Unit 5 Machine Learning

Lecture 1

What is learning

- Learning is the processof acquiring new or modifying existing knowledge, behaviors, skills, values, or preferences
- Evidence that learning has occurred may be seen in changes in behavior from simple to complex.

What is Machine Learning?

- The machine Learning denotes changes in the systems that are adaptive in the sense that they enable the system to do the same task more effectively the next time.

- Like human learning from past experiences, computer system learns from data, which represent some "past experiences" of an application domain.

- Therefore, Machine learning gives computers the ability to learn without being explicitly programmed.

Why Machine Learning ?

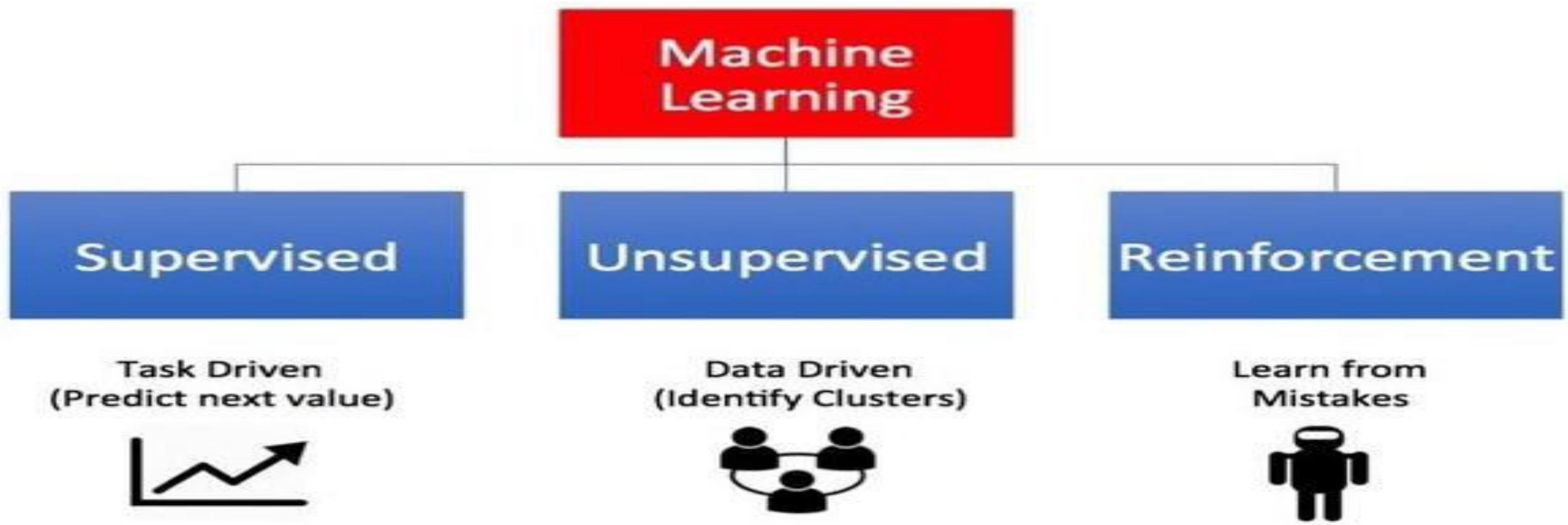
- To learn a target function (relation between input and output)that can be used to predict the values of a discrete class attribute,
  - e.g., male or female, and high-risk or low risk, etc.

- To model the underlying structure or distribution in the data in order to learn more about the data.

- To learns behavior through trial-and-error interactions with a dynamic environment.

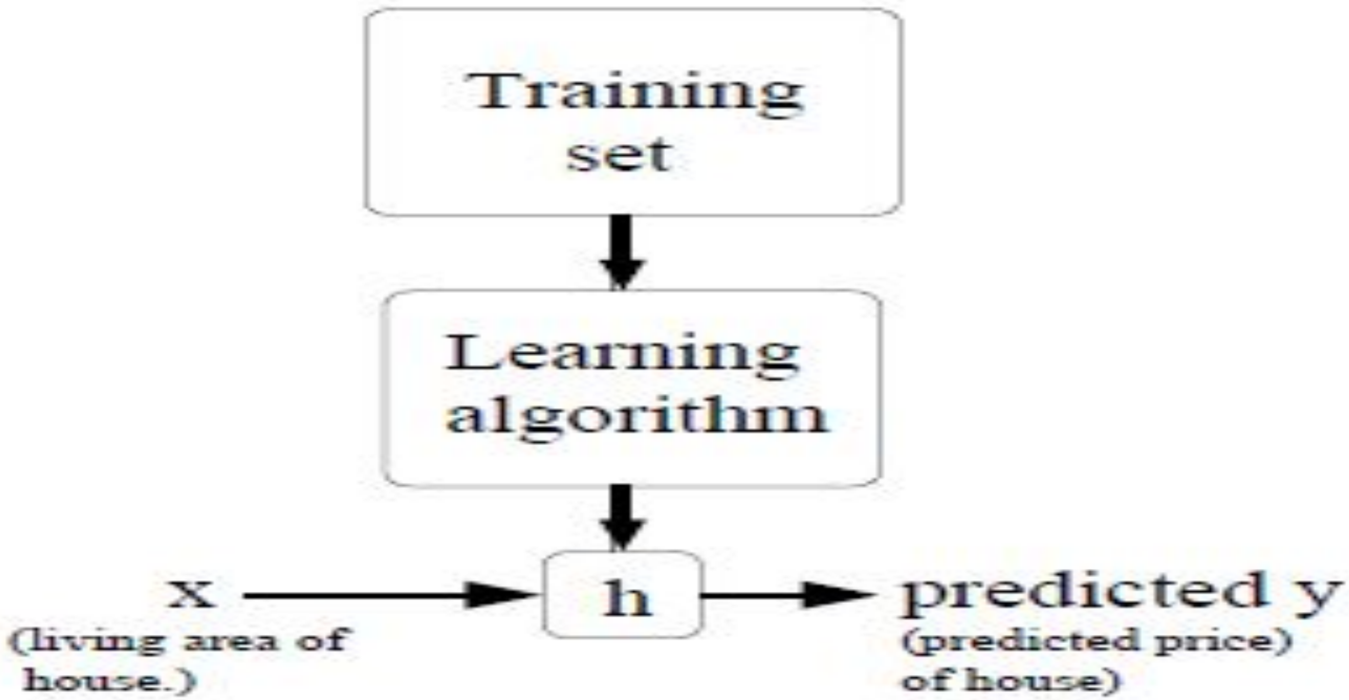
Types

- Based on training set machine learning algorithms are classified into the following three categories:



