

2. Input vector : $\begin{bmatrix} 8 \\ 5 \\ 0 \end{bmatrix}$

$$\text{softmax}(z_i) = \frac{e^{z_i}}{\sum_j e^{z_j}}$$

$$e^8 = 2980.96$$

$$e^5 = 148.41$$

$$e^0 = 1$$

$$\begin{aligned} \sum_j e^{z_j} &= 2980.96 + 148.41 + 1 \\ &= 3130.37 \end{aligned}$$

$$\text{softmax}(8) = \frac{e^8}{3130.37} \approx 0.952$$

$$\text{softmax}(5) = \frac{e^5}{3130.37} \approx 0.0474$$

$$\text{softmax}(0) = \frac{e^0}{3130.37} \approx 0.000319$$

Output probability : $\begin{bmatrix} 0.952 \\ 0.0474 \\ 0.000319 \end{bmatrix}$

$$\text{Loss} = - \sum_i y_i \log(p_i)$$

Target vector $\vec{y} = [1, 0, 0]$

$$\text{Loss} = - [1 \cdot \log(0.952) + 0 \cdot \log(0.0474) + 0 \cdot \log(0.000319)]$$

$$\text{Loss} = -\log(0.952) \approx 0.0496$$

$$3. \quad y = \sigma(z) = \sigma(x \cdot w + b)$$

$$z = x \cdot w + b$$

Neuron 1

$$x = [1, 8, -6, -2]$$

$$w = [1.8, 0.1, 2.7, 3.5]$$

$$b = 2.3$$

$$z = x \cdot w + b$$

$$z = (1 \times 1.8) + (8 \times 0.1) + (-6 \times 2.7) + (-2 \times 3.5) + 2.3$$

$$z = -18.3$$

$$y_1 = \sigma(z) = \frac{1}{1 + e^{-z}} = \frac{1}{1 + e^{-18.3}} \approx 1.12 \times 10^{-8}$$

Neuron 2

$$x = [11, 81, -16, -21]$$

$$w = [7.8, 0.1, 3.7, 3.8]$$

$$b = 2.5$$

$$2 \times 2 \times 2 \times 16$$

$$= (11 \times 7.8) + (81 \times 0.1) + (-16 \times 3.7) + (-21 \times 3.8) + 2.5$$

$$= -42.6$$

$$y_2 = \frac{1}{1 + e^{-2}} = \frac{1}{1 + e^{42.6}}$$

$$\approx 3.15 \times 10^{-10}$$

$$\text{Loss} = -\frac{1}{2} \sum_{i=1}^2 (t_i \log(y_i) + (1-t_i) \log(1-y_i))$$

$$t = [1, 1]$$

$$t_1 = 1, \quad t_2 = 1$$

$$\text{Loss} = -\frac{1}{2} (\log(y_1) + \log(y_2))$$

$$= -\frac{1}{2} (\log(1.12 \times 10^{-8}) + \log(3.15 \times 10^{-10}))$$

$$= -\frac{1}{2} (-18.31 - 42.60)$$

$$= 30.455$$

$$\boxed{\text{Loss} = 30.46}$$

$$2. \text{ Input channel } 1 = \begin{bmatrix} 23 & 145 & 1 \\ 34 & 132 & 10 \\ 76 & 145 & 32 \end{bmatrix}$$

$$\text{Filter channel } 1 = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}$$

$$(23 \times -1) + (145 \times 1) + (34 \times -1) + (132 \times 1) = 104$$

$$\text{Output } 12 = \begin{bmatrix} -104 & -156 \\ 107 & -135 \end{bmatrix}$$

$$\text{Output channel } 2 = \begin{bmatrix} 11 & 221 & 12 \\ 13 & 190 & 65 \\ 45 & 196 & 56 \end{bmatrix}$$

$$\text{Filter channel } 2 = \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix}$$

$$\text{Output } 22 = \begin{bmatrix} 2 & -31 \\ 32 & 6 \end{bmatrix}$$

$$\text{Output channel } 3 = \begin{bmatrix} 34 & 154 & 75 \\ 85 & 130 & 89 \\ 56 & 178 & 90 \end{bmatrix}$$

$$\text{Filter channel } 3 = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}$$

Output 32, $\begin{bmatrix} -36 & -14 \\ 12 & -1 \end{bmatrix}$

Final output

$$2 \begin{bmatrix} 104 & -156 \\ 107 & -135 \end{bmatrix} + \begin{bmatrix} 2 & -31 \\ 32 & -6 \end{bmatrix} + \begin{bmatrix} -36 & -14 \\ 12 & -1 \end{bmatrix}$$

$$2 \begin{bmatrix} 70 & -201 \\ 151 & -130 \end{bmatrix}$$