NLP - LSTM Assignment.

Name: Zain Al Abidin 211-6260.

vocabulary = {the, students, opened, their, laptops, books, exams, minds}

Task 1

1 → Skip Gram Target Weight Matrix

The	3	2	3	4
students	0	3	1	4
opened	3	2	0	2
their	1	3	0	2
laptops	2	2	0	0
books	4	2	4	2
exam	2	2	4	3
minds	4	3	4	2

2 → Skip Gram Context Weight Matrix.

The	1	0	1	4
Students	1	1	1	1
opened	2	0	2	2
their	0	1	2	1
laptops	1	4	0	1
books	4	0	4	4
exam	0	1	3	3
minds	1	2	3	0

3 → Average Embeddings

the	2	1	2	4
students	0.5	2	1	2.5
opened	2.5	1	1	2
their	0.5	2	1	1.5
laptops	1.5	3	0	0.5
books	4	1	4	3
exams	1	1.5	3.5	3
mids	2.5	2.5	3.5	1

Task 2

the students opened their ____ ?

the ⇒ [0 0 0 0 3 2 3 4]

$$f_t = \sigma (W_f * I + B_f)$$

$$\sigma \left(\begin{bmatrix} 48 \\ 37 \\ 50 \\ 28 \end{bmatrix} \right) = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$O_t = \sigma(W_O * I + B_O)$$

$$\sigma\left(\begin{bmatrix}31\\26\\32\\60\end{bmatrix}\right) = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}$$

$$i_t = \sigma(W_i * I + B_i)$$

$$\sigma\left(\begin{bmatrix}17\\47\\22\\32\end{bmatrix}\right) = \begin{bmatrix}0.99\\1\\0.99\\1\end{bmatrix}$$

$$\tilde{C_t} = \sigma(W_c * I + b_c)$$

$$\sigma\left(\begin{bmatrix}29\\42\\35\\53\end{bmatrix}\right) = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}$$

$$C_t = f_t \times C_{t-1} + i_t * \tilde{C_t}$$

$$\begin{bmatrix}1\\1\\1\\1\end{bmatrix} \times \begin{bmatrix}0\\0\\0\\0\end{bmatrix} + \begin{bmatrix}0.99\\1\\0.99\\1\end{bmatrix} \times \begin{bmatrix}1\\1\\1\\1\end{bmatrix} \Rightarrow \begin{bmatrix}0.99\\1\\0.99\\1\end{bmatrix}$$

$$h_t = O_t * C_t$$

$$\Rightarrow \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 0.99 \\ 1 \\ 0.99 \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 0.99 \\ 1 \\ 0.99 \\ 1 \end{bmatrix}$$

Students = $\begin{bmatrix} 0.99 & 1 & 0.99 & 1 & 0 & 3 & 1 & 4 \end{bmatrix}$

$$f_t = \sigma\left(W_f * I + B_f\right)$$

$$= \sigma\left(\begin{bmatrix} 57.96 \\ 30.95 \\ 47.5 \\ 25.9 \end{bmatrix}\right) \Rightarrow \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$i_t = \sigma\left(W_i * I + B_i\right)$$

$$\sigma\left(\begin{bmatrix} 34.8 \\ 44.9 \\ 26.9 \\ 26.9 \end{bmatrix}\right) \Rightarrow \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$O_t = \sigma\left(W_o * I + B_o\right)$$

$$\sigma\left(\begin{bmatrix} 49.9 \\ 21.4 \\ 30.4 \\ 56.9 \end{bmatrix}\right) \Rightarrow \begin{bmatrix} 1 \\ 0.99 \\ 1 \\ 1 \end{bmatrix}$$

$$\tilde{C_t} = \tanh(W_c * I + b_c)$$

$$\sigma\left(\begin{bmatrix} 34.9 \\ 49.9 \\ 43.9 \\ 43.9 \end{bmatrix}\right) \Rightarrow \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C_t}$$

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 0.99 \\ 1 \\ 0.99 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1.99 \\ 2 \\ 2.99 \\ 2 \end{bmatrix}$$

$$h_t = O_t * C_t \Rightarrow \begin{bmatrix} 1 \\ 0.99 \\ 1 \\ 1 \end{bmatrix} \times \begin{bmatrix} 1.99 \\ 2 \\ 2.99 \\ 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1.99 \\ 1.98 \\ 2.99 \\ 2 \end{bmatrix}$$

