

- Consider penultimate layers of node.
- Expression for $\frac{\partial E_n(w)}{\partial a_i^{[2]}}$. Use $\delta_i^{[3]}$.

$$\delta_i^{[3]} \equiv \frac{\partial E_n(w)}{\partial a_i^{[3]}}$$

Chain rule:-

$$\frac{\partial E_n(w)}{\partial a_i^{[2]}} = \boxed{\frac{\partial E_n(w)}{\partial a_i^{[3]}}} \times \frac{\partial a_i^{[3]}}{\partial a_i^{[2]}}$$

↓
 $\delta_i^{[3]}$

blocks of $a_1^{[3]}$
share nothing

$$\therefore a_i^{[3]} = w_{i1}^{[3]} h(a_1^{[2]}) + w_{i2}^{[3]} h(a_2^{[2]}) + w_{i3}^{[3]} h(a_3^{[2]})$$

$$\frac{\partial a_i^{[3]}}{\partial a_i^{[2]}} = \boxed{w_{ii}^{[3]} h'(a_i^{[2]})}$$

$$\therefore \frac{\partial E_n(w)}{\partial a_i^{[2]}} = \delta_i^{[3]} w_{ii}^{[3]} h'(a_i^{[2]}) \quad \text{--- (1)}$$

Use this to calc $\frac{\partial E_n(w)}{\partial w_{ii}^{[2]}}$

$$\frac{\partial E_n(w)}{\partial w_{ii}^{[2]}} = \frac{\partial E_n(w)}{\partial a_i^{[2]}} \times \frac{\partial a_i^{[2]}}{\partial w_{ii}^{[2]}}$$

$$\therefore a_i^{[2]} = w_{i1}^{[2]} h(a_1^{[1]}) + w_{i2}^{[2]} h(a_2^{[1]}) + w_{i3}^{[2]} h(a_3^{[1]})$$

$$\frac{\partial a_i^{[2]}}{\partial w_{ii}^{[2]}} = h(a_i^{[1]}) \quad \text{--- (2)}$$

(1) and (2) in (1)

$$\frac{\partial E_n(w)}{\partial w_{ii}^{[2]}} = \delta_i^{[3]} w_{ii}^{[3]} h'(a_i^{[2]}) h(a_i^{[1]})$$

X — X