 6th have been used in evolutionary algorithms.

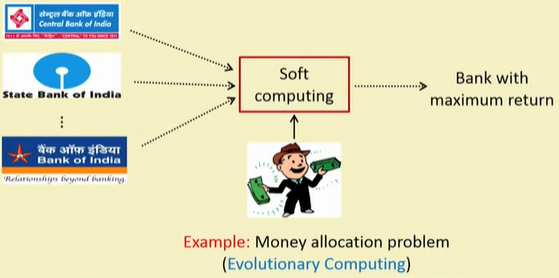
******





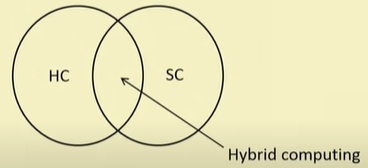


******



1. Neural Networks ----- operate in a numeric framework, and are well known for their learning and generalization capabilities.
2. Fuzzy Systems ----- operate in a linguistic framework, and their strength lies in their capability to handle linguistic information and perform approximate reasoning. Knowledge representation via Fuzzy if-then rules
3. Evolutionary Algorithms ----- provide powerful search and optimization methodologies.

All the three facets of soft computing differ from one another in their time scales of operation and in the extent to which they embed a priori knowledge. These methodologies form the core of Soft Computing. 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.



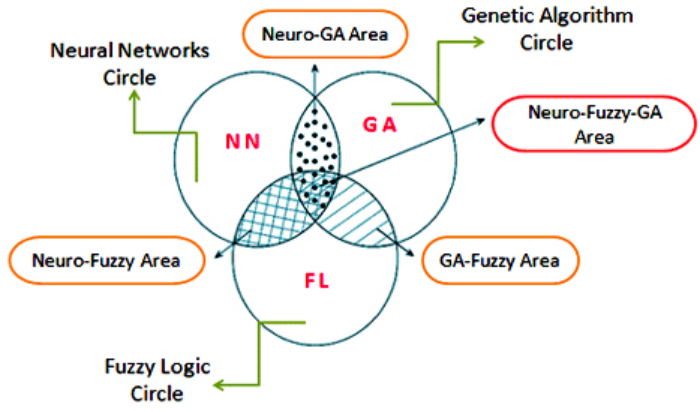
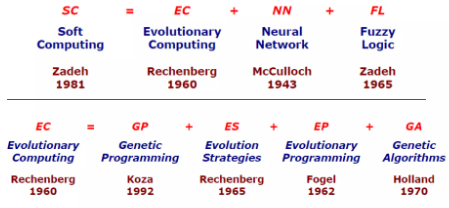


Figure: Integration of neural networks area (NN), fuzzy logic system area (FL) and genetic algorithm area (GA) technologies.

**Evolution of Soft Computing**

******

**Artificial Neural Network**

**I think the brain is essentially a computer, and consciousness is like a computer program. It will cease to run when computer is turned off. Theoretically, it could be recreated on a Neural Network, but that would require all one’s memories.**

**-Stephen Hawking**

**Theoretical, physicist, cosmologist and author**

**\*\*\***

Dr. Robert Hecht-Nielsen, the inventor of the first neuro-computers, defines neural network as:

***“A computing system made up of a number of simple, highly interconnected processing elements, which process information by their dynamic state response to external inputs”.***

Thus one thing which is worth noticed here is that the most prominent component of neural computers is neural networks; both artificial and biological neural networks are prominently required to be discussed for building the basic concept of neural computing as well as soft computing.

After many years of research on human brains, researchers discovered the fact that **brains stores information in the form of patterns.** Some of these patterns are highly complicated and give us the capability of recognizing faces from various angles. This technique of storing information in the form of patterns, exploiting those patterns in solving problems, evidenced to be the foundation of Neural Computing.

* Does not utilize traditional programming mechanisms
* It involves the creation of massively parallel network and the training of that network to solve specific problems especially in situations where it is not possible to follow the so called protocols or rules that lead to the solution of a particular problem.
* It is a promising field of computing, which can proved especially in the applications of pattern recognition.

Every day research on biological neurons promise a basic understanding of the natural thinking mechanism. Neural Networks (NN) adopt various learning mechanisms of which supervised learning and unsupervised learning methods are very much popular and widely used in practical applications.

* In supervised learning, a learning signal (teacher) is assumed to be present during the learning procedure, i.e., the network tries to minimize the error between the target (desired) output presented by the ‘teacher’ and the actual output, to improve the performance.
* Whereas, in un-supervised learning, there is no teacher available to present the desired output; as a result, the network tries to learn by itself, by organizing and reorganizing the input instances of the problem.

Though NN architectures have been broadly classified as

* Single layer feed-forward networks, Multilayer feed-forward networks, Recurrent networks
* Other NN architectures have also been evolved. Some of them include Back-propagation network, Perceptron, ADALINE (Adaptive Linear Element), associative memory, Boltzmann machine, self-organizing feature map, adaptive resonance theory and Hopfield network.

Briefly, a neural network can be considered as a mystery device that is capable of predicting an output pattern after recognizing a presented input pattern. Once effectively trained (programmed), it is also capable of recognizing similarities when a new input pattern is presented, again resulting a predicted output pattern. Thus could be successfully applied to the field of pattern recognition, image processing, data compression, forecasting, optimization and many others

**Fuzzy Logic**

The major difference between conventional computing (hard computing) and human brains is that, human brains have the capability to interpret the imprecise and unambiguous set of information provided as input by different organs like eyes, ears, tongue or other sensory organs. The aim of soft computing is to inhibit this property of human brain, and fuzzy logic plays a vital role in this task, by providing logical algorithms to handle such semantic information.

The origin of Fuzzy Logic is emerged from the theory of Fuzzy sets introduced by Sir Lotfi A. Zadeh in the mid ‘60s.