Supervised Learning

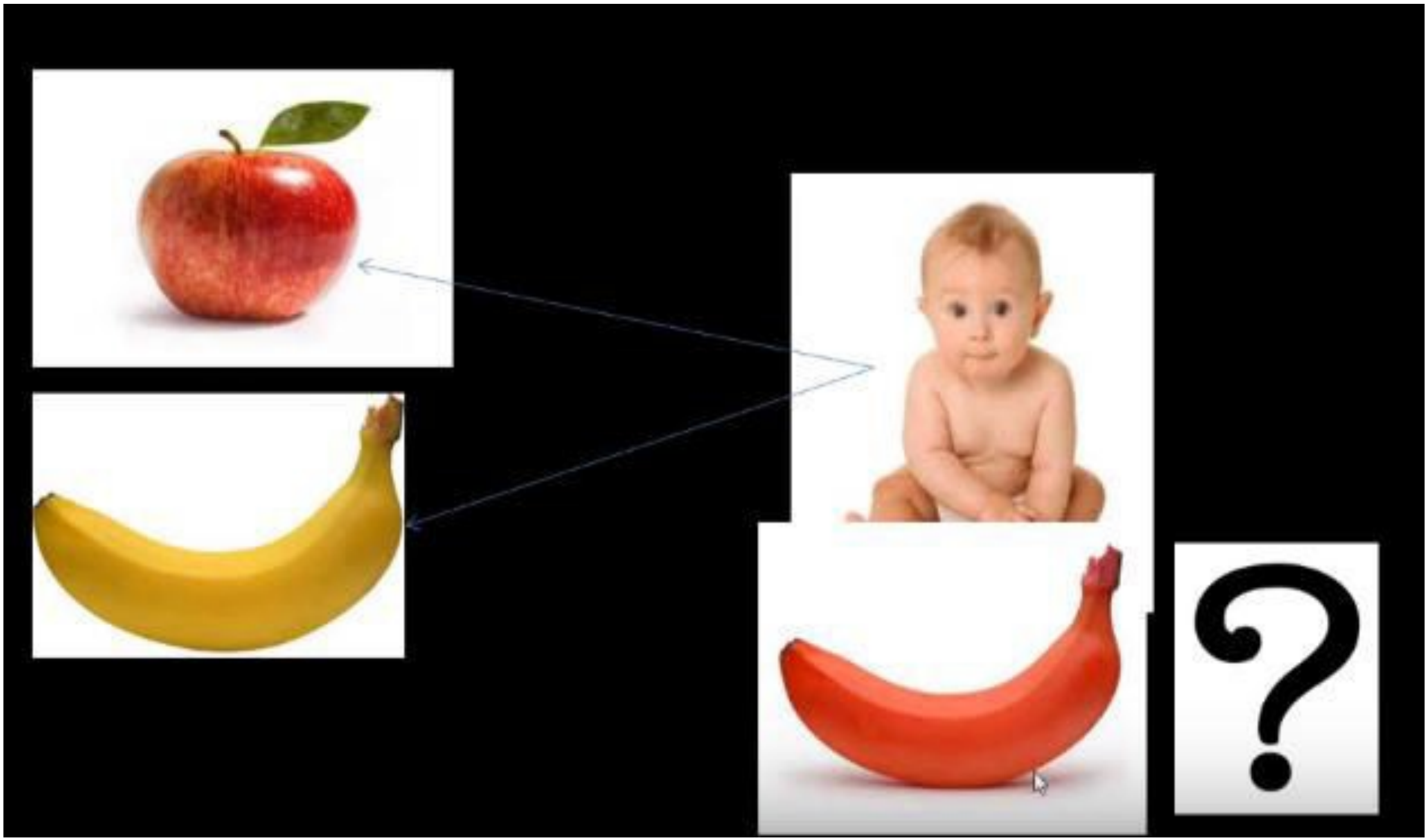
- Supervised learning is where you have input variables (x) and an output variable (Y) and you use an algorithm to learn the mapping function from the input to the output. Y = f(X)
- Learning stops when the algorithm achieves an acceptable level of performance.
- when you have new input data (x) that you can predict the output variables (Y) for that data


Training
set
Learning
algorithm
h
x
(living area of
house.)
predicted y
(predicted price)
of house)

Supervised Learning

- It is called supervised learning because the process of an algorithm learning from the training dataset can be thought of as a teacher supervising the learning process.
- We know the correct answers, the algorithm iteratively makes predictions on the training data and is corrected by the teacher.

Supervised Learning



Unsupervised Learning

- Unsupervised learning is where you only have input data (X) and no corresponding output variables(targets/ labels).
- The goal for unsupervised learning is to model the underlying structure or distribution in the data in order to learn more about the data.
- These are called unsupervised learning because unlike supervised learning above there is no correct answers and there is no teacher.
- Algorithms are left to their own devises to discover and present the interesting structure in the data.

E.g., Clustering

Unsupervised Learning

Unsupervised Learning

Reinforcement Learning

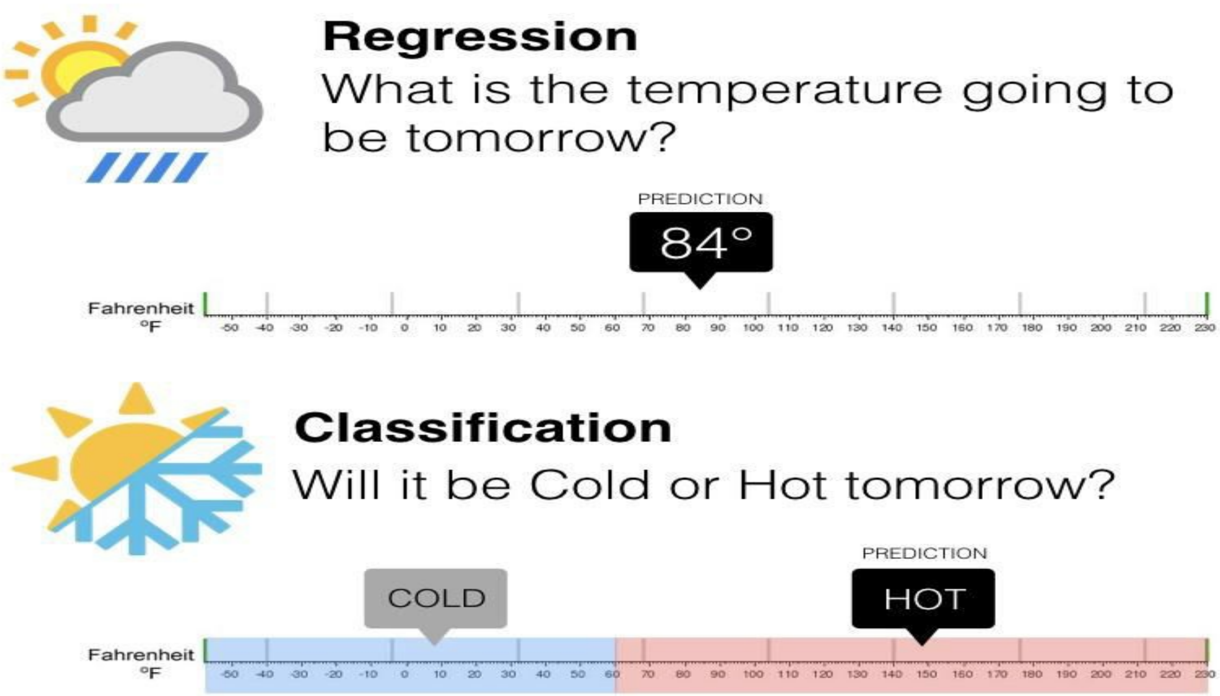
- Is learning behavior through trial-and-error interactions with a environment.
- Is learning how to act in order to maximize a reward (Encouragements).
- Reinforcement learning emphasizes learning feedback that evaluates the learner's performance without providing standards of correctness in the form of behavioral targets.
- Example: Bicycle learning, game playing, etc.

