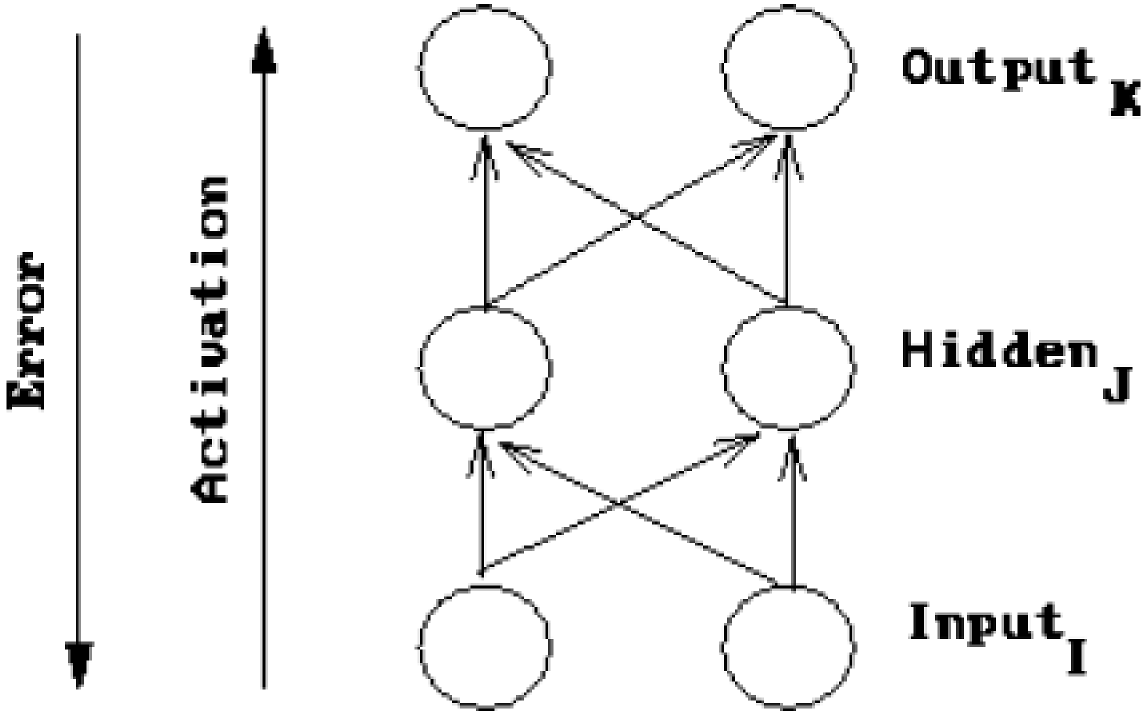
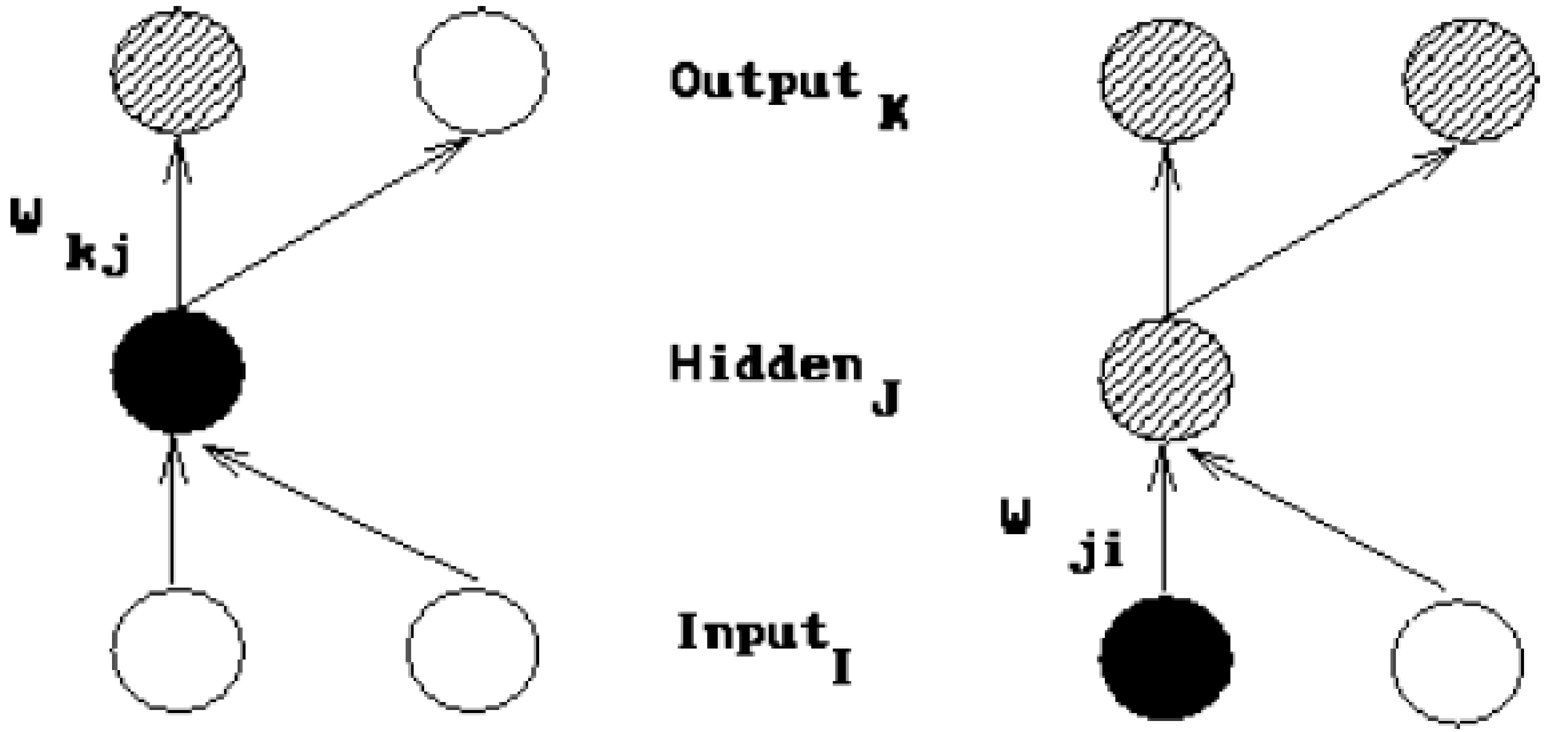
Back Propagation Algorithm