* It is also termed as ‘many-valued logic’ and deals with approximate rather than exact reasoning.
* Fuzzy logic variables acquire a truth value that ranges in probabilistic degree between 0 and1.
* For example, in case of temperature sensing, a system exhibiting Fuzzy logic will answer slightly warm or fairly cold, rather than just warm and cold.
* Thus, it provides a foundation for the development of new tools for dealing with natural languages and natural decision making capabilities.
* The Fuzzy Logic is comprised of four fundamental aspects-
* Logical - deals with logical systems in which solutions or results are not based on exact truth, rather to a certain extent or matter of degree of truth.
* Set- theoretic - encapsulated in the Fuzzy set theory provides a systematic algorithm to handle set of imprecise information linguistically and to perform mathematical operations on this information by imposing the membership functions.
* Relational - comprised in the fuzzy relations, is focused on fuzzy dependencies, correlations and fuzzy rule sets. Most of the practical implementation of fuzzy logic to real world applications is related to Fuzzy reasoning. For instance, fuzzy if-then rules are served as a basis of many automated systems like temperature sensor, automatic toll booth gate controller, fire alarms, earthquake predictors etc.
* Epistemic - is mainly concerned with knowledge representation, natural languages, linguistics and expert systems. Probabilistic reasoning is also a part of it.

Further, the right usage of fuzzy if then rules are very important for a fuzzy inference system (FIS) that can effectively model human like common sense in machinery. Fuzzy inference system can easily interpret the incoming imprecise information by using fuzzy if-then rules, but the limitation of fuzzy inference system is its adequacy in dealing with the changing external conditions. Therefore a neuro-fuzzy expert system was developed by utilizing the learning algorithms of neural network.

In soft computing, the term, Fuzzy logic has been used in two different notions:

* Broadly, it is considered as a system comprised of concepts, principles, and methods for dealing with imprecise information that is supposed to be approximate rather than exact.
* Secondly, it is viewed as an abstraction of the various many-valued logics, which have been studied in the area of semantic logic since the beginning of the twentieth century.

**Evolutionary Computation**

**Some people would claim that things like joy, love and beauty belong to a different category from science and can-not be described in scientific terms, but I think they can now be explained by the theory of evolution.**

**Stephen Hawking**

**Theoretical, physicist, cosmologist and author**

If we look at the history of human evolution, we will find that our today’s intelligence is not achieved over night. It is the outcome of countless years of biological evolution. A deep study of the biological evolutionary process, as to how the natural intelligence developed, from the brains of early-man to today’s intelligent brains, could clear the point of developing intelligence in machines. Thus the field of Evolutionary Computation (EC) developed.

* An expert system, also known as a knowledge based system, is a computer based system that can make intelligent decisions by emulating the decision making abilities of human experts.
* Expert systems are rule based systems and they are part of the artificial intelligence.
* Expert systems have the abilities that they can change their decisions and make new decisions based on the external factors. Some expert systems are designed to take place of a human in an application, while some others are designed to aid the human.
* It is also comprised of Genetic Algorithm (GA) which helps in understanding the principles of imitating the evolution of individual structures through processes of selection, mutation and reproduction.
* To be practically advantageous in solving optimization problems.
* It can be applied to problems whose satisfactory heuristic solutions are not available. Involving the Heuristically informed search techniques, which enables to reduce the extremely large search space by utilizing GA as a candidate technique for the same purpose; it helps in inhibiting the capacity for population-based systematic random searches. Simulated annealing is also one algorithm for solving such heuristic search problems as well as optimization problems.
* Globally applied to solve many problems including optimization, automatic programming, machine learning, operations research, bioinformatics, and social systems.
* Some application areas of expert systems are: online medical systems for diagnosing a problem, financial loan/credit decisions, legal matters, robotics, and engineering design. One of the main problems in expert systems is the knowledge acquisition.

The main components of an expert system are***: knowledge base, interface engine, and user interface.***

* The knowledge base is probably the most important part of any expert system. This is where the intelligence of the system is stored. Expert systems in general can acquire new knowledge by their sensors or by training and extend their knowledge bases so that they can easily respond to new problems. The knowledge is stored in the form of IF-THEN-ELSE statements.
* The interface engine is between the knowledge base and the user. The interface engine makes decisions by following the conditions and the requirements before it comes to an outcome and presents a solution to the user.
* The user interface is usually in the form of natural language used daily by the user in everyday life.
* There are basically two types of programming languages: algorithmic and symbolic. Traditional programming languages such as Pascal, Basic C, and Fortran are algorithmic, also known as procedural languages, where it is difficult to implement logical inferences in these languages. Several symbolic languages have been developed over the years for expert systems development, such as Prolog, Lisp, Clips and so on.

.

**Application of Soft Computing**

We can implement the concept of soft computing in many applications; we will discuss a few of them below:

* Handwriting recognition
* It is one of the demanding parts of computer science. It can translate multilingual documents and sort the various scripts accordingly.
* It uses the concept of “block-level technique” where the system recognizes the particular script from a number of script documents given.
* It uses a Discrete Cosine Transform (DCT) and discrete wavelets Transform (DWT) together, which classify the scripts according to their features.
* Image Processing and Data Compression
* Image analysis is one of the most important parts of the medical field. It is a high-level processing technique which includes recognition and bifurcation of patterns.
* Using soft computing solves the problem of computational complexity and efficiency in the classification. Techniques of soft computing include Genetic Algorithms, Genetic Programming, Classifier Systems, Evolution Strategies, artificial life, and a few others, which are used here.
* These algorithms give the fastest solutions to pattern recognition. These help in analyzing the medical images obtained from microscopes as well as examine the X-rays.
* Automotive systems and manufacturing
* The use of soft computing has solved a major misconception that the automobile industry is slow to adapt. Fuzzy logic is a technique used in vehicles to build classic control methods.
* 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, antiskid steering, etc.
* 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.
* 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.
* Power systems
  + 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.
* Recent developments in soft computing
* People have started using techniques of soft computing like fuzzy sets theory, neural nets, fuzzy neuro system, adaptive neuro-fuzzy inference system (ANFIS), for driving various numerical simulation analysis.
* Soft computing has helped in modeling the processes of machines with the help of artificial intelligence.
* Also, there are certain areas where soft computing is in budding stages only and is expected to see a massive evolution:
* Big Data
* Recommender system
* Behavior and decision science
* Mechanical Engineering
* Computer Engineering
* Civil Engineering
* So, soft computing is implemented in applications that require human-like thinking such as games like poker and chess etc.

