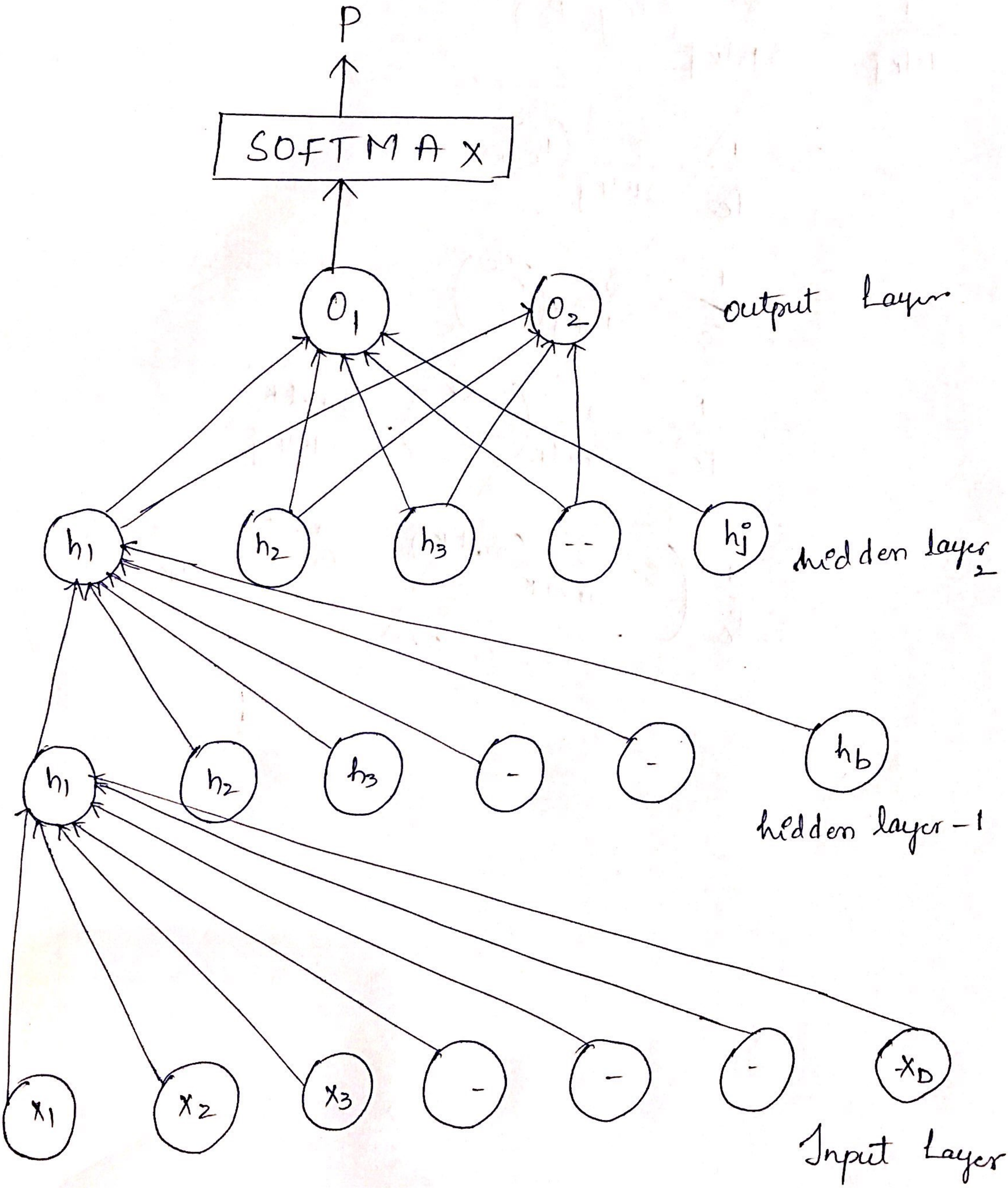
## Multi-class classification using 4 layer NN :-





(7)

$$\frac{\partial J}{\partial W_{kj}} = \frac{\partial}{\partial W_{kj}} (\log P_c)$$

$$= \frac{1}{P_c} \cdot \frac{\partial}{\partial W_{kj}} (P_c)$$

$$= \frac{1}{P_c} \cdot \frac{\partial}{\partial W_{kj}} \left( \frac{e^{O_c}}{S} \right)$$

$$= \frac{1}{P_c} \cdot \frac{\partial}{\partial net_k} \left( \frac{e^{net_c}}{S} \right) \cdot \frac{\partial net_c}{\partial W_{kj}}$$

$$= \frac{1}{P_c} \left( \frac{S \cdot \frac{\partial}{\partial net_k}(net_c) - e^{net_c} \cdot \frac{\partial}{\partial net_k}(S)}{S^2} \right)$$

$$= \frac{1}{P_c} \left( \frac{S \cdot \frac{\partial}{\partial net_k} e^{net_c} - e^{net_c} \cdot \frac{\partial}{\partial net_k} S}{S^2} \right)$$