Prof. G. Panda
IIT Bhubaneswar



- The subscript k denotes the output layer.
- The subscript j denotes the hidden layer.
- The subscript i denotes the input layer.



Output K
Hidden J
Input I
W kj
W ji

Notations:

- w_{kj} denotes a weight from the hidden to the output layer.
- w_{ji} denotes a weight from the input to the hidden layer.
- a denotes an activation value.
- t denotes a target value.
- net denotes the net input.

Review of calculus rules

$$\frac{d(e^u)}{dx} = e^u \frac{du}{dx} \qquad \frac{d(g+h)}{dx} = \frac{dg}{dx} + \frac{dh}{dx} \qquad \frac{d(g^n)}{dx} = ng^{n-1}\frac{dg}{dx}$$

Gradient descent on error

$$E = \frac{1}{2}\sum_k (t_k - a_k)^2$$

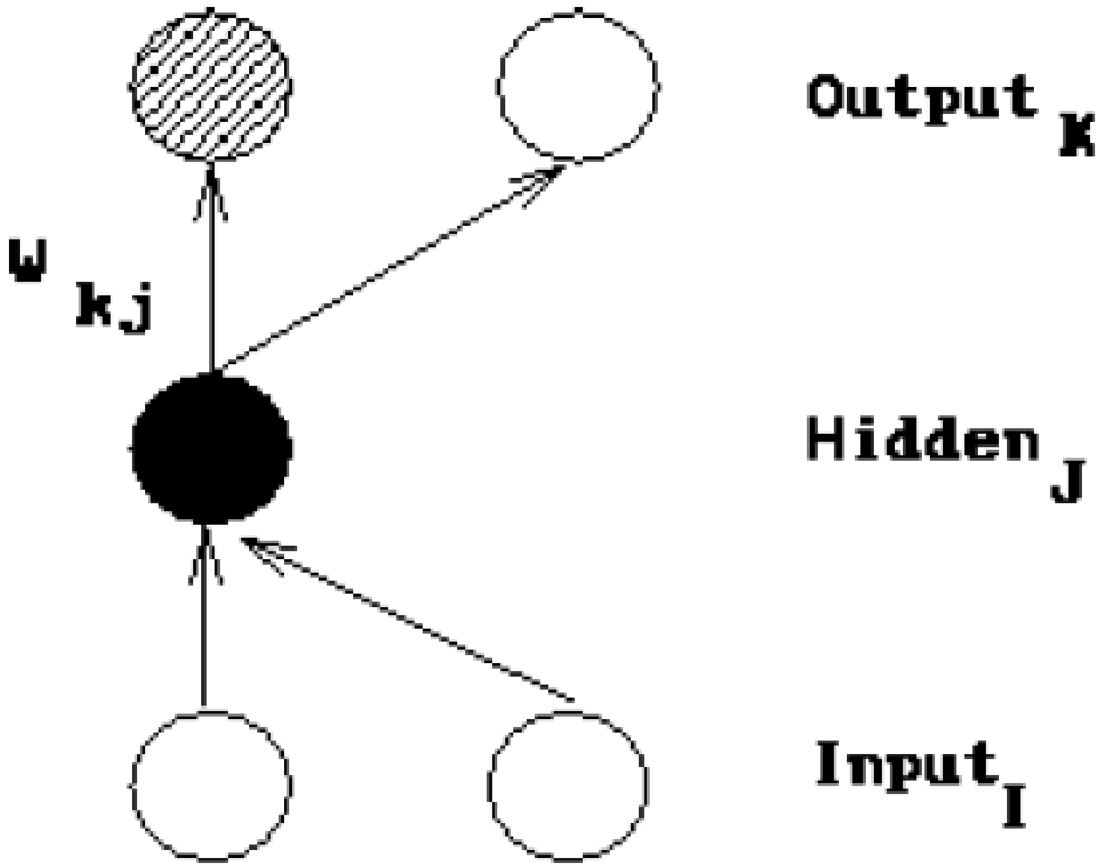
We want to adjust the network's weights to reduce this overall error

$$\Delta W \propto -\frac{\partial E}{\partial W}$$

Start at output layer:

$$\Delta w_{kj} \propto -\frac{\partial E}{\partial w_{kj}}$$

$$\Delta w_{kj} = -\varepsilon \frac{\partial E}{\partial a_k} \frac{\partial a_k}{\partial net_k} \frac{\partial net_k}{\partial w_{kj}}$$



Derivative of error with respect to activation

$$\frac{\partial E}{\partial a_k} = \frac{\partial(\frac{1}{2}(t_k - a_k)^2)}{\partial a_k} = -(t_k - a_k)$$

Derivative of activation with respect to net input