So, this is all about soft computing, it plays an important role in developing intelligent machines using the field of science and engineering. Soft computing reflects the fact that, unlike modern computers, the human brain has a remarkable ability to store and process information that is not clear, imprecise and uncertain and even can’t be categorised.

**UNIT-2**

**Unit-2- Neural Networks:** Basic concepts of Neural Networks, Model of Artificial Neuron, Neural Network Architectures, Feed forward neural network, Back propagation network (BPN) Radial Basis function network, Supervised, Unsupervised, and Reinforcement Learning, Back propagation Learning.

**Basic concepts of Neural Networks**

Computers have high processing power and memory and can create a complex numerical problem in a short time easily. But the real-world activities are associated with vision, speech, pattern recognition, and natural language, Computer fails while performing these tasks. This is because, computers, with or without Artificial Intelligence, require an algorithmic approach i.e. the problem has to be presented as an algorithm, and where the real problem function cannot be constructed as an algorithm. Also, computers cannot do what our brain does i.e. learn and adapt.

***The human brain is fascinating, and even the most advanced computers cannot compete with the brain in one thing - our brain reads, and analyses the problem, and finally adapts to the situation.***

Keeping this in mind, scientists have developed many computer models inspired by Biological Neural Networks (BNN) that can provide a solution to the problems of natural functions. One of them is Artificial Neural Networks or ANN.

**Computation in the Brain**

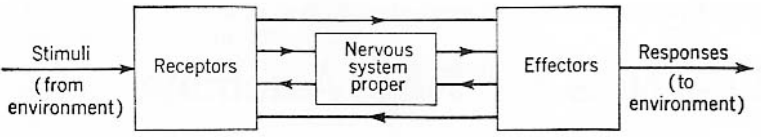
The neural system is not just the neural network; the overall neural system of the human body comprised of three primary constituents: Receptors, A neural network, and Effectors.

1. The function of receptors is to receive the stimuli either internally (in the form of signals from internal organs like heart, liver etc.) or externally (through sensory organs, like eyes, ears, tongue etc.), this information is then passed onto the neurons in the form of electrical impulses.
2. The neural network then in response to the incoming information took a decision of the outputs.
3. At last, the effectors convert incoming electrical impulses out of the neural network into responses to the external signals received by the atmosphere.

Example: Suppose, a ball of fire is coming with enormous speed towards you?

1. Receptors, will send this information to the neural network,
2. Neural network will take a decision in response to the incoming danger, by generating a signal like run, or lean down, or move to right or left, according to the circumstances.
3. In response to the generated signal, the effectors will make the body parts to respond, like legs will start running or anything like that.

Figure shows the bidirectional communication between stages for feedback.



**Figure: Three stages of Biological neural network (BNN).**

**Biological Nervous System**

Biological nervous system is the most important part of many living things, in particular, human beings.

* There is a part called brain at the centre of human nervous system.
* In fact, any biological nervous system consists of a large number of interconnected processing units called neurons.
* Each neuron is approximately 10µ and they can operate in parallel.
* Typically, a human brain consists of approximately 1011 neurons communicating with each other with the help of electrical impulses.

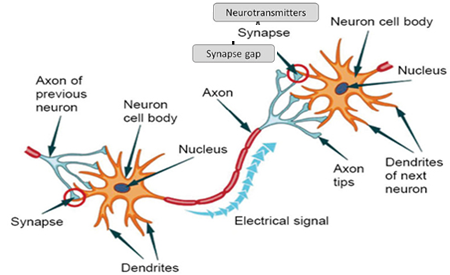
**Neuron: Basic Unit of Nervous System**

The human brain consists of millions of neural cells that process information. Each cell performs like a simple processor. The collection and interconnection of all these cells makes a neural network. The primary component of the neural network is called a neuron. As shown in Figure, a neuron is mainly comprised of three modules: dendrites, soma, and axon.

1. **Dendrites** are the tree-like structure that receives the signal from neighboring neurons and passes on to the cell body.
2. Each neuron consists of a body, called the **Soma**. Soma manages the incoming activations, performs various operations on them and results into output activations.
3. **Axons** play the role of transmission lines to carry activations to other neurons.
4. **Synapses** allow signal transmission between the axons and dendrites. Thus, a complex network of neurons is created in the human brain.

<https://www.youtube.com/watch?v=OvVl8rOEncE>

1. There is a gap between the nerve endings of one neuron and dendrite of the following neuron where signals are transmitted as chemical signals called **neuro-transmitters**.



**Figure: A Biological Neuron.**

* There is a chemical in each neuron called neurotransmitter.
* A signal (also called sense) is transmitted across neurons by this chemical.
* That is, all inputs from other neuron arrive to neurons through dendrites.
* These signals are accumulated at the synapse of the neuron and then serve as the output to be transmitted through the neuron.
* An action may produce an electrical impulse, which usually lasts for about a millisecond.
* Note that this pulse generated due to an incoming signal and all signal may not produce pulses in axon unless it crosses a threshold value.
* Also, note that an action signal in axon of a neuron is commutative signals arrive at dendrites which summed up at soma.

**Neural Network**

Neural networks reflect the behavior of the human brain, allowing computer programs to recognize patterns and solve common problems in the fields of AI, machine learning, and deep learning.

Neural networks were first proposed in 1944 by Warren McCullough and Walter Pitts. They built a circuitry system intended to approximate the functioning of the human brain that ran simple algorithms.