Supervised Learning Algorithm

- Classification:
  - To predict the outcome of a given sample where the output variable is in the form of categories(discrete). Examples include labels such as, sick and healthy.
- Regression:

  -To predict the outcome of a given sample where the output variable is in the form of real values(continuous). Examples include real-valued labels denoting the amount of rainfall, the height of a person

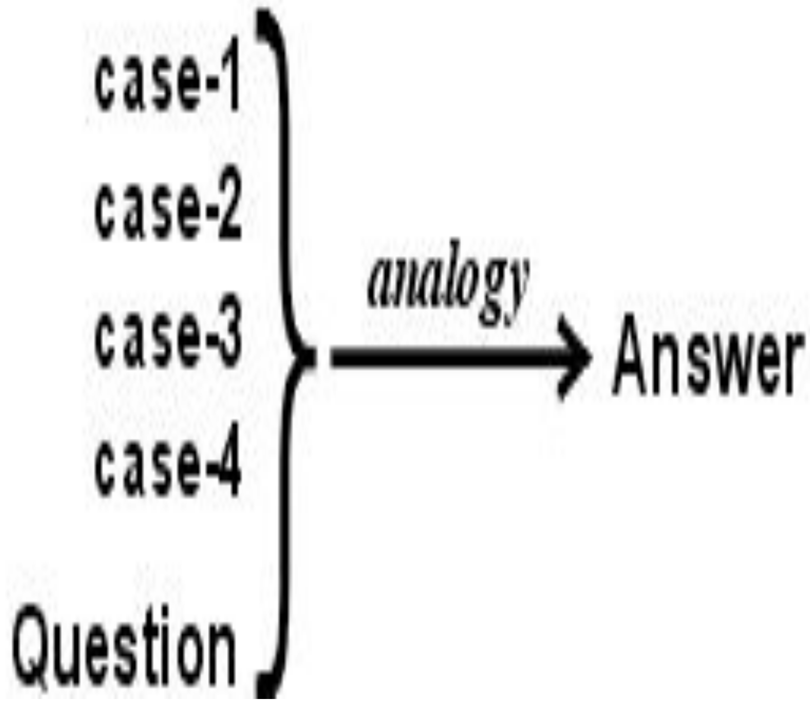
Contd.



Regression
What is the temperature going to be tomorrow?
PREDICTION
84°
Fahrenheit °F
-50 -40 -30 -20 -10 0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230
Classification
Will it be Cold or Hot tomorrow?
PREDICTION
COLD
HOT
Fahrenheit °F
-50 -40 -30 -20 -10 0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230

Learning by Analogy:

- Reasoning by analogy generally involves abstracting details from a particular set of problems and resolving structural similarities between previously distinct problems.
- Analogical reasoning refers to this process of recognition and then applying the solution from the known problem to the new problem.
- Such a technique is often identified as *case- based reasoning*. Analogical learning generally involves developing a set of mappings between features of two instances.


case-1
case-2
case-3
case-4
Question
analogy
Answer

The question in above figure represents some known aspects of a new case, which has unknown aspects to be determined.

- In deduction, the known aspects are compared (by a version of structure mapping called *unification*) with the premises of some implication.
- Then the unknown aspects, which answer the question, are derived from the conclusion of the implication.
- In analogy, the known aspects of the new case are compared with the corresponding aspects of the older cases.
- The case that gives the best match may be assumed as the best source of evidence for estimating the unknown aspects of the new case.
- The other cases show alternative possibilities for those unknown aspects; the closer the agreement among the alternatives, the stronger the evidence for the conclusion

Retrieve:

- Given a target problem, retrieve cases from memory that are relevant to solving it.
- A case consists of a problem, its solution, and, typically, annotations about how the solution was derived. For example, suppose Fred wants to prepare blueberry pancakes.
- The procedure he followed for making the plain pancakes, together with justifications for decisions made along the way, constitutes Fred's retrieved case.

2. Reuse:

- Map the solution from the previous case to the target problem.
- This may involve adapting the solution as needed to fit the new situation.
- In the pancake example, Fred must adapt his retrieved solution to include the addition of blueberries.

3. Revise:

- Having mapped the previous solution to the target situation, test the new solution in the real world (or a simulation) and, if necessary, revise.
- Suppose Fred adapted his pancake solution by adding blueberries to the batter.
- After mixing, he discovers that the batter has turned blue – an undesired effect.
- This suggests the following revision: delay the addition of blueberries until after the batter has been ladled into the pan.

4. Retain:

- After the solution has been successfully adapted to the target problem, store the resulting experience as a new case in memory.
- Fred, accordingly, records his newfound procedure for making blueberry pancakes, thereby enriching his set of stored experiences, and better preparing him for future pancake-making demands.

