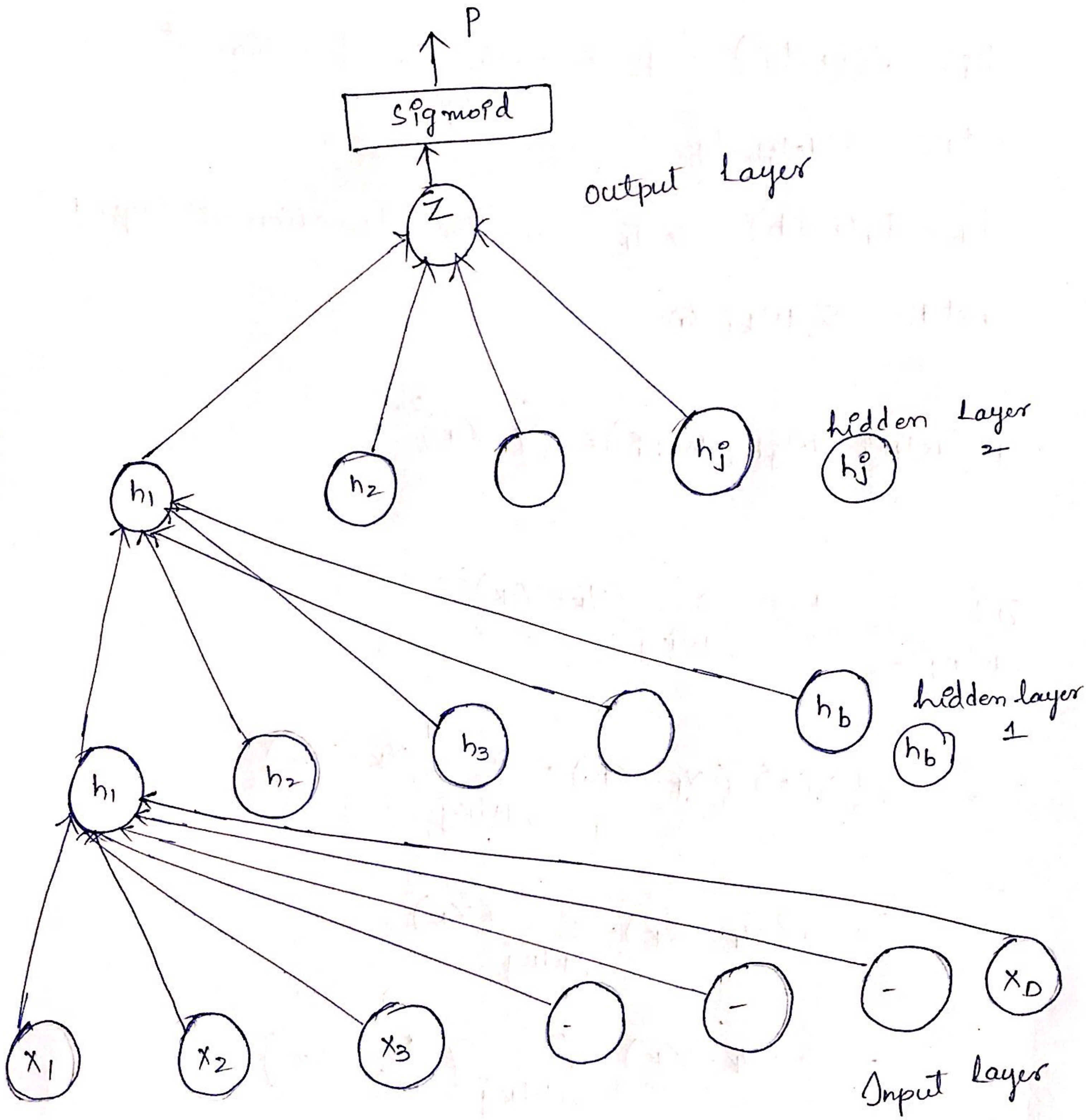
4- Layer NN for Binary Classification Using Squared Loss :-



- All the nodes are connected

Scanned with CamScanner

$Z_K = P = f(\text{net } z) \longrightarrow$ Function f is Sigmoid

$\text{net } z = \sum W_{kj} \cdot h_j$

$h_j = f_2(\text{net}_j)$ $\quad f_2 \rightarrow$ Activation function at layer 2.