## A neural network is a method in artificial intelligence that teaches computers to process data in a way that is inspired by the human brain. It is a type of machine learning process, called deep learning that uses interconnected nodes or neurons in a layered structure that resembles the human brain. It creates an adaptive system that computers use to learn from their mistakes and improve continuously. Thus, artificial neural networks attempt to solve complicated problems, like summarizing documents or recognizing faces, with greater accuracy.

## NeuronsNeurons

* Scientists agree that our brain has around 100 billion neurons.
* These neurons have hundreds of billions connections between them.
* Neurons are the fundamental units of our brain and nervous system.
* The neurons are responsible for receiving input from the external world, for sending output (commands to our muscles), and for transforming the electrical signals in between.
* In fact, the human brain is a highly complex structure viewed as a massive, highly interconnected network of simple processing elements called neurons.
* Artificial neural networks (ANNs) or simply we refer it as neural network (NNs), which are simplified models (i.e. imitations) of the biological nervous system, and obviously, therefore, have been motivated by the kind of computing performed by the human brain.
* The behavior of a biological neural network can be captured by a simple model called **artificial neuron or perceptron**.

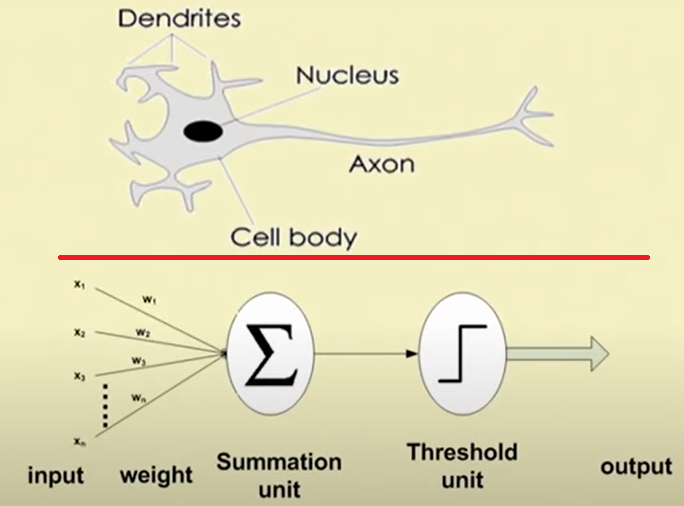
**How do neural networks work?**

The human brain is the inspiration behind neural network architecture. Human brain cells, called neurons, form a complex, highly interconnected network and send electrical signals to each other to help humans process information. Similarly, an artificial neural network is made of artificial neurons that work together to solve a problem. Artificial neurons are software modules, called nodes, and artificial neural networks are software programs or algorithms that, at their core, use computing systems to solve mathematical calculations.

**Simple neural network architecture**

A basic neural network has interconnected artificial neurons in three layers:

* **Input Layer -** Information from the outside world enters the artificial neural network from the input layer. Input nodes process the data, analyze or categorize it, and pass it on to the next layer.
* **Hidden Layer** - Hidden layers take their input from the input layer or other hidden layers. Artificial neural networks can have a large number of hidden layers. Each hidden layer analyzes the output from the previous layer, processes it further, and passes it on to the next layer.
* **Output Layer -** The output layer gives the final result of all the data processing by the artificial neural network. It can have single or multiple nodes. For instance, if we have a binary (yes/no) classification problem, the output layer will have one output node, which will give the result as 1 or 0. However, if we have a multi-class classification problem, the output layer might consist of more than one output node.

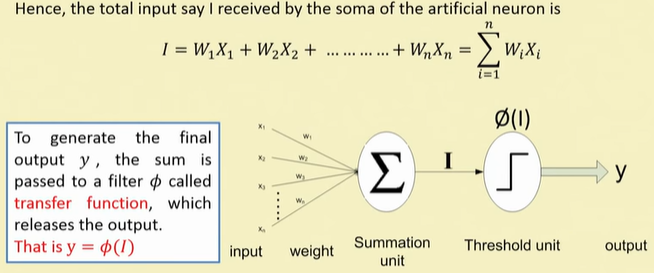


We may note that a neutron is a part of an interconnected network of nervous system and serves the following.

* + Compute input signals
  + Transportation of signals (at a very high speed)
  + Storage of information
  + Perception, automatic training and learning

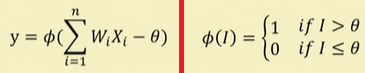
We also can see the analogy between the biological neuron and artificial neuron. Truly, every component of the model (i.e. artificial neuron) bears a direct analogy to that of a biological neuron. It is this model which forms the basis of neural network (i.e. artificial neural network)

* Note that, a biological neuron receives all inputs through the dendrites, sums them and produces an output if the sum is greater than a threshold value.
* The input signals are passed on to the cell body through the synapse, which may accelerate or retard an arriving signal.
* It is this acceleration or retardation of the input signals that is modelled by the weights.
* An effective synapse, which transmits a stronger signal will have a correspondingly larger weights while a weak synapse will have smaller weights.
* Thus, weights here are multiplicative factors of the Inputs to account for the strength of the synapse



* A very commonly known transfer function is the thresholding function denoted as Ø***.***
* In this thresholding function, sum (i.e. I) is compared with a threshold value say θ.
* If the value of *I* is greater than θ, then the output is 1 else it is 0 (this is just like a simple linear filter).

In other words,



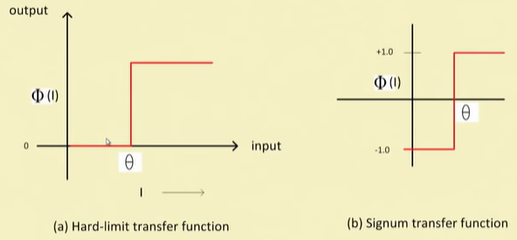
Such a Ø is called a **step function** also called as **Heaviside function**.

