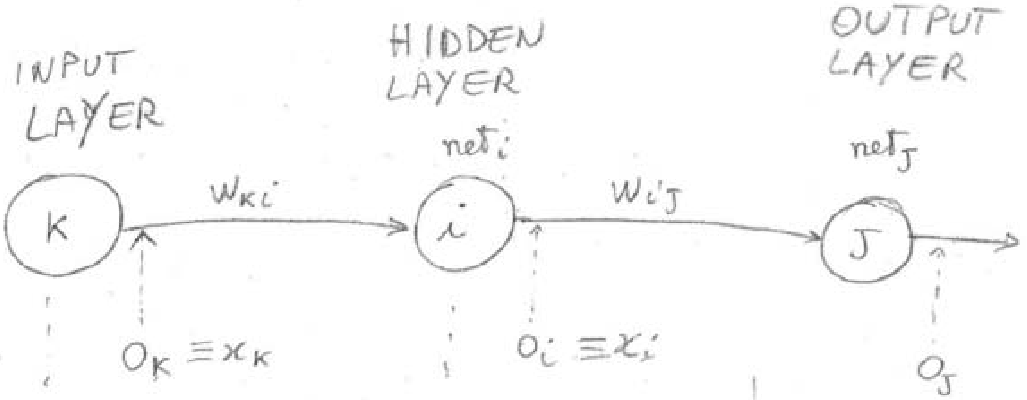
E for a training sample: $\frac{1}{2} \sum_J (t_J - O_J)^2$



$$\frac{\partial E}{\partial w_{iJ}} = \frac{\partial E}{\partial O_J} \frac{\partial O_J}{\partial net_J} \frac{\partial net_J}{\partial w_{iJ}} =$$

O_i

$$= \underbrace{-(t_J - O_J)\, f'(net_J)}_{\delta_J}\, x_i = -\delta_J x_i$$

$$net_J = \sum_{i=0}^{m} x_i w_{iJ} = \sum_{i=0}^{m} O_i w_{iJ}$$

$$O_J = f(net_J)$$

$$\frac{\partial E}{\partial w_{Ki}} = \frac{\partial E}{\partial O_J} \frac{\partial O_J}{\partial net_J} \frac{\partial net_J}{\partial O_i} \frac{\partial O_i}{\partial net_i} \frac{\partial net_i}{\partial w_{Ki}} =$$

O_K

$$= \underbrace{-\delta_J\, w_{iJ}\, f'(net_i)}_{\delta_i}\, x_K = -\delta_i x_K$$

If, e.g., neuron i is connected to more output neurons, we have:

$$\delta_i = f'(net_i) \sum_J \delta_J w_{iJ}$$