$\text{net}_j = \sum W_{jb} \cdot h_b$

$h_b = f_1(\text{net } b) \longrightarrow f_1$ activation function at layer 1

$\text{net } b = \sum W_{bi} \cdot x_i$

$$J(W_{kj}, W_{jb}, W_{bi}) = (y - Z_K)^2$$

$$\frac{\partial J}{\partial W_{kj}} = \frac{1}{2} \cdot \frac{\partial}{\partial W_{kj}} (Y_K - Z_K)^2$$

$$= \left(\frac{1}{\cancel{2}}\right)(\cancel{2})(Y_K - Z_K) \cdot \frac{\partial}{\partial W_{kj}} (Y_K - Z_K)$$

$$= -(Y_K - Z_K) \cdot \frac{\partial}{\partial W_{kj}} (Z_K)$$

$$= -(Y_K - Z_K) \cdot \frac{\partial}{\partial W_{kj}} \left(\frac{1}{1 + e^{-\text{net } z}}\right)$$

$$= -(Y_K - Z_K) \cdot \frac{\partial}{\partial W_{kj}} \left(\frac{1}{1 + e^{-\text{net } z}}\right)$$

Scanned with CamScanner

$$= -(y_k - z_k)\cdot \frac{e\;1}{(1+e^{-netz})^2}\cdot \frac{\partial}{\partial w_{kj}}(1+e^{-netz})$$

$$= -(y_k - z_k)\cdot \frac{1}{(1+e^{-netz})^2}\cdot \frac{\partial}{\partial netz} e^{-netz} \cdot \frac{\partial\, netz}{\partial w_{kj}}$$

$$= -(y_k - z_k)\cdot \frac{e^{-netz}}{(1+e^{-netz})^2}\cdot \frac{\partial}{\partial w_{kj}}\left(\Sigma w_{kj}\cdot h_j\right)$$

$$= -(y_k - z_k)\cdot \frac{e^{-netz}}{(1+e^{-netz})^2}\cdot h_j$$

$$\boxed{w_{kj}^{new} = w_{kj}^{old} + \lambda\,(y_k - z_k)\cdot \frac{e^{-netz}}{(1+e^{-netz})^2}\, h_j}$$

$$\frac{\partial J}{\partial w_{jb}} = \sum_k \frac{1}{2}\cdot \frac{\partial J}{\partial w_{jb}}(y_k - z_k)^2$$

In this Case k=1 So the Sum doesnot matter

$$= \cancel{\left(\frac{1}{2}\right)}\cancel{(2)}\,(y_k - z_k)\cdot \frac{\partial}{\partial w_{jb}}(y_k - z_k)$$

$$= -(y_k - z_k)\cdot \frac{\partial}{\partial w_{jb}}(z_k)$$

$$= -(y_k - z_k)\cdot \frac{\partial}{\partial w_{jb}}\left(\frac{1}{1+e^{-netz}}\right)$$

Scanned with CamScanner

$$= -(Y_K - Z_K) \cdot \frac{\partial F}{\partial net z} \cdot \frac{\partial net z}{\partial w_{jb}}$$

$$= -(Y_K - Z_K) \cdot \frac{e^{-net z}}{(1+e^{-net z})^2} \cdot \frac{\partial}{\partial w_{jb}} \left(\Sigma \, w_{kj} \cdot h_j \right)$$

$$= -(Y_K - Z_K) \cdot \frac{e^{-net z}}{(1+e^{-net z})^2} \cdot w_{kj} \cdot \frac{\partial}{\partial w_{jb}} F_2(net j)$$

$$= -(Y_K - Z_K) \cdot \frac{e^{-net z}}{(1+e^{-net z})^2} \cdot w_{kj} \cdot \frac{\partial}{\partial w_{jb}} F_2(net j)$$

$$= -(Y_K - Z_K) \cdot \frac{e^{-net z}}{(1+e^{-net z})^2} \cdot w_{kj} \cdot F_2'(net j) \cdot \frac{\partial \, \Sigma w_{jb} \cdot h_b}{\partial w_{jb}}$$

$$= -(Y_K - Z_K) \cdot \frac{e^{-net z}}{(1+e^{-net z})^2} \cdot w_{kj} \cdot F_2'(net j) \cdot h_b$$

$$\boxed{w_{jb}^{New} = w_{jb}^{old} + \lambda (Y_K - Z_K) \cdot \frac{e^{-net z}}{(1+e^{-net z})^2} \cdot w_{kj} \cdot F_2'(net j) \cdot h_b}$$