**Hard-limit transfer function:** The transformation we have just discussed is called hard-limit transfer function. It is generally used in perception neuron. In other words,



**Linear transfer function:** The output of the transfer function is made equal to its input (normalized) and its lies in the range of -1.0 to +1.0. It is also known as Signum or Quantizer function and it defined as

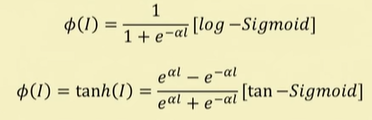




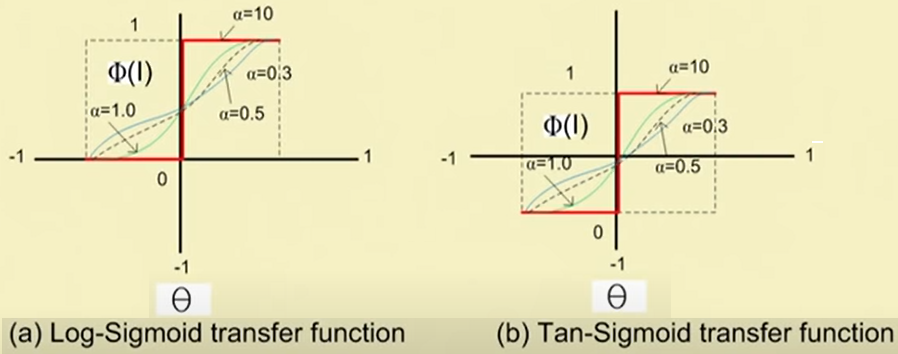
**Note:** Transfer functions are used for transformation purposes i.e., translates the input signals to output signals. On the other hand, activation function checks for the output if it meets a certain threshold and either outputs zero or one. Some examples of non-linear transfer functions are softmax and sigmoid.

**Transfer Functions in ANN**

**Sigmoid transfer function:** This function is a continuous function that varies gradually between the asymptotic values 0 and 1 (called log-sigmoid) or -1 and +1 (called Tan- sigmoid) threshold function and is given by



Here α is the coefficient of transfer function.



**Advantages of ANN**

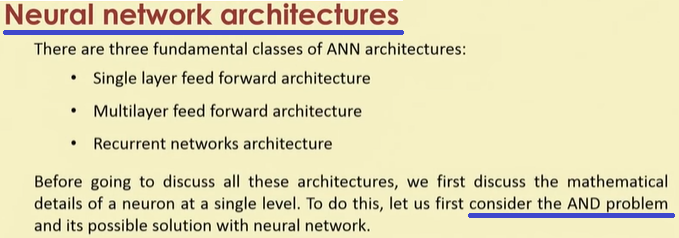
* ANNs exhibits mapping capabilities, that is, they can map input patterns to their associated output pattern.
* The ANNs learn by examples. Thus, an ANN architecture can be trained with known example of a problem before they are tested for their inference capabilities on unknown instance of the problem. In other words, they can identify new objects previous untrained.
* The ANNs possess the capability to generalize. This is the power to apply in application where exact mathematical model to problem are not possible.
* The ANNs are robust system and fault tolerant. They can therefore, recall full patterns from incomplete, partial or noisy patterns.
* The ANNS can process information in parallel, at high speed and in a distributed manner.
* Thus a massively parallel distributed processing system made up of highly interconnected (artificial) neural computing elements having ability to learn and acquire knowledge is possible.

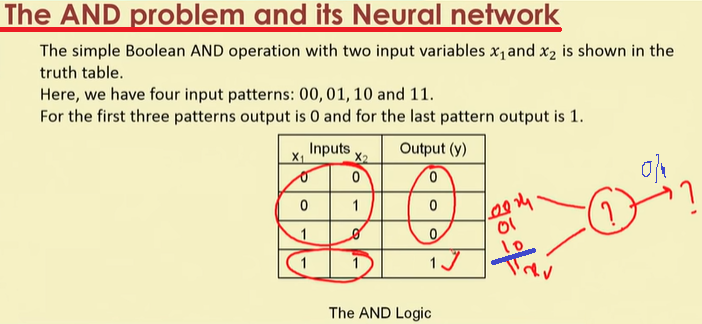
# ****Neural Network Applications****

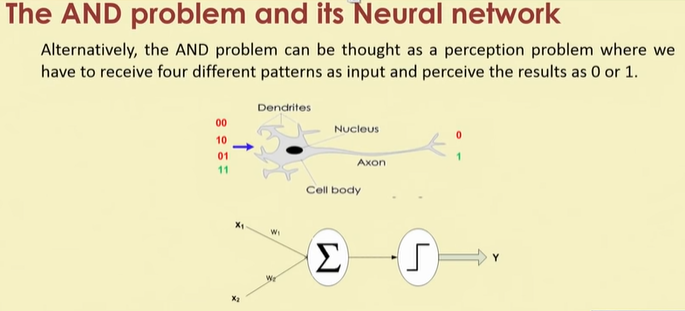
Artificial Neural Network (ANN) is based on the processing of human brain. It is developed to simplify tasks that are easy for human but difficult for machines. The algorithms can be used to model complex patterns and prediction problems with the help of ANN.

