DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
(ARTIFICIAL INTELLIGENCE & MACHINE LEARNING)

Delta Learning Rule (LMS-Widrow Hoff)

- A network with single linear unit is called Adaptive Linear Neuron(Adaline).
- In Adaline, input-output relationship is linear.
- It uses bipolar activation function.
- Adaline neuron can be trained using Delta rule or Least Mean Square(LMS) rule or widrow-hoff rule.

- The delta rule is derived from gradient descent method.
- The delta rule updates the weights between the connections so as to minimize the difference between the net input to the output unit and the target value.
- The aim is to minimize the error over all training patterns.
- This is done by reducing the error for each pattern, one at a time.
- The delta rule for adjusting the weight of ith pattern

$$\Delta w_i = \alpha (t - y_{in}) x_i$$

α = learning rate

x = input

$y_{in} = \sum_{i=1}^{n} x_i w_i$

Architecture:

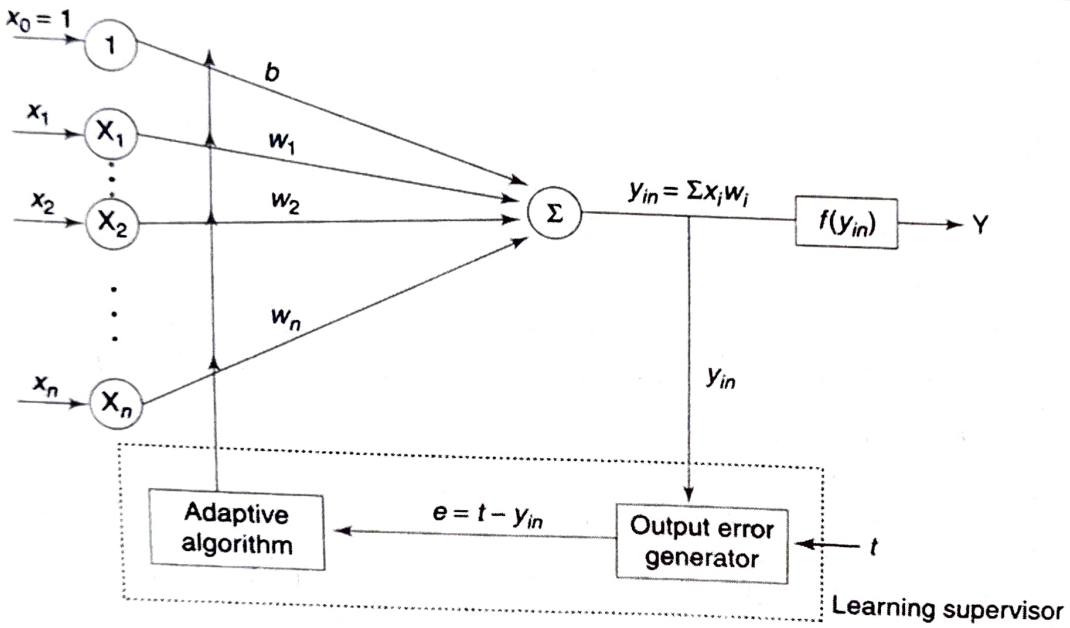

Fig. Adaline model

- The adaline model is shown in above figure.
- Weights have signs.
- Inputs are either of 2 values (+1 or -1).
- Initially random weights are assigned.
- The net input calculated is applied to a activation function that restores the output to +1 or -1.

Training algorithm:

The adaline training algorithm is as follows:

Step 0: Weights and bias are set to some random value (not zero).
Set learning rate parameter α.

Step 1: Perform step 2-6 when stopping condition is false.

Step 2: Perform step 3-5 for each bipor training pair.

Step 3: Set activations for input units i=1 to n

$$x_i = s_i$$

Step 4: calculate net input to output unit.

$$y_{in} = b + \sum_{i=1}^{n} x_i w_i$$

Step 5: Update the weights and bias for i= 1 to n.

$$w_i(new) = w_i(old) + \alpha(t - y_{in})x_i$$

$$b(new) = b(old) + \alpha(t - y_{in})$$

Step 6: If error is smaller than specified, stop the training process, else continue.

Error = $\Sigma (t - y_{in})^2$