Transformational Analogy:

- Suppose you are asked to prove a theorem in plane geometry.
- You might look for a previous theorem that is very similar and copy its proof, making substitutions when necessary.
- The idea is to transform a solution to a previous problem in to solution for the current problem. The following figure shows this process,

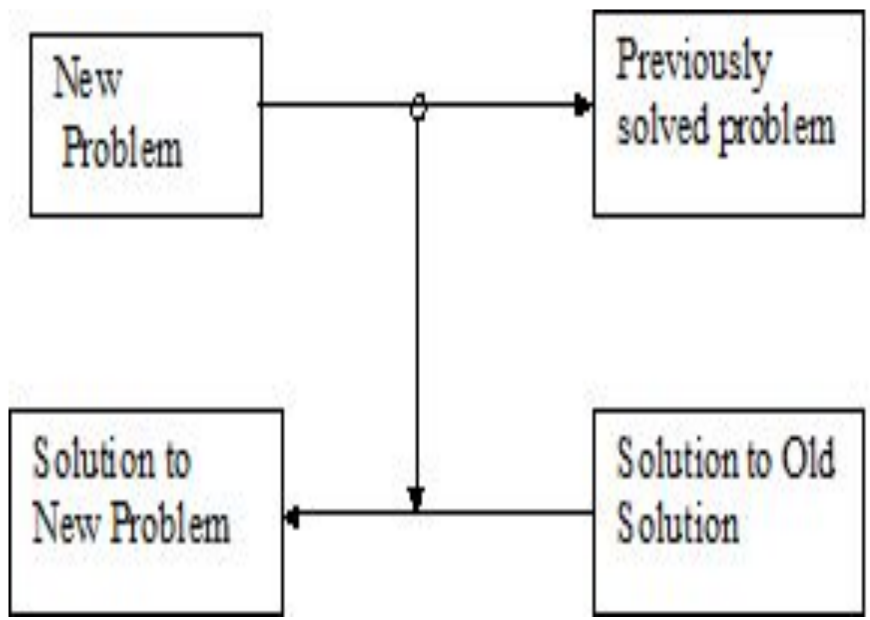


Fig: Transformational Analogy

Derivational Analogy:

- Notice that transformational analogy does not look at how the old problem was solved, it only looks at the final solution.
- Often the twists and turns involved in solving an old problem are relevant to solving a new problem.
- The detailed history of problem solving episode is called derivation, Analogical reasoning that takes these histories into account is called derivational analogy.

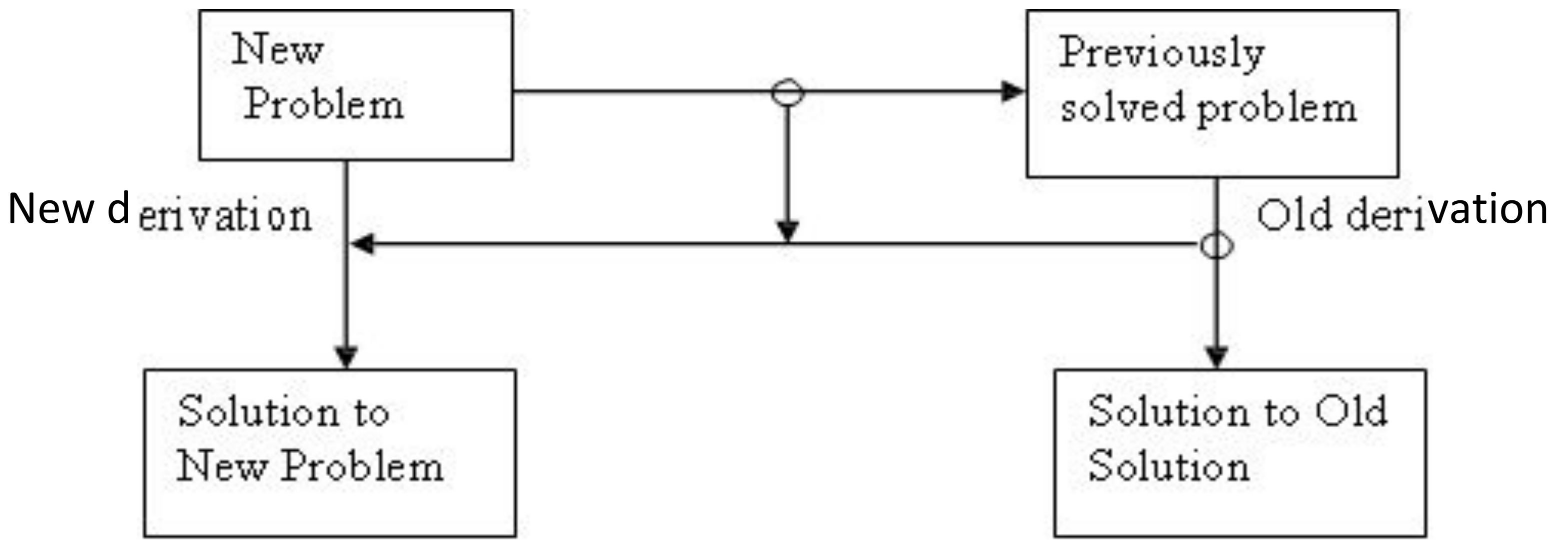
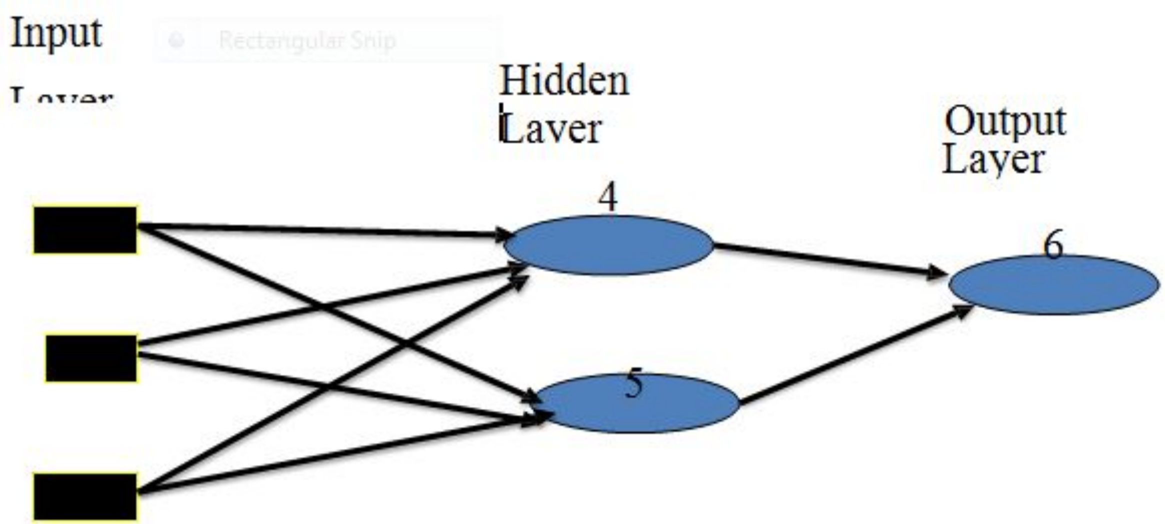


Fig: Derivational Analogy

Artificial Neural Networks

- A neural network is composed of number of nodes or units , connected by links. Each link has a numeric weight associated with it.



- Artificial neural networks are programs design to solve any problem by trying to mimic the structure and the function of our nervous system.

