6427 Exercise LSTM

Q: (i) $f_1 = ?$ comment?

(ii) $c_1 = ?$

(iii) $h_1 = ?$

$$\begin{bmatrix} i \\ f \\ o \\ g \end{bmatrix} = \begin{bmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{bmatrix} W \begin{bmatrix} h_{(t-1)} \\ x_t \end{bmatrix}$$

$$c_t = f \odot c_{t-1} + i \odot g$$

$$h_t = o \odot \tanh(c_t)$$

Solution (i) ① let $t=0$

$$h_0 = o_0 \odot \tanh(c_0)$$

② let $t=1$

$$h_1 = o_1 \odot \tanh(c_1)$$

$$c_1 = f_1 \odot c_0 + i_1 \odot g_1$$

$$\begin{bmatrix} i_1 \\ f_1 \\ o_1 \\ g_1 \end{bmatrix} = \begin{bmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{bmatrix} W \begin{bmatrix} h_0 \\ x_1 \end{bmatrix} \quad (1)$$

$$f_1 = \sigma\left(W_f \begin{bmatrix} h_0 \\ x_1 \end{bmatrix}\right)$$

$$= \sigma\left(\begin{bmatrix} 0.1 & 0.2 & 0.5 & 0.6 \\ 0.3 & 0.4 & 0.7 & 0.8 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 2 \\ 1 \end{bmatrix}\right)$$

$$= \sigma \begin{bmatrix} 1.6 \\ 2.2 \end{bmatrix}$$

$$= \begin{bmatrix} 0.8320 \\ 0.9002 \end{bmatrix} \qquad f_1 = \begin{bmatrix} 0.822 \\ 0.900 \end{bmatrix} \text{ (round to 3 decimal places)}$$

The sell state at $t=1$ retains most of the memory from the previous cell state $t=0$

(ii) $C_1 = f_1 \odot C_0 + i_1 \odot g_1$

$$= \begin{bmatrix} 0.822 \\ 0.900 \end{bmatrix} \odot \begin{bmatrix} 0.1 \\ 0.2 \end{bmatrix} + \begin{bmatrix} 0.3 \\ 0.4 \end{bmatrix} \odot \begin{bmatrix} 0.5 \\ 0.6 \end{bmatrix}$$

$$= \begin{bmatrix} 0.0822 \\ 0.1800 \end{bmatrix} + \begin{bmatrix} 0.15 \\ 0.24 \end{bmatrix}$$

$$= \begin{bmatrix} 0.2322 \\ 0.4200 \end{bmatrix}$$

$C_1 = \begin{bmatrix} 0.232 \\ 0.420 \end{bmatrix}$ (round to 3 decimal places)

(ii) $h_1 = o_1 \odot \tanh(c_1)$

$$= \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix} \odot \tanh \begin{bmatrix} 0.232 \\ 0.420 \end{bmatrix}$$

$$= \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix} \odot \begin{bmatrix} 0.2279 \\ 0.3969 \end{bmatrix}$$

$$= \begin{bmatrix} 0.0916 \\ 0.2382 \end{bmatrix}$$