Opened = [1.99 1.98 1.99 2 3 2 0 2]

$$f_t = \sigma\left(W_f * I + B_f\right)$$

$$\sigma\left(\begin{bmatrix} 47.9 \\ 29.9 \\ 48.9 \\ 28.9 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$i_t = \sigma(W_i * I + B_i)$$

$$\sigma\left(\begin{bmatrix} 46.7 \\ 45.4 \\ 26.9 \\ 47.8 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \\ 0.99 \\ 1 \end{bmatrix}$$

$$O_t = \sigma(W_o * I + B_o)$$

$$\sigma\left(\begin{bmatrix} 58.8 \\ 40.9 \\ 31.9 \\ 63.8 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$C_t^{\sim} = \sigma(W_c * I + B_c)$$

$$\sigma\left(\begin{bmatrix} 50.84 \\ 43.93 \\ 37.4 \\ 56.9 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$C_t = f_t \times C_{t-1} + i_t * C_t^{\sim}$$

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1.99 \\ 2 \\ 1.99 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2.99 \\ 3 \\ 2.99 \\ 3 \end{bmatrix}$$

$$h_t = O_t * C_t$$

$$\begin{bmatrix}1\\1\\1\\1\end{bmatrix}\begin{bmatrix}2.99\\3\\2.99\\3\end{bmatrix} = \begin{bmatrix}2.99\\3\\2.99\\3\end{bmatrix}$$

their = $\begin{bmatrix}2.99 & 3 & 2.99 & 3 & 1 & 3 & 0 & 2\end{bmatrix}$

$$f_t = \sigma\left(W_f * I + B_f\right)$$

$$\sigma\left(\begin{bmatrix}55.4\\58.5\\52.5\\64.3\end{bmatrix}\right) = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}$$

$$i_t = \sigma\left(W_i * I + B_i\right)$$

$$\sigma\left(\begin{bmatrix}65.4\\55.7\\37.3\\54.3\end{bmatrix}\right) = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}$$

$$O_t = \sigma\left(W_o * I + B_o\right)$$

$$\sigma\left(\begin{bmatrix}74.6\\39.7\\35.5\\72.4\end{bmatrix}\right) = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}$$

$$\tilde{C_t} = \sigma\left(W_c * I + b_c\right)$$

$$\sigma\left(\begin{bmatrix} 55.4 \\ 58.5 \\ 52.5 \\ 69.3 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$C_t = f_t * C_{t-1} \quad i_t * \tilde{C_t}$$

$$\begin{bmatrix} 3.99 \\ 4 \\ 3.99 \\ 4 \end{bmatrix}$$

$$h_t = O_t \times C_t$$

$$= \begin{bmatrix} 3.99 \\ 4 \\ 3.99 \\ 4 \end{bmatrix}$$

Output layer → Softmax

$$\begin{bmatrix} 58.64 \\ 37.99 \\ 59.7 \\ 39.8 \\ 58.8 \\ 43.8 \\ 66.8 \\ 55.7 \end{bmatrix} = \begin{bmatrix} 2.6 \times 10^{-4} \\ 2.56 \times 10^{-13} \\ 8.8 \times 10^{-4} \\ 1.84 \times 10^{-12} \\ 3.35 \times 10^{-4} \\ 1.05 \times 10^{-10} \\ 0.99 \\ 1.6 \times 10^{-5} \end{bmatrix}$$

The students opened their exams

Now using context weight matrix

The $[0\ 0\ 0\ 0\ 1\ 0\ 1\ 4]$

$$f_t = \sigma\begin{bmatrix}28\\27\\30\\14\end{bmatrix} = \begin{bmatrix}1\\1\\1\\0.99\end{bmatrix} \qquad i_t = \sigma\begin{bmatrix}5\\31\\11\\14\end{bmatrix} = \begin{bmatrix}0.99\\1\\0.99\\0.99\end{bmatrix}$$

$$O_t = \sigma\begin{bmatrix}26\\18\\23\\40\end{bmatrix} = \begin{bmatrix}1\\0.99\\0.99\\1\end{bmatrix} \qquad C_t^{\sim} = \sigma\begin{bmatrix}17\\29\\19\\21\end{bmatrix} = \begin{bmatrix}0.99\\1\\0.99\\0.99\end{bmatrix}$$

$$C_t = \begin{bmatrix}0.99\\1\\0.99\\0.99\end{bmatrix} \qquad h_t = \begin{bmatrix}0.99\\0.99\\0.98\\0.99\end{bmatrix}$$