Scanned with CamScanner

$$\frac{\partial J}{\partial W_{bi}} = \frac{1}{2} \frac{\partial \sum (Y_K - Z_K)^2}{\partial W_{bi}}$$

$$= \sum_K \left(\frac{1}{\cancel{2}}\right)(\cancel{2})(Y_K - Z_K) \cdot \frac{\partial}{\partial W_{bi}}(Y_K - Z_K)$$

$K = 1$ so with/without summation its the same value.

$$= -(Y_K - Z_K) \cdot \frac{\partial}{\partial W_{bi}}(Z_K)$$

$$= -(Y_K - Z_K) \cdot \frac{\partial}{\partial W_{bi}}\left(\frac{1}{1+e^{-netz}}\right)$$

$$= -(Y_K - Z_K) \cdot \frac{\partial}{\partial netz}\left(\frac{1}{1+e^{-netz}}\right) \cdot \frac{\partial\, netz}{\partial W_{bi}}$$

$$= -(Y_K - Z_K) \cdot \frac{e^{-netz}}{(1+e^{-netz})^2} \cdot \frac{\partial (netz)}{\partial W_{bi}}$$

$$= -(Y_K - Z_K) \cdot \frac{e^{-netz}}{(1+e^{-netz})^2} \cdot \frac{\partial}{\partial W_{bi}} \sum W_{Kj} \cdot h_j$$

$$= -(Y_K - Z_K) \cdot \frac{e^{-netz}}{(1+e^{-netz})^2} \cdot \sum_j W_{Kj} \cdot \frac{\partial}{\partial W_{bi}}(h_j)$$

Scanned with CamScanner

$$\sum_j W_{kj} \cdot \frac{\partial}{\partial W_{bi}} \cdot h_j$$

$$h_j = f_2(net_j)$$

$$= \sum_j W_{kj} \cdot \frac{\partial}{\partial W_{bi}} f_2(net_j)$$

$$= \sum_j W_{kj} \cdot f_2'(net_j) \cdot \frac{\partial}{\partial W_{bi}}(net_j)$$

$$= \sum_j W_{kj} \cdot f_2'(net_j) \cdot \frac{\partial}{\partial W_{bi}}\left(\sum W_{jb} \cdot h_b\right)$$

$$= \sum_j W_{kj} \cdot f_2'(net_j) \cdot W_{jb} \cdot \frac{\partial}{\partial W_{bi}}(h_b)$$

$$= \sum_j W_{kj} \cdot f_2'(net_j) \cdot W_{jb} \cdot \frac{\partial}{\partial W_{bi}}(f_1(net_b))$$

$$= \sum_j W_{kj} \cdot f_2'(net_j) \cdot W_{jb} \cdot f_1'(net_b) \cdot \frac{\partial}{\partial W_{bi}}(net_b)$$

$$= \sum_j W_{kj} \cdot f_2'(net_j) \cdot W_{jb} \cdot f_1'(net_b) \cdot \frac{\partial}{\partial W_{bi}}\left(\sum W_{bi} \cdot x_i\right)$$

$$= \sum_j W_{kj} \cdot f_2'(net_j) \cdot W_{jb} \cdot f_1'(net_b) \cdot x_i$$

Scanned with CamScanner

$$W_{bi}^{new} = W_{bi}^{old} + \eta\,(y_k - z_k) \cdot \frac{e^{-net_z}}{(1+e^{-net_z})^2} \left(\sum W_{kj} \cdot F_2'(net_j) \cdot W_{jb} \cdot F_1'(net_b) \cdot x_i \right)$$

Scanned with CamScanner