

Assignment - 3

RNN's (Recursive Neural Networks)

$$CE(y, \hat{y}) = - \sum_i y_i \log(\hat{y}_i)$$

$$y \in \mathbb{R}^{1 \times 5}$$

$$h^{(1)} = \max(0, \underbrace{[h_{\text{Left}}^{(1)}, h_{\text{right}}^{(1)}]w^{(1)} + b^{(1)}}_{z^{(1)}})$$

$$\hat{y} = \underbrace{\text{softmax}(h^{(1)}U + b^{(2)})}_{z^{(2)}}$$

$$w^{(1)} \in \mathbb{R}^{2d \times d}$$

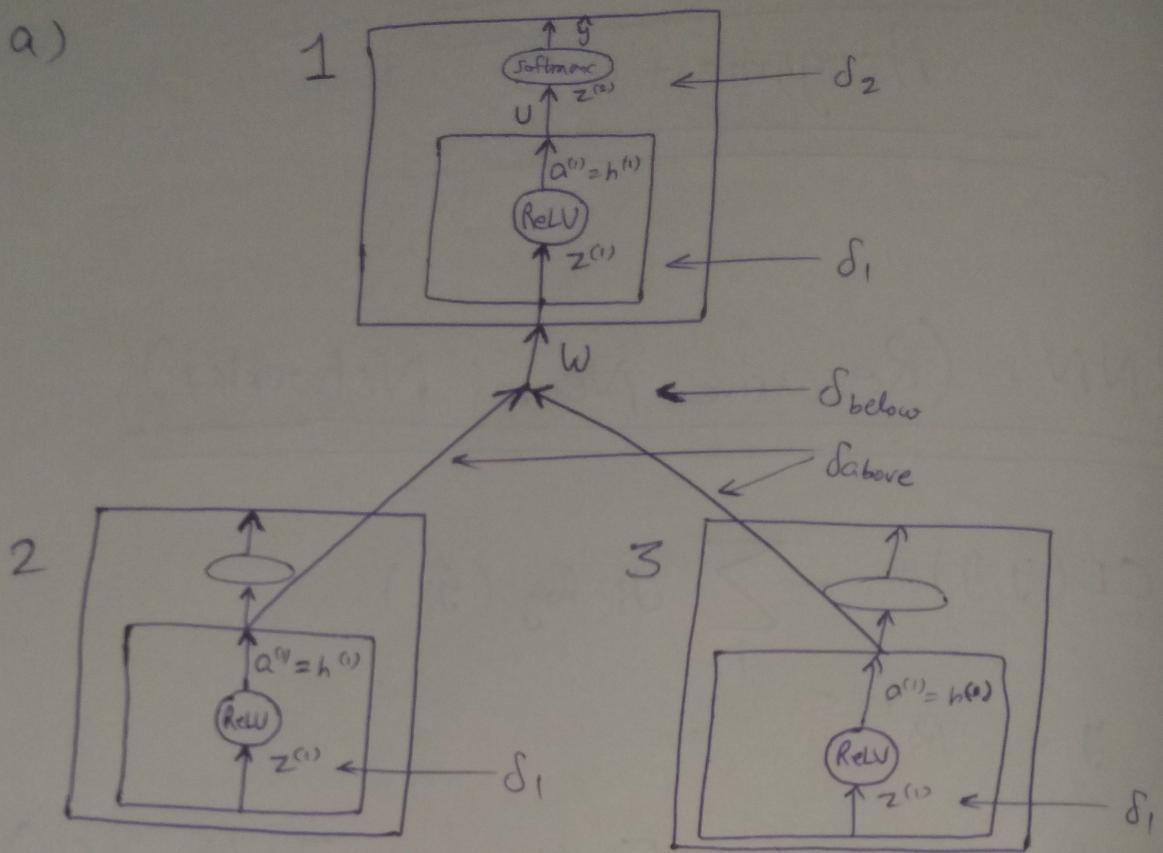
$$b^{(1)} \in \mathbb{R}^{1 \times d}$$

$$U \in \mathbb{R}^{d \times 5}$$

$$b^{(2)} \in \mathbb{R}^{1 \times 5}$$

$$L \in \mathbb{R}^{M \times d}$$

(a)



for root node

$$\delta_{\text{above}} = 0$$

$$\frac{\partial J}{\partial U} = h^{(1)T} \delta_2$$

$$\delta_2 = \hat{y} - y$$

$$\frac{\partial J}{\partial b^{(2)}} = \delta_2$$

for all other (intermediate) nodes :

$$\frac{\partial J}{\partial h^{(1)}} = \delta_2 U^T + \delta_{\text{above}}$$

$$\delta_1 = (\delta_2 U^T + \delta_{\text{above}}) \odot \text{ReLU}'(z_1)$$

$$\begin{aligned} \text{ReLU}'(z_1) &= \begin{cases} 1, & z_1 > 0 \\ 0, & \text{o/w} \end{cases} = \mathbb{1}(\text{ReLU}(z_1) > 0) \\ &= \mathbb{1}(h^{(1)} > 0) \end{aligned}$$

$$\delta_1 = (\delta_2 U^T + \delta_{\text{above}}) \odot \mathbb{1}(h^{(1)} > 0)$$

$$\frac{\partial J}{\partial w^{(1)}} = [h_{\text{Left}}^{(1)}, h_{\text{Right}}^{(1)}]^T \delta_1$$

$$\frac{\partial J}{\partial b^{(1)}} = \delta_1$$

$$\delta_{\text{below}} = \delta_1 w^T$$

$$[\delta_{\text{Left, above}}, \delta_{\text{Right, above}}] = \delta_{\text{below}} = \delta_1 w^T$$

For Leaf nodes;

$$\frac{\partial J}{\partial L_i} = \delta_2 U^T + \delta_{\text{above}}$$