RADHIKA PATWARI
18CS10062

15/3/2021

Class Test 3 (CS60010)

1. (a) i) $f_t = 4$
ii) $i_t = 5$
iii) $o_t = 8$
iv) $C_t = 3$
v) $C_{t-1} = 1$
vi) $\tilde{C}_t = 7$

(b) (i) $C_t = (C_{t-1} \odot f_t) + (i_t \odot \tilde{C}_t)$
(3×1) (3×1) (3×1) (3×1)

$$= \begin{bmatrix} 0.2 \\ 0.4 \\ 0.8 \end{bmatrix} \odot \begin{bmatrix} 1 \\ 0.5 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0.5 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0 \\ 0.1 \end{bmatrix}$$

$$= \begin{bmatrix} 0.2 \\ 0.2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0.5 \\ 0 \\ 0.05 \end{bmatrix} = \begin{bmatrix} 0.7 \\ 0.2 \\ 0.05 \end{bmatrix}$$

$$= \{0.7, 0.2, 0.05\}$$

(ii) $h_t = o_t \odot \tanh(C_t)$

$$= \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \odot \tanh \begin{bmatrix} 0.7 \\ 0.2 \\ 0.05 \end{bmatrix}$$

$$= \begin{bmatrix} \tanh(0.7) \\ 0 \\ 0 \end{bmatrix} = 0.604367\ldots$$

$$= 0.6044$$

$$\tanh(0.7) = \frac{e^{0.7} - e^{-0.7}}{e^{0.7} + e^{-0.7}} = \frac{e^{2\times0.7} - 1}{e^{2\times0.7} + 1}$$

$$= \frac{e^{1.4} - 1}{e^{1.4} + 1} = 0.60436$$

---

3. $C_t = 3\times1$, $X_t = 2\times1$

(a) $W_f = 3\times2$

(b) $b_f = 3\times1$

(c) Parameters = $W_f, b_f, W_i, b_i, W_o, b_o, W_c, b_c$

Total no. of parameters = 8 variables.

W_f = 6 parameters

b_f = 3 parameters

$= 4\times6 + 4\times3$

$= 24 + 12 = 36$ parameters in 1 LSTM cell.

---

2. (a) $e_{kj} = (S_{k-1} \cdot h_j)$

(b) 0

5. Depthwise convolution

Input: 200 × 200 × 3

filter = 4 × 4 stride = 1, padding = 1

Output $= \left\lfloor \frac{200 - 4 + 2}{1} \right\rfloor + 1 = 199$

Output: 199 × 199 × 3

Computation = 4 × 4 × 200 × 200 × 3 = 1920000

Point wise convolution

Input: 199 × 199 × 3, filter = 1 × 1 × 3 } 10 such filters.

Output = 199 × 199 × 10

Computation = 199 × 199 × 3 × 10 = 1188030

∴ Total computation = 4 × 4 × 200 × 200 × 3 + 199 × 199 × 30

= 3,108,030

= 3M × 4 bytes

= 12 Mb

∴ Output after depthwise convolution = 199 × 199 × 3

Output after pointwise convolution = 199 × 199 × 10

4. $c_{t+1} = (c_t \odot f_{t+1}) + (i_{t+1} \odot \tilde{c}_{t+1})$

(a)

$$f_t = \sigma(W_f [h_t, x_t] + b_f)$$

$$\frac{\partial L}{\partial h_T} = \frac{\partial V_T}{\partial h_T} \times \frac{\partial \hat{y}^T}{\partial V_T} \times \frac{\partial L}{\partial \hat{y}^T}$$

$$c_T = (c_{T-1} \odot f_T) + (i_T \odot \tilde{c}_T)$$

$$h_T = o_T \odot \tanh(c_T)$$

$$\frac{\partial h_T}{\partial c_T} = (o_T)(1 - \tanh^2(c_T))$$

$$\frac{\partial L}{\partial c_T} = \frac{\partial h_T}{\partial c_T} \times \frac{\partial L}{\partial h_T}$$

$$= (o_t)(1 - \tanh^2(c_t)) \times \frac{\partial L}{\partial h_T}$$