$$\frac{\partial a_k}{\partial net_k} = \frac{\partial(1 + e^{-net_k})^{-1}}{\partial net_k} = \frac{e^{-net_k}}{(1 + e^{-net_k})^2}$$

Using $1 - \frac{1}{1 + e^{-net_k}} = \frac{e^{-net_k}}{1 + e^{-net_k}}$ we can write $a_k(1 - a_k)$

Derivative of net input with respect to weight

$$\frac{\partial net_k}{\partial w_{kj}} = \frac{\partial (w_{kj} a_j)}{\partial w_{kj}} = a_j$$

Weight change rule for output to hidden weight

$$\Delta w_{kj} = \varepsilon \overbrace{(t_k - a_k) a_k (1 - a_k)}^{\delta_k} a_j$$

$$\Delta w_{kj} = \varepsilon \delta_k a_j$$

Weight change rule for hidden to input weight

$$\Delta w_{ji} \propto -[\sum_k \frac{\partial E}{\partial a_k} \frac{\partial a_k}{\partial net_k} \frac{\partial net_k}{\partial a_j}] \frac{\partial a_j}{\partial net_j} \frac{\partial net_j}{\partial w_{ji}}$$

$$= \varepsilon [\sum_k \overbrace{(t_k - a_k) a_k (1 - a_k)}^{\delta_k} w_{kj}] a_j (1 - a_j) a_i$$

$$= \varepsilon \overbrace{[\sum_k \delta_k w_{kj}] a_j (1 - a_j)}^{\delta_j} a_i$$

$$\Delta w_{ji} = \varepsilon \delta_j a_i$$

Final equations:

Weight change rule for output to hidden weight

$$\delta_k = (t_k - a_k)a_k(1 - a_k)$$

$$w_{kj}(n+1) = w_{kj}(n) + \varepsilon\delta_k\alpha_j$$

Weight change rule for hidden to input weight

$$\delta_j = (\sum_k \delta_k\, w_{kj}(n)\,)a_j(1 - a_j)$$

$$w_{ji}(n+1) = w_{ji}(n) + \varepsilon\delta_j\alpha_i$$

Thank you