Artificial neural network model:

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

$$sum = w_1 x_1 + w_2 x_2 + \ldots\ldots + w_n x_n$$

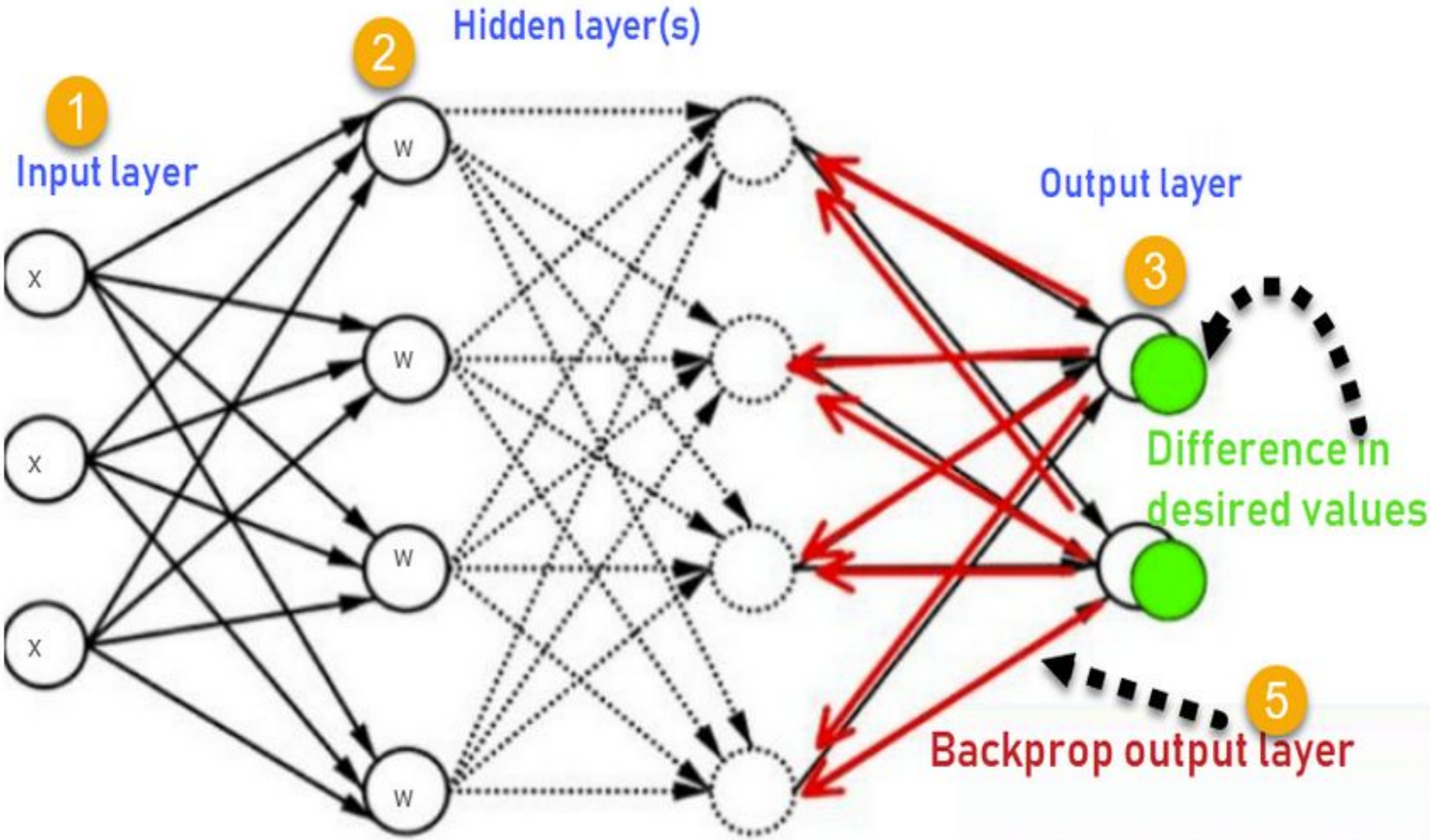
- These products are simply summed, fed through the transfer function f() to generate result and output

Back Propagation Algorithm

- **Back propagation** is the essence of neural network training.
- It is the method of fine-tuning the weights of a neural network based on the error rate obtained in the previous iteration.
- Proper tuning of the weights allows to reduce error rates and make the model reliable by increasing its generalization.
- Backpropagation in neural network is a short form for "backward propagation of errors."
- It is a standard method of training artificial neural networks.
- This method helps calculate the gradient of a loss function with respect to all the weights in the network.

How Back Propagation Algorithm Works?

- The Back propagation algorithm in neural network computes the gradient of the loss function for a single weight by the chain rule.
- It efficiently computes one layer at a time, unlike a native direct computation.
- It computes the gradient, but it does not define how the gradient is used.
- It generalizes the computation in the delta rule.
- Consider the following Back propagation neural network example diagram to understand:



How Backpropagation Algorithm Works

- Inputs X, arrive through the pre connected path
- Input is modeled using real weights W. The weights are usually randomly selected.
- Calculate the output for every neuron from the input layer, to the hidden layers, to the output layer.
- Calculate the error in the outputs

  $Error_B$= Actual Output – Desired Output
- Travel back from the output layer to the hidden layer to adjust the weights such that the error is decreased.
- Keep repeating the process until the desired output is achieved

Why We Need Backpropagation?

Most prominent advantages of Backpropagation are:

- Back propagation is fast, simple and easy to program
- It has no parameters to tune apart from the numbers of input
- It is a flexible method as it does not require prior knowledge about the network
- It is a standard method that generally works well
- It does not need any special mention of the features of the function to be learned.

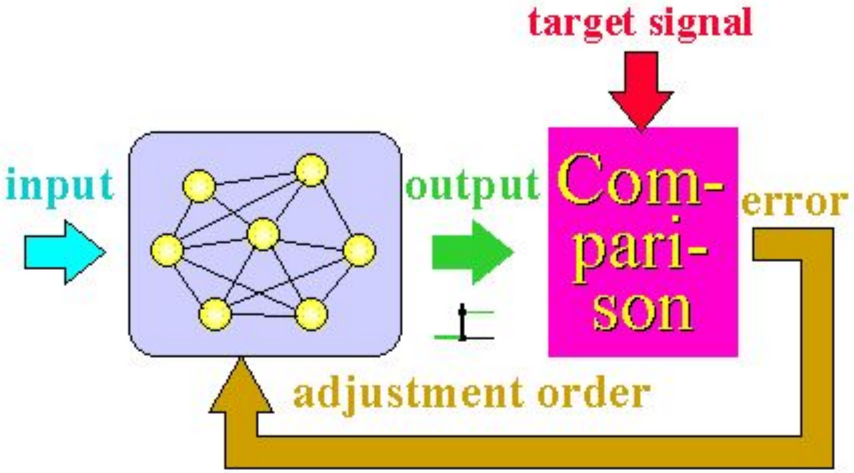
Learning in Neural Networks:

- Learning:
  - Learning in neural networks is carried out by adjusting the connection weights among neurons.
  - There is no algorithm that determines how the weights should be assigned in order to solve specific problems. Hence, the weights are determined by a learning process
- Learning may be classified into two categories:
  - Supervised Learning
  - Unsupervised Learning

Learning in Neural Networks:

1) Supervised Learning:

- In supervised learning, the network is presented with inputs together with the target (teacher signal) outputs.
- Then, the neural network tries to produce an output as close as possible to the target output by adjusting the values of internal weights.
- The most common supervised learning method is the "**error correction method**".
  - Neural networks are trained with this method in order to reduce the error (difference between the network's output and the desired output) to zero.