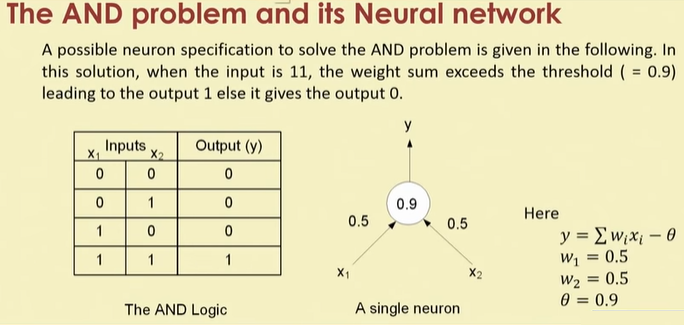
* **Telecommunications:** Image and data compression, automated information services, real-time spoken language translation
* **Electronics:** Code sequence prediction, IC chip layout, chip failure analysis, machine vision, voice synthesis.
* **Anomaly Detection:** can be trained to generate an output when something unusual occurs that misfits the pattern.
* **Speech:** Speech recognition, speech classification, text to speech conversion.
* **Signal Processing:** Neural networks can be trained to process an audio signal and filter it appropriately in the hearing aids.
* **Medical:** Cancer cell analysis, EEG and ECG analysis, prosthetic design, transplant time optimizer
* **Aerospace:** Autopilot aircrafts, aircraft fault detection.
* **Industrial:** Manufacturing process control, product design and analysis, quality inspection systems, welding quality analysis, paper quality prediction, chemical product design analysis, dynamic modelling of chemical process systems, machine maintenance analysis, project bidding, planning, and management.
* **Software:** Pattern Recognition in facial recognition, optical character recognition, etc.
* **Automotive:** Automobile guidance systems.
* **Time Series Prediction:** ANNs are used to make predictions on stocks and natural calamities.
* **Signal Processing:** Neural networks can be trained to process an audio signal and filter it appropriately in the hearing aids.
* **Military:** Weapon orientation and steering, target tracking, object discrimination, facial recognition, signal/image identification.
* **Financial:** Real estate appraisal, loan advisor, mortgage screening, corporate bond rating, portfolio trading program, corporate financial analysis, currency value prediction, document readers, credit application evaluators.
* **Transportation:** Truck Brake system diagnosis, vehicle scheduling, routing systems.
* **Control:** ANNs are often used to make steering decisions of physical vehicles.

**ANN ARCHITECTURES**







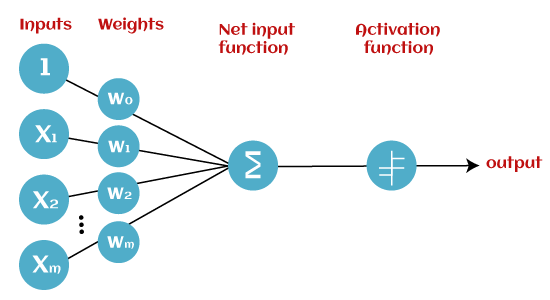


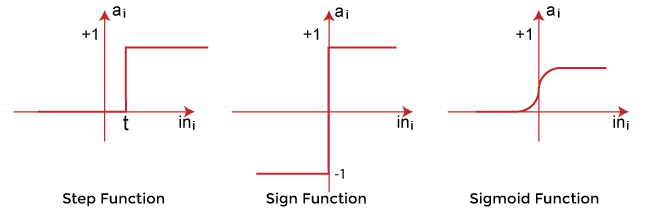
**Artificial Neuron/Perceptrons**

Perceptron is Machine Learning algorithm for supervised learning of various binary classification tasks.

* A **Perceptron** is an **Artificial Neuron**
* It is the simplest possible **Neural Network**
* **Neural Networks** are the building blocks of **Machine Learning**.

Mr. Frank Rosenblatt invented the perceptron model as a binary classifier which contains three main components. These are as follows:



* **Input Nodes or Input Layer:** This is the primary component of Perceptron which accepts the initial data into the system for further processing. Each input node contains a real numerical value.
* **Wight and Bias:** Weight parameter represents the strength of the connection between units. This is another most important parameter of Perceptron components. Weight is directly proportional to the strength of the associated input neuron in deciding the output. Further, Bias can be considered as the line of intercept in a linear equation.
* **Activation Function:** These are the final and important components that help to determine whether the neuron will fire or not. Activation Function can be considered primarily as a step function.

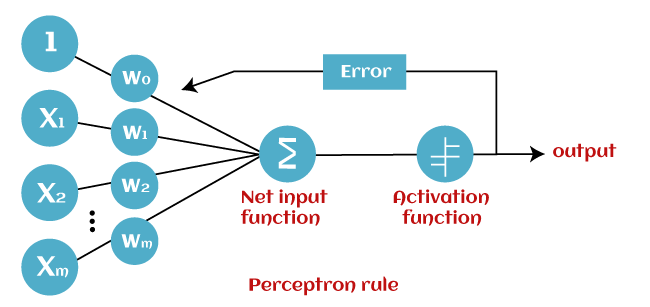
Types of Activation functions:

* Sign function
* Step function, and
* Sigmoid function

The data scientist uses the activation function to take a subjective decision based on various problem statements and forms the desired outputs. Activation function may differ (e.g., Sign, Step, and Sigmoid) in perceptron models by checking whether the learning process is slow or has vanishing or exploding gradients.

## How does Perceptron work?

In Machine Learning, Perceptron is considered as a single-layer neural network that consists of four main parameters named input values (Input nodes), weights and Bias, net sum, and an activation function. The perceptron model begins with the multiplication of all input values and their weights, then adds these values together to create the weighted sum. Then this weighted sum is applied to the activation function 'f' to obtain the desired output. This activation function is also known as the **step function** and is represented by **'f'**.



This step function or Activation function plays a vital role in ensuring that output is mapped between required values (0,1) or (-1,1). It is important to note that the weight of input is indicative of the strength of a node. Similarly, an input's bias value gives the ability to shift the activation function curve up or down.

Perceptron model works in two important steps as follows:

**Step-1:** In the first step first, multiply all input values with corresponding weight values and then add them to determine the weighted sum. Mathematically, we can calculate the weighted sum as follows:

∑wi\*xi = x1\*w1 + x2\*w2 +…wn\*xn

Add a special term called **bias 'b'** to this weighted sum to improve the model's performance.

**∑wi\*xi + b**

**Step-2:** In the second step, an activation function is applied with the above-mentioned weighted sum, which gives us output either in binary form or a continuous value as follows:

**Y = f(∑wi\*xi + b)**

## Types of Perceptron Models

Based on the layers, Perceptron models are divided into two types. These are as follows:

1. Single-layer Perceptron Model
2. Multi-layer Perceptron model

### Single Layer Perceptron Model: This is one of the easiest Artificial neural networks (ANN) types. A single-layered perceptron model consists feed-forward network and also includes a threshold transfer function inside the model. The main objective of the single-layer perceptron model is to analyze the linearly separable objects with binary outcomes.

