Shreeya Kokate
CWID : 20005256

PAGE No.
DATE / /

Assignment 1 (Bonus) : RNN

Q1] Binary Addition

⇒ Let's consider,

$x^{(t)}$ as vector 2×1 with x_1 & x_2 as input at time t.

$h^{(t)}$ as vector 3×1 with h_1, h_2 & h_3 as hidden unit at time t.

$y^{(t)}$ as a scalar of output.

Also,

$$U = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$W = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$b_h = \begin{bmatrix} -0.5 \\ -1.5 \\ -2.5 \end{bmatrix}$$

$$V = \begin{bmatrix} 1 & -1 & 1 \end{bmatrix}$$

$$by = -0.5$$

So, for $t \geq 1$ we can get,

$$h^{(t)} = Ux^{(t)} + Wh^{(t-1)} + b_h$$

$$y^{(t)} = Vh^{(t)} + by$$

Lets expand $h^{(t)}$ and $y^{(t)}$ obtain,

$$h_1^{(t)} = x_1^{(t)} + x_2^{(t)} + h_1^{(t-1)} - 0.5$$

$$h_2^{(t)} = x_1^{(t)} + x_2^{(t)} + h_2^{(t-1)} - 1.5$$

$$h_3^{(t)} = x_1^{(t)} + x_2^{(t)} + h_3^{(t-1)} - 2.5$$

$$y^{(t)} = h_1^{(t)} - h_2^{(t)} + h_3^{(t)} - 0.5$$

Q2] Given computations :-

$$i^{(t)} = \sigma(w_{ix} x^{(t)} + w_{ih} h^{(t-1)})$$

$$f^{(t)} = \sigma(w_{fx} x^{(t)} + w_{fh} h^{(t-1)})$$

$$o^{(t)} = \sigma(w_{ox} x^{(t)} + w_{oh} h^{(t-1)})$$

$$g^{(t)} = \tanh(w_{gx} x^{(t)} + w_{gh} h^{(t-1)})$$

$$c^{(t)} = f^{(t)} c^{(t-1)} + i^{(t)} g^{(t)}$$

$$h^{(t)} = o^{(t)} \tanh(c^{(t)})$$

a] To derive backprop through time equations for the activations & gates :-

$$\frac{\partial h^{(t)}}{\partial h^{(t+1)}} = 1 + \overline{i^{(t+1)}} \frac{\partial i^{(t+1)}}{\partial h^{(t+1)}} + \overline{f^{(t+1)}} \frac{\partial f^{(t+1)}}{\partial h^{(t+1)}} + \overline{o^{(t+1)}} \frac{\partial o^{(t+1)}}{\partial h^{(t+1)}} + \overline{g^{(t+1)}} \frac{\partial g^{(t+1)}}{\partial h^{(t+1)}}$$

$$= 1 + \overline{i^{(t+1)}} \sigma^{-1}(w_{ix} x^{(t+1)} + w_{ih} h^{(t+1)}) \cdot w_{ih}$$

$$+ \overline{f^{(t+1)}} \sigma^{-1}(w_{fx} x^{(t+1)} + w_{fh} h^{(t+1)}) \cdot w_{fh}$$

$$+ \overline{o^{(t+1)}} \sigma^{-1}(w_{ox} x^{(t+1)} + w_{oh} h^{(t+1)}) \cdot w_{oh}$$

$$+ \overline{g^{(t+1)}} \tanh^{-1}(w_{gx} x^{(t+1)} + w_{gh} h^{(t+1)}) \cdot w_{gh}$$

$$\overline{c^{(t)}} = \overline{h^{(t)}} \frac{\partial h^{(t)}}{\partial c^{(t)}} + \overline{c^{(t+1)}} \frac{\partial c^{(t+1)}}{\partial c^{(t)}}$$

$$= \overline{h^{(t)}}\, o^{(t)} \tanh'(c^{(t)}) + \overline{c^{(t+1)}}\, f^{(t)}$$

$$\overline{g^{(t)}} = \overline{c^{(t)}} \frac{\partial c^{(t)}}{\partial g^{(t)}}$$

$$= \overline{c^{(t)}}\, i^{(t)}$$

$$\overline{o^{(t)}} = \overline{h^{(t)}} \frac{\partial h^{(t)}}{\partial o^{(t)}}$$

$$= \overline{h^{(t)}} \tanh(c^{(t)})$$

$$\overline{f^{(t)}} = \overline{c^{(t)}} \frac{\partial c^{(t)}}{\partial f^{(t)}}$$

$$= \overline{c^{(t)}}\, c^{(t-1)}$$

$$\overline{i^{(t)}} = \overline{c^{(t)}}\, g^{(t)}$$