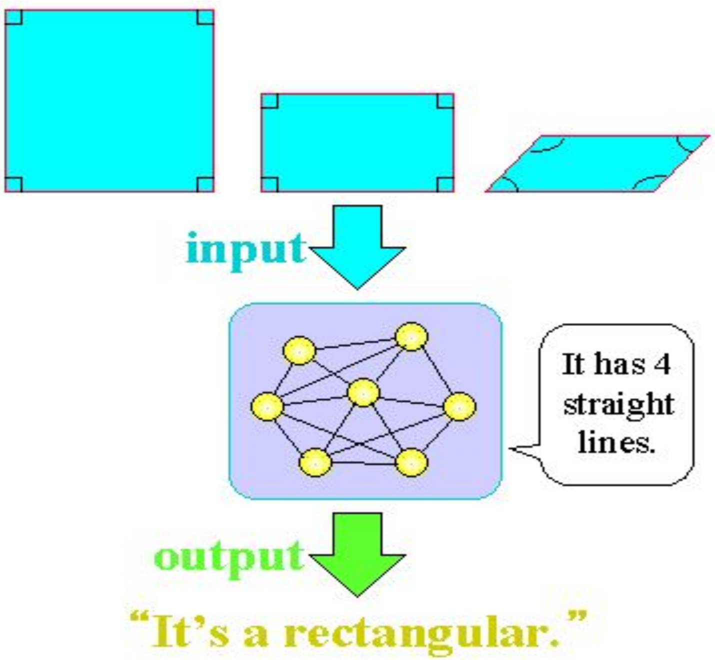


Learning in Neural Networks:

2) Unsupervised Learning:

– In unsupervised learning, there is no teacher (target signal/output) from outside and the network adjusts its weights in response to only the input patterns

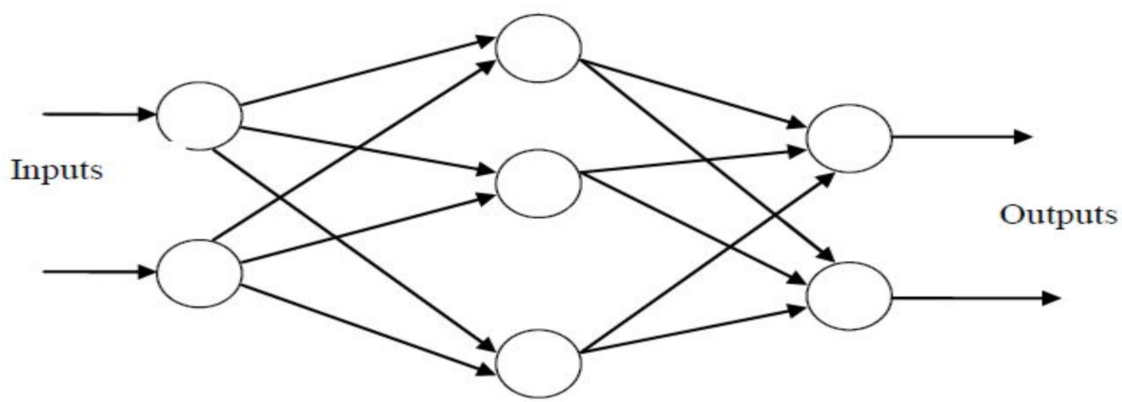
– A typical example of unsupervised learning is **Hebbian learning.**



Network Architecture

Feed-forward networks:

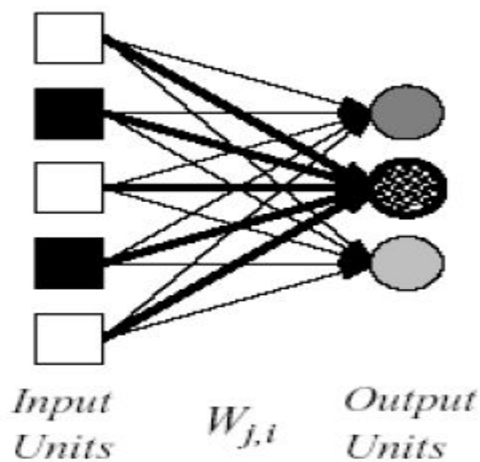
- Feed-forward ANNs allow signals to travel one way only; from input to output.
- Feed-forward ANNs tend to be straight forward networks that associate inputs with outputs.



Types of Feed Forward Neural Network:

a) Single-layer neural networks

A neural network in which all the inputs connected directly to the outputs is called a single-layer neural network.



Types of Feed Forward Neural Network:

a) Single-layer Feed Forward neural networks :

Two types:

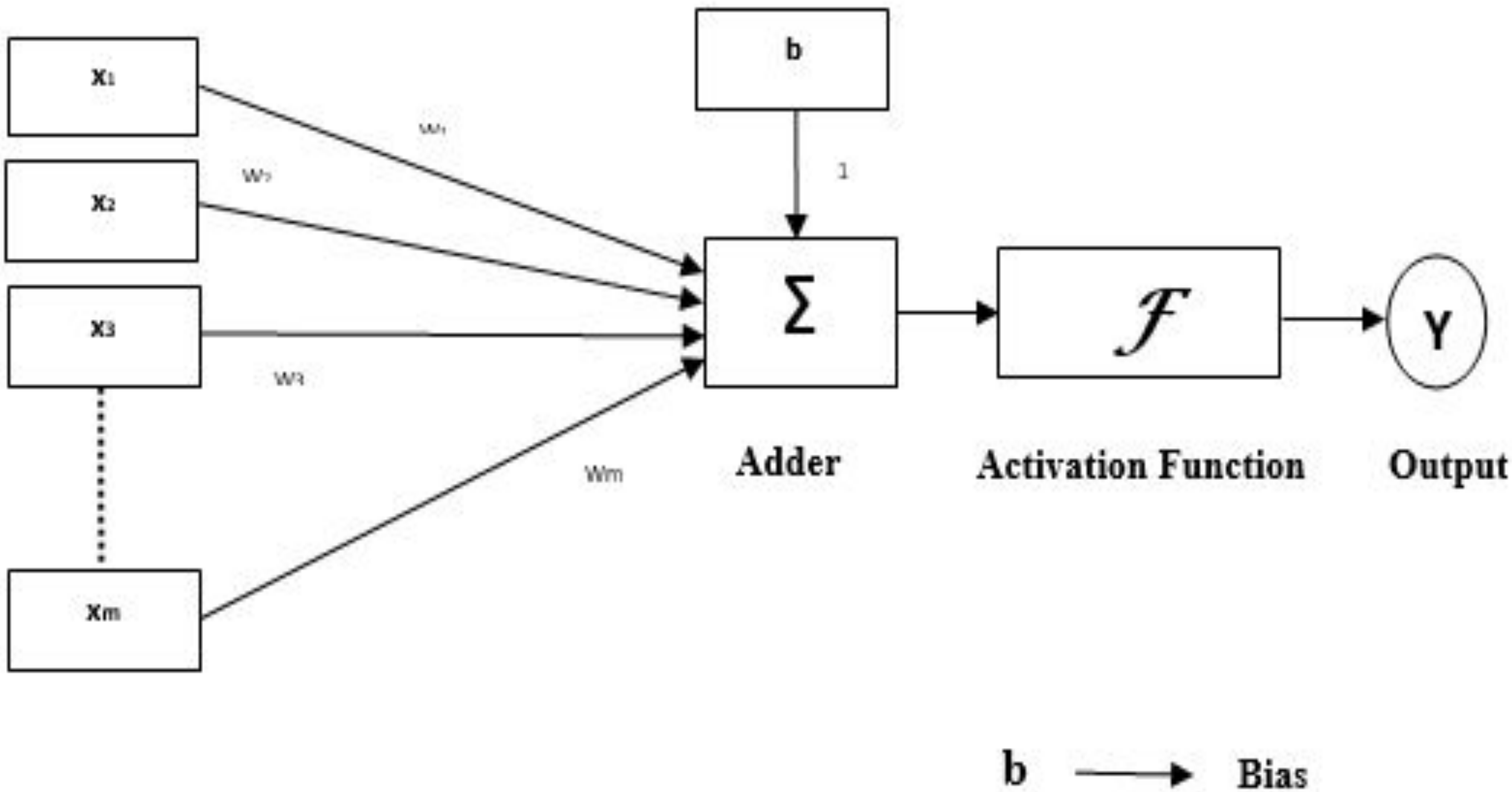
- Perceptron, and
- ADLINE

Perceptron

- Developed by Frank Rosenblatt by using McCulloch and Pitts model, perceptron is the basic operational unit of artificial neural networks.
- It employs supervised learning rule and is able to classify the data into two classes.
- Operational characteristics of the perceptron:
  - It consists of a single neuron with an arbitrary number of inputs along with adjustable weights, but the output of the neuron is 1 or -1 depending upon the input. It also consists of a bias whose weight is always 1.
- Following figure gives a schematic representation of the perceptron.

Perceptron

- Following figure gives a schematic representation of the perceptron:



Perceptron

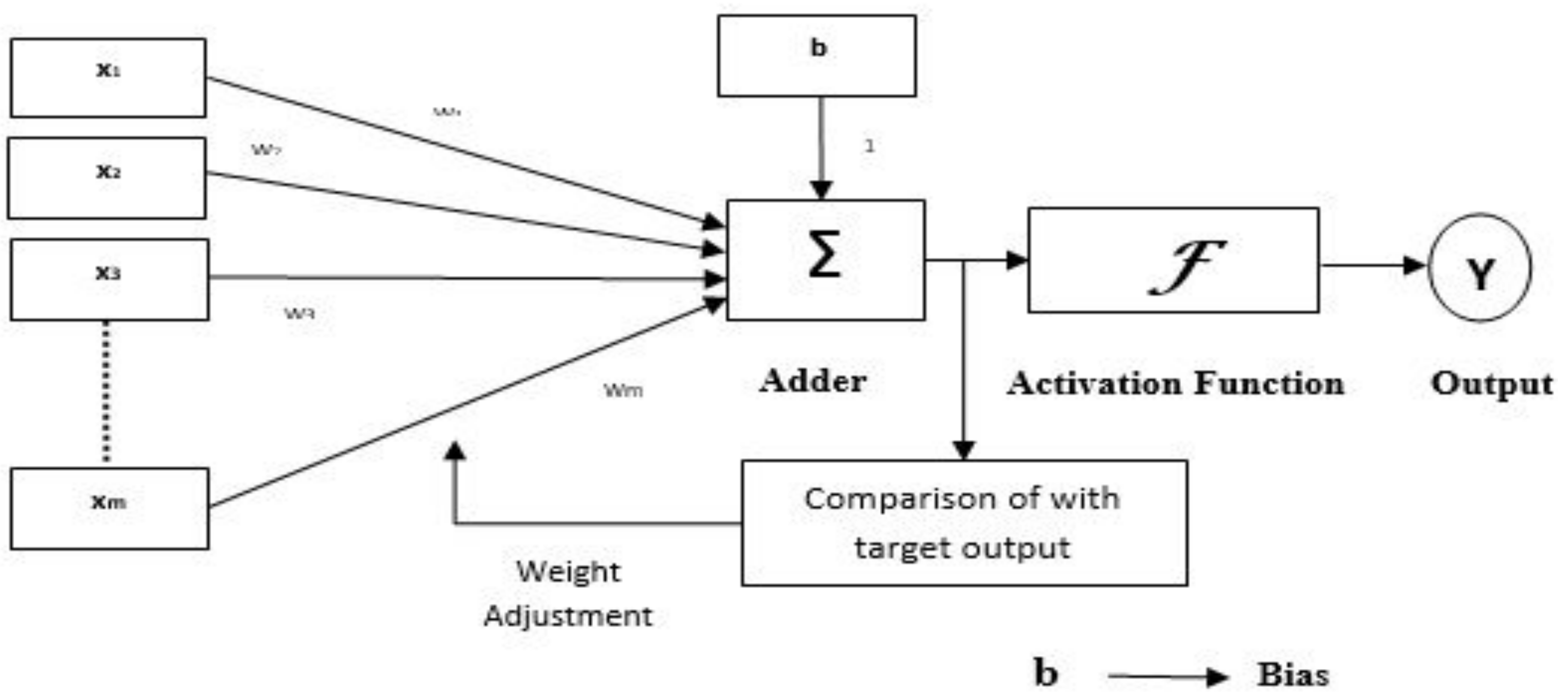
- Perceptron thus has the following three basic elements:
  - Links
  - Adder
  - Activation function

Adaptive Linear Neuron (ADALINE)

- ADALINE which stands for Adaptive Linear Neuron, is a network having a single linear unit.
- It was developed by Widrow and Hoff in 1960. Some important points about ADALINE are as follows:
  - It uses bipolar activation function.
  - It uses delta rule for training to minimize the Mean-Squared Error (MSE) between the actual output and the desired/target output.
  - The weights and the bias are adjustable.