In a single layer perceptron model, its algorithms do not contain recorded data, so it begins with inconstantly allocated input for weight parameters. Further, it sums up all inputs (weight). After adding all inputs, if the total sum of all inputs is more than a pre-determined value, the model gets activated and shows the output value as +1.

If the outcome is same as pre-determined or threshold value, then the performance of this model is stated as satisfied, and weight demand does not change. However, this model consists of a few discrepancies triggered when multiple weight inputs values are fed into the model. Hence, to find desired output and minimize errors, some changes should be necessary for the weights input.

**"Single-layer perceptron can learn only linearly separable patterns."**

### Multi-Layered Perceptron Model

Like a single-layer perceptron model, a multi-layer perceptron model also has the same model structure but has a greater number of hidden layers.

The multi-layer perceptron model is also known as the Backpropagation algorithm, which executes in two stages as follows:

* **Forward Stage:** Activation functions start from the input layer in the forward stage and terminate on the output layer.
* **Backward Stage:** In the backward stage, weight and bias values are modified as per the model's requirement. In this stage, the error between actual output and demanded originated backward on the output layer and ended on the input layer.

Hence, a multi-layered perceptron model has considered as multiple artificial neural networks having various layers in which activation function does not remain linear, similar to a single layer perceptron model. Instead of linear, activation function can be executed as sigmoid, TanH, ReLU, etc., for deployment.

***A multi-layer perceptron model has greater processing power and can process linear and non-linear patterns.*** Further, it can also implement logic gates such as AND, OR, XOR, NAND, NOT, XNOR, NOR.

**Advantages of Multi-Layer Perceptron:**

* A multi-layered perceptron model can be used to solve complex non-linear problems.
* It works well with both small and large input data.
* It helps us to obtain quick predictions after the training.
* It helps to obtain the same accuracy ratio with large as well as small data.

**Disadvantages of Multi-Layer Perceptron:**

* In Multi-layer perceptron, computations are difficult and time-consuming.
* In multi-layer Perceptron, it is difficult to predict how much the dependent variable affects each independent variable.
* The model functioning depends on the quality of the training.

## Perceptron Function: Perceptron function ''f(x)'' can be achieved as output by multiplying the input 'x' with the learned weight coefficient 'w'. Mathematically, we can express it as follows:

**f(x)=1; if w.x+b>0**

**otherwise, f(x)=0**

* 'w' represents real-valued weights vector
* 'b' represents the bias
* 'x' represents a vector of input x values.

## Characteristics of Perceptron

The perceptron model has the following characteristics.

1. Perceptron is a machine learning algorithm for supervised learning of binary classifiers.
2. In Perceptron, the weight coefficient is automatically learned.
3. Initially, weights are multiplied with input features, and the decision is made whether the neuron is fired or not.
4. The activation function applies a step rule to check whether the weight function is greater than zero.
5. The linear decision boundary is drawn, enabling the distinction between the two linearly separable classes +1 and -1.
6. If the added sum of all input values is more than the threshold value, it must have an output signal; otherwise, no output will be shown.

## Limitations of Perceptron Model

1. The output of a perceptron can only be a binary number (0 or 1) due to the hard limit transfer function.
2. Perceptron can only be used to classify the linearly separable sets of input vectors. If input vectors are non-linear, it is not easy to classify them properly.

**How Do You Train An Algorithm?**

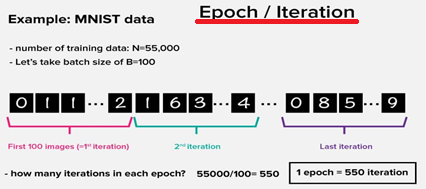
Neural networks are trained like any other algorithm. You want to get some results and provide information to the network to learn from. For example, we want our neural network to distinguish between photos of cats and dogs and provide plenty of examples.

Delta is the difference between the data and the output of the neural network. We use calculus magic and repeatedly optimize the weights of the network until the delta is zero. Once the delta is zero or close to it, our model is correctly able to predict our example data.

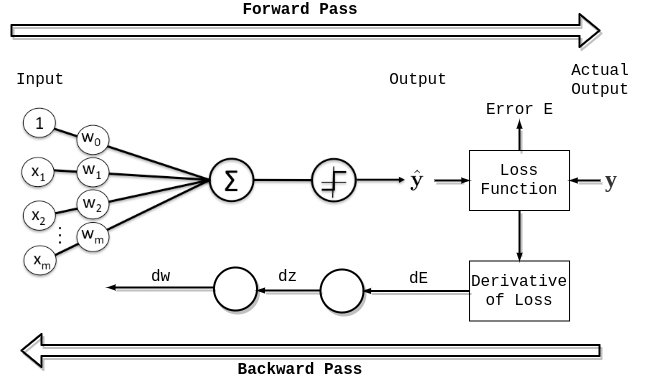
**Iteration -** This is a kind of counter that increases every time the neural network goes through one training set. In other words, this is the total number of training sets completed by the neural network.

**Epoch -** The epoch increases each time we go through the entire set of training sets. The more epochs there are, the better is the training of the model.

**Batch -** Batch size is equal to the number of training examples in one forward/backward pass. The higher the batch size, the more memory space you’ll need.

