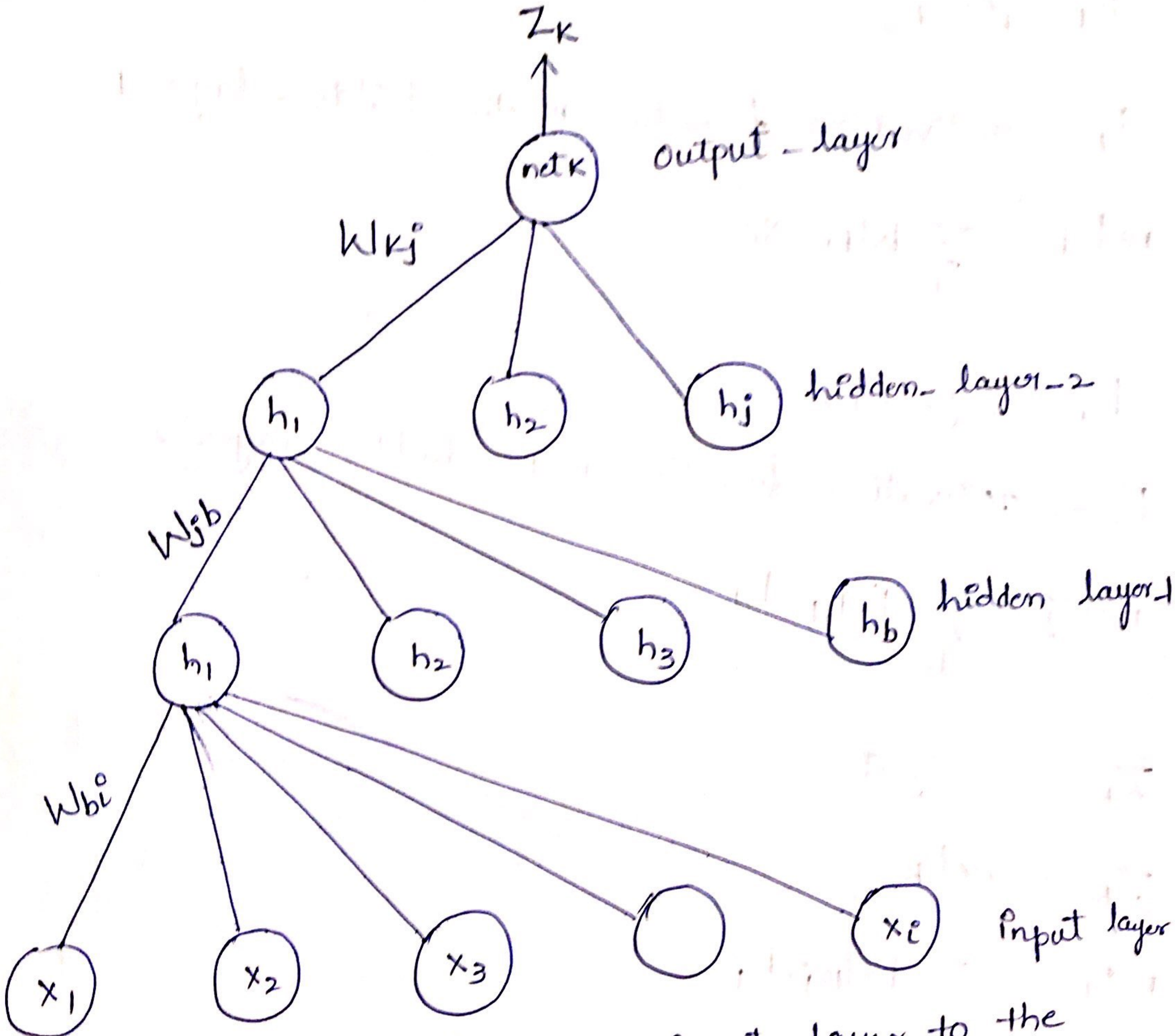
→ 4- Layer Neural Network for Regression



W_{bi} → Weights from the input layer to the hidden_layer_1

W_{jb} → Weights from hidden_layer_1 to hidden_layer_2

$W_{kj} \rightarrow$ Weights from the hidden_layer_2 to the ② output.

$h_b = f_1(net_b)$

$f_1 \rightarrow$ activation function in the hidden_layer_1

$net_b = \sum_i W_{bi} \cdot x_i$

$h_j = f_2(net_j)$

$f_2 \rightarrow$ activation function in the hidden_layer_2

$net_j = \sum_b W_{jb} \cdot h_b$

$Z_K \rightarrow$ output

$Z_K = net_K$

$net_K = \sum_j W_{kj} \cdot h_j$

<u>Goal</u> :- Tune the weights in each level in a way that reduces / minimizes the error

$$J(W_{kj}, W_{jb}, W_{bi}) = \frac{1}{2}(y - Z_k)^2$$

$$\frac{\partial J}{\partial W_{kj}} = \frac{\partial}{\partial W_{kj}}\left(\frac{1}{2}(y - Z_k)^2\right)$$

$$= \frac{1}{2} \cdot \frac{\partial}{\partial W_{kj}}(y - Z_k)^2$$

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

$$= (y - Z_k) \cdot \frac{\partial}{\partial W_{kj}}(-Z_k)$$

$$= -(y - Z_k) \cdot \frac{\partial}{\partial W_{kj}}\left(\sum_j W_{kj} \cdot h_j\right)$$

$$= -(y - Z_k) \cdot h_j$$

$$W_{kj}^{New} = W_{kj}^{old} - \lambda \cdot \left(\frac{\partial J}{\partial W_{kj}}\right)$$