Students $[0.99\ \ 0.99\ \ 0.98\ \ 0.99\ \ 1\ \ 1\ \ 1\ \ 1]$

$$f_t = \sigma\begin{bmatrix}27.8\\17.9\\23.9\\18.8\end{bmatrix} = \begin{bmatrix}1\\1\\1\\1\end{bmatrix} \qquad i_t = \sigma\begin{bmatrix}22.9\\26.8\\16.9\\28.8\end{bmatrix} = \begin{bmatrix}0.99\\0.99\\0.99\\1\end{bmatrix}$$

$$O_t = \sigma\begin{pmatrix}32.8\\20.8\\15.8\\37.8\end{pmatrix} = \begin{bmatrix}1\\0.99\\0.99\\1\end{bmatrix} \qquad \tilde{C_t} = \sigma\begin{pmatrix}22.8\\25.8\\21.8\\32.8\end{pmatrix} = \begin{bmatrix}1\\1\\0.99\\1\end{bmatrix}$$

$$C_t = \begin{bmatrix}4.8\\5\\4.8\\4.9\end{bmatrix} \qquad h_t = \begin{bmatrix}4.7\\4.9\\4.7\\4.9\end{bmatrix}$$

their = [1 1 1 1 0 1 2 1]

$$f_t = \sigma\begin{bmatrix}32\\19\\21\\23\end{bmatrix} = \begin{bmatrix}1\\0.99\\0.99\\1\end{bmatrix} \qquad i_t = \sigma\begin{bmatrix}5\\31\\11\\14\end{bmatrix} = \begin{bmatrix}0.99\\1\\0.99\\0.99\end{bmatrix}$$

$$O_t = \sigma\begin{bmatrix}34\\17\\15\\36\end{bmatrix} = \begin{bmatrix}1\\0.99\\0.99\\1\end{bmatrix} \qquad \tilde{C_t} = \sigma\begin{bmatrix}18\\19\\25\\34\end{bmatrix} = \begin{bmatrix}0.99\\1\\1\\1\end{bmatrix}$$

$$C_t = \begin{bmatrix}5.79\\5.94\\5.74\\5.84\end{bmatrix} \qquad h_t = \begin{bmatrix}5.68\\5.88\\5.68\\5.78\end{bmatrix}$$

Output layer ⇒ Softmax.

$$\Rightarrow \begin{bmatrix} 82.9 \\ 53.8 \\ 86.3 \\ 57.1 \\ 83.8 \\ 63.3 \\ 95.2 \\ 80.7 \end{bmatrix} = \begin{bmatrix} 4.27 \times 10^{-6} \\ 3.2 \times 10^{-18} \\ 2.52 \times 10^{-3} \\ 5.25 \times 10^{-16} \\ 1.25 \times 10^{-5} \\ 2.8 \times 10^{-13} \\ 0.99 \\ 9.32 \times 10^{-6} \end{bmatrix}$$

hence exams

Now using average weight matrix.

$$\text{the} = [0\ 0\ 0\ 0\ 2\ 1\ 2\ 4]$$

$$F_t = \sigma\begin{bmatrix} 38 \\ 32 \\ 40 \\ 21 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0.99 \end{bmatrix} \qquad i_t = \sigma\begin{bmatrix} 17 \\ 39 \\ 17 \\ 23 \end{bmatrix} = \begin{bmatrix} 0.99 \\ 1 \\ 0.99 \\ 0.99 \end{bmatrix}$$

$$O_t = \sigma\begin{bmatrix} 31 \\ 24 \\ 27 \\ 51 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.99 \\ 1 \\ 1 \end{bmatrix} \qquad C_t^{\sim} = \sigma\begin{bmatrix} 26 \\ 35 \\ 27 \\ 37 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$C_t = \begin{bmatrix} 0.99 \\ 1 \\ 0.99 \\ 0.99 \end{bmatrix} \qquad h_t = \begin{bmatrix} 0.99 \\ 0.99 \\ 0.99 \\ 0.99 \end{bmatrix}$$