Adaptive Linear Neuron (ADALINE)

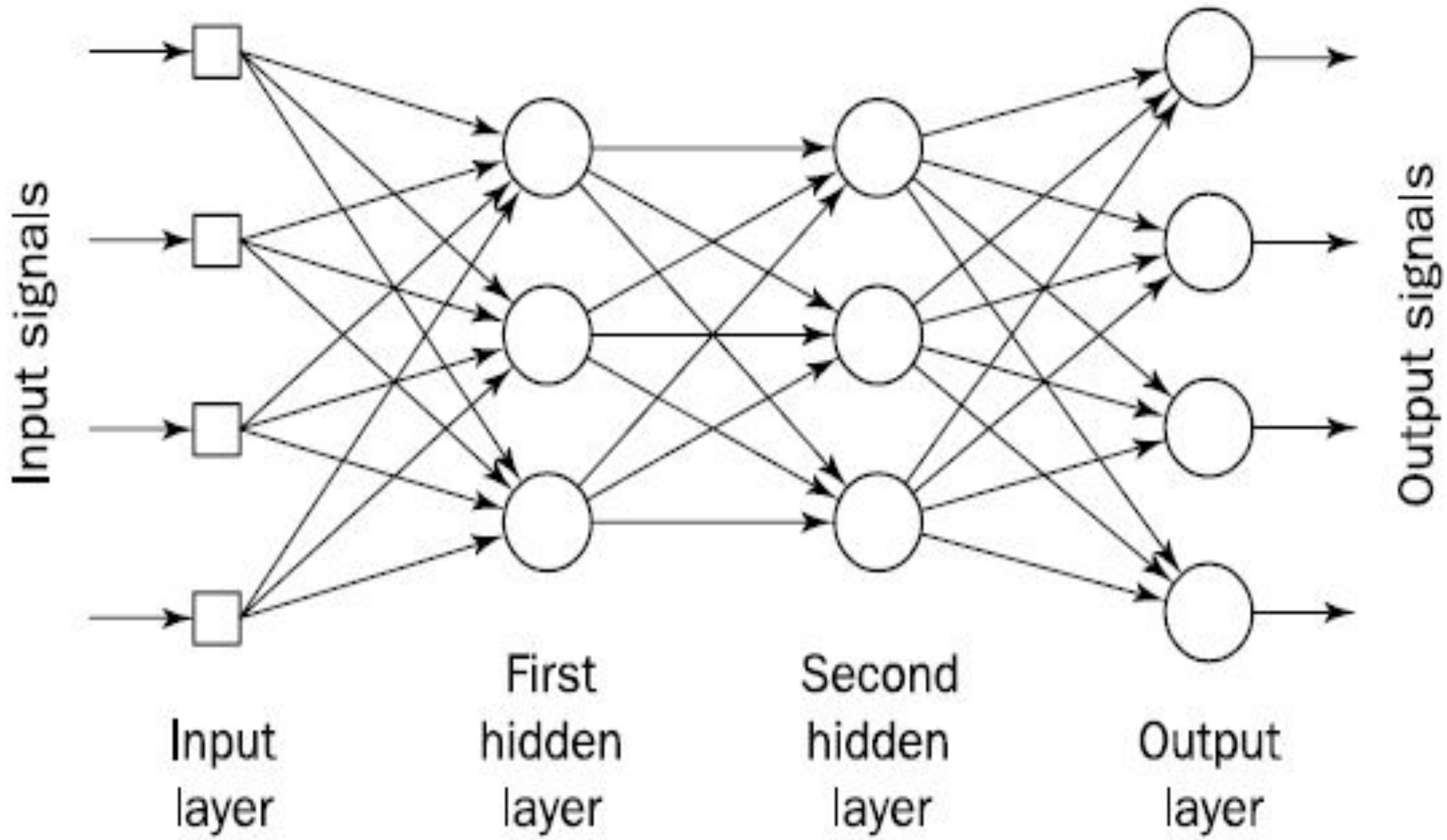
- **Architecture :**
  - The basic structure of ADALINE is similar to perceptron having an extra feedback loop with the help of which the calculated output is compared with the desired/target output.
  - After comparison on the basis of training algorithm, the weights and bias will be updated.



Types of Feed Forward Neural Network:

b) Multilayer neural networks

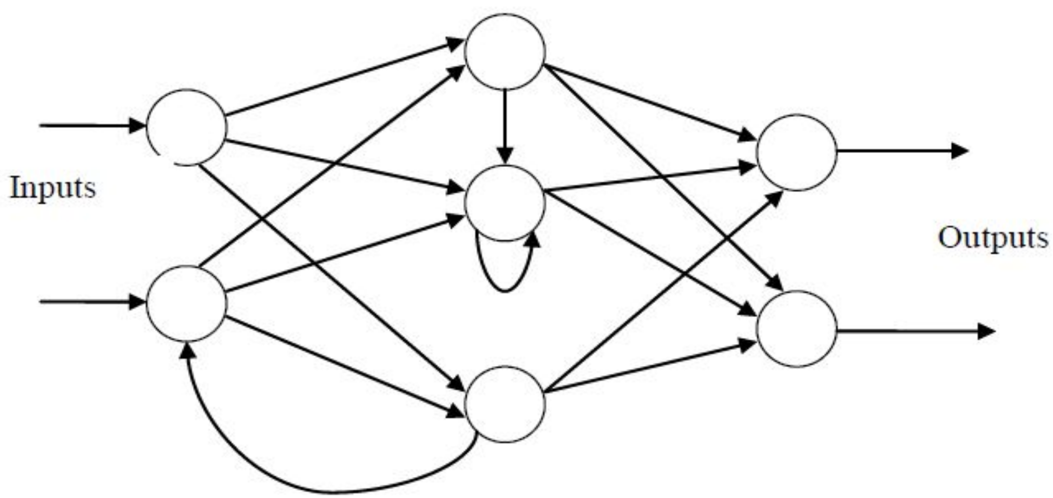
The neural network which contains input layers, output layers and some hidden layers also is called multilayer neural network.



Network Architectures

2) Feedback networks (Recurrent networks:)

- Feedback networks can have signals traveling in both directions by introducing loops in the network.
  - very powerful
  - extremely complicated.
  - dynamic: Their 'state' is changing continuously until they reach an equilibrium point.
- also known as interactive or recurrent.



Applications of Neural Network

- Speech recognition
- Optical character recognition
- Face Recognition
- Pronunciation (NETtalk)
- Stock-market prediction
- Navigation of a car
- Signal processing/Communication
- Imaging/Vision
- ….