$$\boxed{W_{kj}^{New} = W_{kj}^{old} + \lambda (y - Z_k) \cdot h_j}$$

$$\frac{\partial J}{\partial W_{jb}} = \frac{\partial}{\partial W_{jb}}\left(\tfrac{1}{2}(y - Z_K)^2\right)$$

$$= \tfrac{1}{2}\cdot\frac{\partial}{\partial W_{jb}}(y - Z_K)^2$$

$$= \left(\tfrac{1}{\cancel{2}}\right)(\cancel{2})\cdot(y - Z_K)\cdot\frac{\partial}{\partial W_{jb}}(y - Z_K)$$

$$= (y - Z_K)\cdot\frac{\partial}{\partial W_{jb}}(y - Z_K)$$

$$= -(y - Z_K)\cdot\frac{\partial}{\partial W_{jb}}(Z_K)$$

$$= -(y - Z_K)\cdot\frac{\partial}{\partial W_{jb}}(Z_K)$$

$$= -(y - Z_K)\cdot\frac{\partial}{\partial W_{jb}}\,net_K$$

$$= -(y - Z_K)\cdot\frac{\partial}{\partial W_{jb}}\cdot\left(\sum W_{kj}\cdot h_j\right)$$

$$= -(y - Z_K)\sum_K W_{kj}\cdot\frac{\partial}{\partial W_{jb}}(h_j)$$

$$= -(y - Z_K)\sum_K W_{kj}\cdot\frac{\partial}{\partial W_{jb}}F_2(net_j)$$

$$= -(y - Z_K)\sum_K W_{kj}\cdot\frac{\partial\cdot F_2(net_j)}{\partial\, net_j}\cdot\frac{\partial\, net_j}{\partial W_{jb}}$$

$$= -(Y-Z_K)\sum_K W_{Kj} \cdot f_2'(net_j) \cdot \frac{\partial}{\partial W_{jb}} \left(\sum_b W_{jb} \cdot h_b \right)$$

$$= -(Y-Z_K)\sum_K \left(W_{Kj} \cdot f_2'(net_j) \right) h_b$$

$$W_{jb}^{New} = W_{jb}^{old} - \lambda \cdot \frac{\partial J}{\partial W_{jb}}$$

$$\boxed{W_{jb}^{New} = W_{jb}^{old} + \lambda (Y-Z_K)\sum_K \left(W_{Kj} \cdot f_2'(net_j) \right) h_b}$$

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

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

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

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

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

$$= \frac{\partial J}{\partial W_{bi}} = -(Y-Z_K)\cdot \frac{\partial}{\partial W_{bi}}(net\,K).$$

$$= -(Y-Z_K)\cdot \frac{\partial}{\partial W_{bi}}\left(\sum_j W_{Kj}\cdot h_j\right)$$

$$= -(Y-Z_K)\cdot \sum_j W_{Kj}\cdot \frac{\partial}{\partial W_{bi}}(h_j)$$

$$= -(Y-Z_K)\cdot \sum_{K,j} \cdot W_{Kj}\cdot \frac{\partial}{\partial W_{bi}}(F_2(net_j))$$

$$= -(Y-Z_K)\cdot \sum_{K,j} W_{Kj}\cdot \left(\frac{\partial F_2(net_j)}{\partial net_j}\cdot \frac{\partial net_j}{\partial W_{bi}}\right)$$

$$= -(Y-Z_K)\cdot \sum_{K,j} W_{Kj}\cdot F_2'(net_j)\cdot \frac{\partial\, net_j}{\partial W_{bi}}$$

$$= -(Y-Z_K)\cdot \sum_{K,j} W_{Kj}\cdot F_2'(net_j)\cdot \frac{\partial}{\partial W_{bi}}\left(\sum_b W_{jb}\cdot h_b\right)$$

$$= -(Y-Z_K)\cdot \sum_{K,j} W_{Kj}\cdot F_2'(net_j)\cdot W_{jb}\cdot \frac{\partial}{\partial W_{bi}}(h_b)$$

$$= -(Y-Z_K)\cdot \sum_{K,j} W_{Kj}\cdot F_2'(net_j)\cdot W_{jb}\cdot \frac{\partial}{\partial W_{bi}}(F_1(net_b))$$

(13)

$$= -(Y-Z_K)\cdot\sum_{K,j} W_{Kj}\cdot F_2'(net_j)\cdot W_{jb}\cdot\frac{\partial F_1(net_b)}{\partial\, net_b}\cdot\frac{\partial\, net_b}{\partial W_{bi}}$$

$$= -(Y-Z_K)\cdot\sum_{K,j} W_{Kj}\cdot F_2'(net_j)\cdot W_{jb}\cdot F_1'(net_b)\cdot\frac{\partial}{\partial W_{bi}}\left(\sum_i W_{bi}\cdot X_i\right)$$

$$= -(Y-Z_K)\sum_{K,j} W_{Kj}\cdot F_2'(net_j)\cdot W_{jb}\cdot F_1'(net_b)\cdot X_i$$

$$= -(Y-Z_K)\left(\sum_{K,j} W_{Kj}\cdot F_2'(net_j)\cdot W_{jb}\cdot F_1'(net_b)\right)\cdot X_i$$

$$W_{bi}^{New} = W_{bi}^{old} - \lambda\cdot\frac{\partial J}{\partial W_{bi}}$$

$$W_{bi}^{New} = W_{bi}^{old} - \lambda\left(-(Y-Z_K)\left(\sum_{K,j} W_{Kj}\cdot F_2'(net_j)\cdot W_{jb}\cdot F_1'(net_b)\right)\cdot X_i\right)$$

$$\boxed{W_{bi}^{New} = W_{bi}^{old} + \lambda\,(Y-Z_K)\left(\sum_{K,j} W_{Kj}\cdot F_2'(net_j)\cdot W_{jb}\cdot F_1'(net_b)\right)(X_i).}$$