$$\underline{\text{Students}} = \begin{bmatrix} 0.99 & 0.99 & 0.99 & 0.99 & 0.5 & 2 & 1 & 25 \end{bmatrix}$$

$$f_t = \sigma \begin{pmatrix} 42.8 \\ 24.4 \\ 35.9 \\ 22.4 \end{pmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0.99 \end{bmatrix} \qquad i_t = \sigma \begin{pmatrix} 28.8 \\ 35.8 \\ 21.9 \\ 27.8 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0.99 \\ 1 \end{pmatrix}$$

$$O_t = \sigma \begin{pmatrix} 41.3 \\ 21.4 \\ 23.4 \\ 47.3 \end{pmatrix} = \begin{bmatrix} 1 \\ 0.99 \\ 0.99 \\ 1 \end{bmatrix} \qquad \tilde{C_t} = \sigma \begin{pmatrix} 28.8 \\ 37.8 \\ 32.9 \\ 38.3 \end{pmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$C_t = \begin{pmatrix} 1.99 \\ 2 \\ 1.98 \\ 1.98 \end{pmatrix} \qquad h_t = \begin{bmatrix} 1.99 \\ 1.98 \\ 1.96 \\ 1.98 \end{bmatrix}$$

$$\underline{\text{Opened}} = \begin{bmatrix} 1.99 & 1.98 & 1.96 & 1.98 & 2.5 & 1 & 1 & 2 \end{bmatrix}$$

$$f_t = \sigma\begin{pmatrix}45.7\\31.9\\42.8\\32.2\end{pmatrix} = \begin{pmatrix}1\\1\\1\\1\end{pmatrix} \qquad i_t = \sigma\begin{pmatrix}41.1\\48.7\\27.8\\48.7\end{pmatrix} = \begin{pmatrix}1\\1\\1\\1\end{pmatrix}$$

$$O_t = \sigma\begin{pmatrix}55.1\\39.3\\30.8\\64.1\end{pmatrix} = \begin{pmatrix}1\\1\\1\\1\end{pmatrix} \qquad \tilde{C_t} = \sigma\begin{pmatrix}44.1\\43.6\\35.7\\55.1\end{pmatrix} = \begin{pmatrix}1\\1\\1\\1\end{pmatrix}$$

$$C_t = \begin{pmatrix}2\\2\\2\\2\end{pmatrix} \qquad h_t = \begin{pmatrix}2\\2\\2\\2\end{pmatrix}$$

their = $\begin{bmatrix}2 & 2 & 2 & 2 & 0.5 & 2 & 1 & 1.5\end{bmatrix}$

$$f_t = \sigma\begin{pmatrix}49\\25.5\\37\\30.5\end{pmatrix} = \begin{pmatrix}1\\1\\1\\1\end{pmatrix} \qquad i_t = \sigma\begin{pmatrix}44\\43\\29\\42\end{pmatrix} = \begin{pmatrix}1\\1\\1\\1\end{pmatrix}$$

$$O_t = \sigma\begin{pmatrix}54.5\\28.3\\25.5\\54.5\end{pmatrix} = \begin{pmatrix}1\\1\\1\\1\end{pmatrix} \qquad \tilde{C_t} = \sigma\begin{pmatrix}37\\44\\39\\49.5\end{pmatrix} = \begin{pmatrix}1\\1\\1\\1\end{pmatrix}$$

$C_t = \begin{bmatrix} 3 \\ 3 \\ 3 \\ 3 \end{bmatrix}$ $h_t = \begin{bmatrix} 3 \\ 3 \\ 3 \\ 3 \end{bmatrix}$

Output layer = Softmax.

→ Softmax $\begin{bmatrix} 45 \\ 29 \\ 45 \\ 30 \\ 45 \\ 33 \\ 51 \\ 42 \end{bmatrix} = \begin{bmatrix} 2.4 \times 10^{-3} \\ 2.7 \times 10^{-10} \\ 2.4 \times 10^{-3} \\ 7.5 \times 10^{-10} \\ 2.4 \times 10^{-3} \\ 1.51 \times 10^{-8} \\ 0.99 \\ 1.22 \times 10^{-4} \end{bmatrix}$ Hence exams.