When $k == c$

$$= \frac{1}{P_c} \left( \frac{S \cdot e^{O_c} - e^{O_c} \cdot e^{O_k}}{S^2} \right) \cdot \frac{\partial}{\partial W_{kj}} (\Sigma W_{kj} h_j)$$

$$= \frac{1}{P_c} \left( \frac{S \cdot e^{O_c} - e^{O_c} e^{O_k}}{S^2} \right) \cdot h_j$$

$$= \frac{1}{\cancel{P_c}} \cdot \left( \frac{\cancel{e^{o_c}}}{\cancel{s}^2} \right) (s - e^{o_K}) \cdot h_j^o$$ ⑧

$$= \left( 1 - \frac{e^{o_K}}{s} \right) h_j^o$$

$$= (1 - P_K)\, h_j^o$$

When $c \neq K$

$$= \frac{-e^{o_c} \cdot e^{o_K} \cdot h_j^o}{s^2}$$

$$= -P_K \cdot h_j^o$$

By combining both of them

$$\frac{\partial J}{\partial W_{Kj}} = (Y_K - P_K) \cdot h_j$$

$$\boxed{W_{Kj}^{New} = W_{Kj}^{old} + \lambda (Y_K - P_K) \cdot h_j^o}$$

(2)

$$\frac{\partial J}{\partial W_{jb}^{i}} = \sum_K \log P_c$$

$$= \sum_K \frac{1}{P_c} \cdot \frac{\partial}{\partial W_{jb}^{i}} (P_c)$$

$$= \sum_K \left(\frac{1}{P_c}\right) \cdot \frac{\partial F}{\partial net_K} \cdot \frac{\partial net_K}{\partial W_{jb}}$$

$$= \sum_K (Y_K - P_K) \cdot \frac{\partial}{\partial W_{jb}^{i}} (W_{Kj} \cdot h_j^{i})$$

$$= \sum_K (Y_K - P_K) \cdot W_{Kj}^{i} \cdot \frac{\partial}{\partial W_{jb}^{i}} h_j$$

$$= \sum_K (Y_K - P_K) \cdot W_{Kj}^{i} \cdot \frac{\partial F_2}{\partial net_j^{i}} \cdot \frac{\partial net_j^{i}}{\partial W_{jb}^{i}}$$

$$= \sum_K (Y_K - P_K) \cdot W_{Kj}^{i} \cdot F_2'(net_j^{i}) \cdot \frac{\partial}{\partial W_{jb}^{i}} \left(\sum W_{jb}^{i} \cdot h_b\right)$$

$$= \sum_K (Y_K - P_K) \cdot W_{Kj}^{i} \cdot F_2'(net_j^{i}) \cdot h_b$$

$$\boxed{W_{jb}^{i\,New} = W_{jb}^{i\,old} + \lambda \sum_K (Y_K - P_K) \cdot W_{Kj}^{i} \cdot F_2'(net_j^{i}) \cdot h_b}$$

$$\frac{\partial J}{\partial W_{bi}} = \sum_{k,j} \log P_c$$

$$= \sum_{k,j} \frac{1}{P_c} \cdot \frac{\partial}{\partial W_{bi}} (P_c)$$

$$= \sum_{k,j} \frac{1}{P_c} \cdot \frac{\partial}{\partial W_{bi}} \left( \frac{e^{O_c}}{S} \right)$$

$$= \sum_{k,j} \frac{1}{P_c} \cdot \frac{\partial}{\partial net_k} \left( \frac{e^{O_c}}{S} \right) \cdot \frac{\partial net_k}{\partial W_{bi}}$$

$$= \sum_{k,j} (Y_k - P_k) \cdot W_{kj} \cdot \frac{\partial h_j}{\partial W_{bi}}$$

$$= \sum_{k,j} (Y_k - P_k) \cdot W_{kj} \cdot \frac{\partial}{\partial W_{bi}} f_2(net_j)$$

$$= \sum_{k,j} (Y_k - P_k) \cdot W_{kj} \cdot \frac{\partial}{\partial net_j} f_2(net_j) \cdot \frac{\partial}{\partial W_{bi}} (net_j)$$

$$= \sum_{k,j} (Y_k - P_k) \cdot W_{kj} \cdot f_2'(net_j) \cdot W_{jb} \cdot \frac{\partial}{\partial W_{bi}} (h_b)$$

$$= \sum_{k,j} (Y_k - P_k) \cdot W_{kj} \cdot f_2'(net_j) \cdot W_{jb} \cdot f_1'(net_b) \cdot X_i$$

$$\boxed{W_{bi}^{new} = W_{bi}^{old} + \lambda \sum_{k,j} (Y_k - P_k) \cdot W_{kj} \cdot f_2'(net_j) \cdot W_{jb} \cdot f_1'(net_b) \cdot X_i}$$