

①

"RNN Example"

$|V| = 7$

The Students opened their

$V = \{$

the,
Students,
opened,
their,
Laptops,
books,
~~mini~~,
exam $\}$

one hot
created
vector
of size
'7'

$\begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

$\begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

$\begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$

$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$

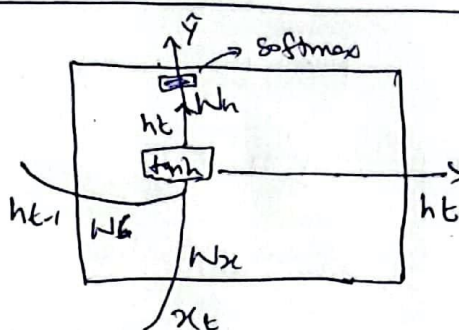
projects to get embeddings

$\begin{bmatrix} 3 \\ 0 \\ 4 \\ 4 \end{bmatrix}$
 e_1

$\begin{bmatrix} 7 \\ 7 \\ 0 \\ 0 \end{bmatrix}$
 e_2

$\begin{bmatrix} 8 \\ 0 \\ -8 \\ 0 \end{bmatrix}$
 e_3

$\begin{bmatrix} 2 \\ 5 \\ 2 \\ 5 \end{bmatrix}$
 e_4



Context
units = '3'

$$I = \begin{bmatrix} h_{t-1} & x_t \end{bmatrix}$$

$1 \times 7 \quad 1 \times 3 \quad 1 \times 4$

$$W_H = \begin{bmatrix} W_c & W_x \end{bmatrix}$$

$3 \times 7 \quad 3 \times 3 \quad 3 \times 4$

$$B_H = \begin{bmatrix} \end{bmatrix}$$

3×1

$$h_t = \begin{bmatrix} \end{bmatrix}$$

1×3

$$\hat{y} = \begin{bmatrix} \end{bmatrix}$$

1×7

$$W_o = \begin{bmatrix} W_h \end{bmatrix} B_o = \begin{bmatrix} \end{bmatrix}$$

$7 \times 3 \quad 7 \times 1$

$$W_H = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ -3 & -3 & -3 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & -2 & -2 & -2 & -2 \end{bmatrix}$$

3×7

$$\textcircled{2} \quad t=0, \quad h_{t-1} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} \quad x_t = \begin{bmatrix} 3 & 0 & 4 & 4 \end{bmatrix}$$

$$a_t = (W_H \cdot \vec{I}) + B_H$$

$$a_t = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ -3 & -3 & -3 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & -2 & -2 & -2 & -2 \end{bmatrix}_{3 \times 7} \cdot \begin{bmatrix} 0 \\ 0 \\ 0 \\ 3 \\ 0 \\ 4 \\ 4 \end{bmatrix}_{7 \times 1} + \begin{bmatrix} 0 \\ 0 \\ 7 \end{bmatrix}_{3 \times 1}$$

$$a_t = \begin{bmatrix} 0 \\ 11 \\ -22 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 7 \end{bmatrix} = \begin{bmatrix} 0 \\ 11 \\ -15 \end{bmatrix}$$

$$h_t = \tanh \begin{bmatrix} 0 \\ 11 \\ -15 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

$$t=1, \quad h_{t-1} = \begin{bmatrix} 0 & 1 & -1 \end{bmatrix} \quad x_t = \begin{bmatrix} 7 & 7 & 0 & 0 \end{bmatrix}$$

$$a_t = (W_H \cdot \vec{I}) + B_H$$

$$a_t = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ -3 & -3 & -3 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & -2 & -2 & -2 & -2 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ -1 \\ 7 \\ 7 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 7 \end{bmatrix}$$

$$a_t = \begin{bmatrix} 0 \\ 14 \\ -28 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 7 \end{bmatrix} = \begin{bmatrix} 0 \\ 14 \\ -21 \end{bmatrix}$$

$$h_t = \tanh \begin{bmatrix} 0 \\ 14 \\ -21 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

$$\textcircled{3} t=2, \quad h_{t-1} = [0 \ 1 \ -1] \quad x_t = [-8 \ 0 \ -8 \ 0]$$

$$a_t = (W_H \cdot I^T) + B_H$$

$$a_t = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ -3 & -3 & -3 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & -2 & -2 & -2 & -2 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ -1 \\ -8 \\ 0 \\ -8 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 7 \end{bmatrix}$$

$$a_t = \begin{bmatrix} 0 \\ -16 \\ 39 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 7 \end{bmatrix} = \begin{bmatrix} 0 \\ -16 \\ 39 \end{bmatrix}$$

$$h_t = \tanh \begin{bmatrix} 0 \\ -16 \\ 39 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

$$t=3, \quad h_{t-1} = [0 \ -1 \ 1] \quad x_t = [2 \ 5 \ 2 \ 5]$$

$$a_t = (W_H \cdot I^T) + B_H$$

$$a_t = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ -3 & -3 & -3 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & -2 & -2 & -2 & -2 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ -1 \\ 1 \\ 2 \\ 5 \\ 2 \\ 5 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 7 \end{bmatrix}$$

$$a_t = \begin{bmatrix} 0 \\ 14 \\ -28 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 7 \end{bmatrix} = \begin{bmatrix} 0 \\ 14 \\ -21 \end{bmatrix}$$

$$h_t = \tanh \begin{bmatrix} 0 \\ 14 \\ -21 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

* Now predicting
output word

$$\begin{matrix} W_0 & \begin{matrix} \rightarrow \\ 7 \times 3 \end{matrix} & \begin{bmatrix} 1 & -2 & -3 \\ 1 & -2 & +3 \\ 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 3 & -5 \\ 1 & 2 & 2 \\ 1 & -2 & -2 \end{bmatrix} & \begin{matrix} h_t & \begin{matrix} 2 \\ 3 \times 1 \end{matrix} \\ \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \end{matrix} & + & \begin{matrix} \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 2 \\ 0 \\ 0 \end{bmatrix} & \begin{matrix} 7 \times 1 \\ B_0 \end{matrix} \end{matrix}$$

Q)

$$b_t = (W_0 * h_t) + B_0$$

$$b_t = \begin{pmatrix} 1 \\ -5 \\ 0 \\ 0 \\ 8 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ -5 \\ 0 \\ 0 \\ 10 \\ 0 \\ 0 \end{pmatrix}$$

$$y_t = \text{Softmax} \begin{pmatrix} 3 \\ -5 \\ 0 \\ 0 \\ 10 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} 0.0009 \\ 0 \\ 0 \\ 0 \\ 0.9989 \\ 0 \\ 0 \end{pmatrix} \begin{array}{l} \text{The} \\ \text{student} \\ \text{opened} \\ \text{their} \\ \text{laptops} \\ \text{books} \\ \text{exam} \end{array}$$

$$y_t = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \underline{\underline{\text{Laptops}}}$$