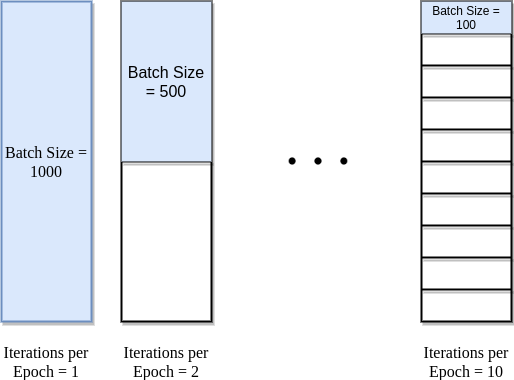


**An epoch means training the neural network with all the training data for one cycle. In an epoch, we use all of the data exactly once. A forward pass and a backward pass together are counted as one pass:**



**An epoch is made up of one or more**[**batch**](https://www.baeldung.com/cs/neural-networks-epoch-vs-iteration#3-batch)**es, where we use a part of the dataset to train the neural network. We call passing through the training examples in a batch an**[**iteration**](https://www.baeldung.com/cs/neural-networks-epoch-vs-iteration#2-iteration)**.**

An epoch is sometimes mixed with an iteration. To clarify the concepts, let’s consider a simple example where we have 1000 data points as presented in the figure below:

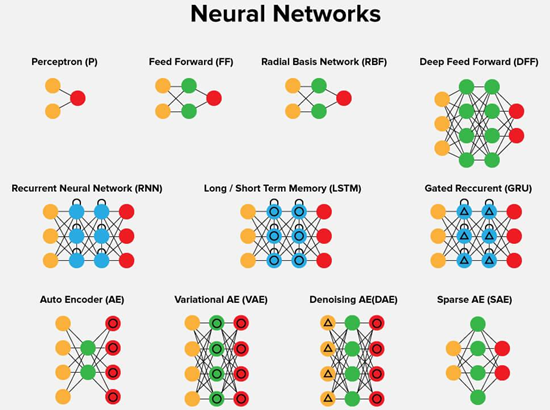


If the batch size is 1000, we can complete an epoch with a single iteration. Similarly, if the batch size is 500, an epoch takes two iterations. So, if the batch size is 100, an epoch takes 10 iterations to complete. Simply, for each epoch, the required number of iterations times the batch size gives the number of data points.

We can use multiple epochs in training. In this case, the neural network is fed the same data more than once.

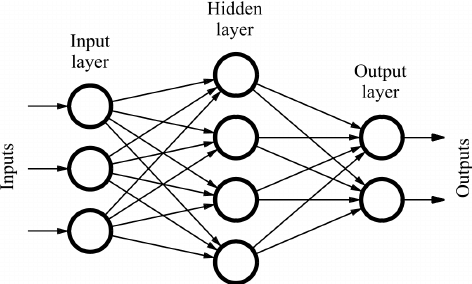
**What Kinds of Neural Networks Exist?**

There are so many different neural networks out there that it is simply impossible to mention them all. If you want to learn more about this variety, [visit the neural network zoo](https://www.asimovinstitute.org/neural-network-zoo/) where you can see them all represented graphically.



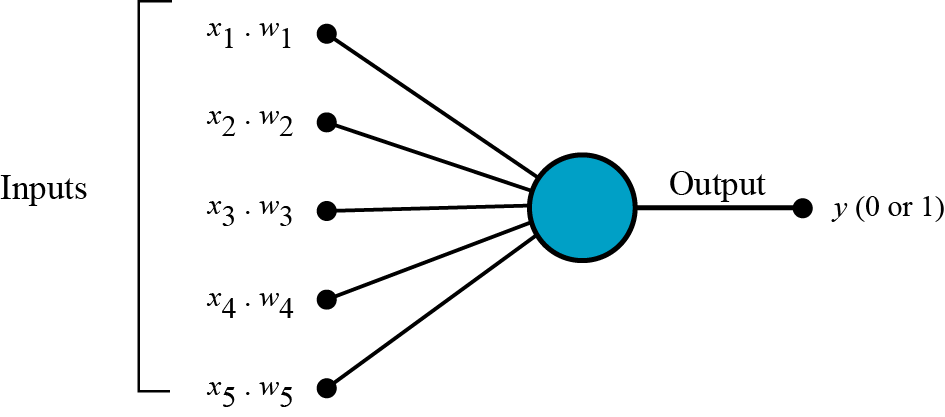
**Feed-Forward Neural Networks**

A Feed Forward [Neural Network](https://deepai.org/machine-learning-glossary-and-terms/neural-network) is an artificial neural network in which the connections between nodes does not form a cycle. The opposite of a feed forward neural network is a [recurrent neural network](https://deepai.org/machine-learning-glossary-and-terms/recurrent-neural-network), in which certain pathways are cycled. The feed forward model is the simplest form of neural network as information is only processed in one direction. While the data may pass through multiple hidden nodes, it always moves in one direction and never backwards.



## How does a Feed Forward Neural Network work?

A Feed Forward Neural Network is commonly seen in its simplest form as a single layer [perceptron](https://deepai.org/machine-learning-glossary-and-terms/perceptron). In this model, a series of inputs enter the layer and are multiplied by the weights. Each value is then added together to get a sum of the weighted input values. If the sum of the values is above a specific threshold, usually set at zero, the value produced is often 1, whereas if the sum falls below the threshold, the output value is -1. The single layer perceptron is an important model of feed forward neural networks and is often used in classification tasks. Furthermore, single layer perceptrons can incorporate aspects of [machine learning](https://deepai.org/machine-learning-glossary-and-terms/machine-learning). Using a property known as the delta rule, the neural network can compare the outputs of its nodes with the intended values, thus allowing the network to adjust its weights through training in order to produce more accurate output values. This process of training and learning produces a form of a gradient descent. In multi-layered perceptrons, the process of updating weights is nearly analogous, however the process is defined more specifically as back-propagation. In such cases, each hidden layer within the network is adjusted according to the output values produced by the final layer.



### Applications of Feed Forward Neural Networks

While Feed Forward Neural Networks are fairly straightforward, their simplified architecture can be used as an advantage in particular machine learning applications. For example, one may set up a series of feed forward neural networks with the intention of running them independently from each other, but with a mild intermediary for moderation. Like the human brain, this process relies on many individual [neurons](https://deepai.org/machine-learning-glossary-and-terms/neuron) in order to handle and process larger tasks. As the individual networks perform their tasks independently, the results can be combined at the end to produce a synthesized, and cohesive output.