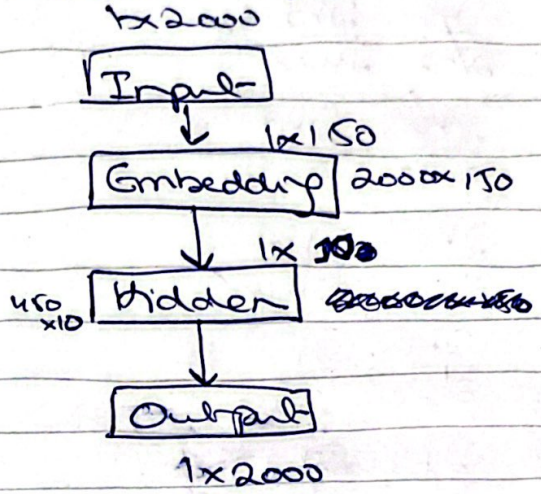
21L-5625
Rohan Javed

Class Activity

Q1)

Total words = 300,000
Vocabulary = 2000
Hidden = 10 units
Context size



$w_h = 450 \times 10$
$w_o = 10 \times 2000$
E = embedding
$I = [E_1, E_2, E_3]$
$h_t = \tanh(w_h I + b_h)$
$O = \text{sigmoid}(w_o h_t + b_o)$

Q2)

$$\frac{\partial L}{\partial w_x} = \sum \frac{\partial g_t}{\partial z_t} \cdot \frac{\partial z_t}{\partial h_t} \cdot \frac{\partial h_t}{\partial h_k} \cdot \frac{\partial h_k}{\partial w_x}$$

$$= \frac{\partial L_1}{\partial \hat{y}_1} \cdot \frac{\partial \hat{y}_1}{\partial z_1} \cdot \frac{\partial z_1}{\partial h_1} \cdot \frac{\partial h_1}{\partial h_0} \cdot \frac{\partial h_0}{\partial w_x} +$$

$$= \frac{\delta L_1}{\delta \hat{y}_1} \cdot \frac{\delta \hat{y}_1}{\partial z_1} \cdot \frac{\partial z_1}{\partial h_1} \cdot \frac{\delta h_1}{\partial h_1} \cdot \frac{\partial h_1}{\partial w_x} +$$

$$= \frac{\partial L_1}{\partial \hat{y}_1} \cdot \frac{\partial \hat{y}_1}{\partial z_1} \cdot \frac{\partial z_1}{\partial h_1} \cdot \frac{\partial h_1}{\partial h_2} \cdot \frac{\delta h_2}{\partial w_x} +$$

$$\frac{\partial L_1}{\partial \dot{y}_1}\cdot\frac{\partial \dot{y}_1}{\partial z_1}\cdot\frac{\partial z_1}{\partial h_1}\cdot\frac{\partial h_1}{\partial h_3}\cdot\frac{\partial h_3}{\partial w_x} +$$

$$\frac{\partial L_2}{\partial \dot{y}_2}\cdot\frac{\partial \dot{y}_2}{\partial z_2}\cdot\frac{\partial z_2}{\partial h_2}\cdot\frac{\partial h_2}{\partial h_0}\cdot\frac{\partial h_0}{\partial w_x} +$$

$$\frac{\partial L_2}{\partial \dot{y}_2}\cdot\frac{\partial \dot{y}_2}{\partial z_2}\cdot\frac{\partial z_2}{\partial h_2}\cdot\frac{\partial h_2}{\partial h_1}\cdot\frac{\partial h_1}{\partial w_x} +$$

$$\frac{\partial L_2}{\partial \dot{y}_2}\cdot\frac{\partial \dot{y}_2}{\partial z_2}\cdot\frac{\partial z_2}{\partial h_2}\cdot\frac{\partial h_2}{\partial h_2}\cdot\frac{\partial h_2}{\partial w_x} +$$

$$\frac{\partial L_2}{\partial \dot{y}_2}\cdot\frac{\partial \dot{y}_2}{\partial z_2}\cdot\frac{\partial z_2}{\partial h_3}\cdot\frac{\partial h_3}{\partial h_3}\cdot\frac{\partial h_3}{\partial w_x} +$$

$$\frac{\partial L_3}{\partial \dot{y}_3}\cdot\frac{\partial \dot{y}_3}{\partial z_3}\cdot\frac{\partial z_3}{\partial h_3}\cdot\frac{\partial h_3}{\partial h_0}\cdot\frac{\partial h_0}{\partial w_x} +$$

$$\frac{\partial L_3}{\partial \dot{y}_3}\cdot\frac{\partial \dot{y}_3}{\partial z_3}\cdot\frac{\partial z_3}{\partial h_3}\cdot\frac{\partial h_3}{\partial h_1}\cdot\frac{\partial h_1}{\partial w_x} +$$

$$\frac{\partial L_3}{\partial \dot{y}_3}\cdot\frac{\partial \dot{y}_3}{\partial z_3}\cdot\frac{\partial z_3}{\partial h_3}\cdot\frac{\partial h_3}{\partial h_2}\cdot\frac{\partial h_2}{\partial w_x} +$$

$$\frac{\partial L_3}{\partial \dot{y}_3}\cdot\frac{\partial \dot{y}_3}{\partial z_3}\cdot\frac{\partial z_3}{\partial h_3}\cdot\frac{\partial h_3}{\partial h_3}\cdot\frac{\partial h_3}{\partial w_x}$$