## National University of Computer and Emerging Sciences, Lahore Campus

| | | | | |
|---|---|---|---|---|
| | Course: | Natural Language Processing | Course Code: | CS 4063 |
| | Program: | BS(Data Science) | Semester: | Fall 2024 |
| | Duration: | 30 Minutes | Total Marks: | 10 |
| | Paper Date: | 6-May-2024 | Weight | |
| | Section: | 8A | Page(s): | 2 |
| | Exam: | Quiz 4 | | |

Q1) Given the following weight matrices (Wq, Wk, Wv) and embedding vectors (x1, x2, x3), calculate output of self-attention layer (z1 ,z2, and z3). [5 Marks]

$$Wq = \begin{bmatrix} 0 & 0.7 & 1 \\ 1 & 2 & 0.3 \\ 0.5 & 1 & 1 \end{bmatrix}, \quad Wk = \begin{bmatrix} 1.8 & 1 & 1 \\ 1 & 0.5 & 0.7 \\ 0.2 & 1.5 & 0.9 \end{bmatrix}, \quad Wv = \begin{bmatrix} 1 & 1.3 & 0.4 \\ 2 & 1 & 2 \\ 1 & 1.5 & 0.2 \end{bmatrix}$$

$$x1 = [1 \quad 0.3 \quad 0.4], \quad x2 = [1.5 \quad 0.5 \quad 1], \quad x3 = [0.3 \quad 1 \quad 0.8]$$

Q2) What advantages does the transformer model offer over the attention based (biLSTM) encoder decoder model? [3 Marks]

Q3) Write the attention equations for LSTM based encoder decoder model. At time stamp t, encoder hidden state is represented as ht, decoder hidden state as st [2 Marks]

**Department of Computer Science**

# National University of Computer and Emerging Sciences, Lahore Campus

| | | | |
|---|---|---|---|
| Course: | Natural Language Processing | Course Code: | CS 4063 |
| Program: | BS(Data Science) | Semester: | Fall 2024 |
| Duration: | 20 Minutes | Total Marks: | 7 |
| Paper Date: | 2 April-2024 | Weight | |
| Section: | 8B | Page(s): | 2 |
| Exam: | LSTM Quiz  v2 | | |

We've thoroughly practiced employing LSTM (Long Short-Term Memory) in our previous assignment to forecast upcoming work tasks. The current objective involves computing values for the below given tasks.

1. Compute embedding from the given target weight matrix based on One Hot vector: [0 1 0 0]
2. Compute value for forget gate from the data given below.
3. Compute $C_t$ & $h_t$ value from all supporting values given below.
4. Write Equations for finding $C_t$ & $h_t$.

**Target Weight Matrix:**

| | | | |
|---|---|---|---|
| 1 | 1 | 3 | 4 |
| 2 | 3 | 3 | 4 |
| 4 | 1 | 1 | 0 |
| 5 | 0 | 2 | 4 |

**Weight Matrix for Input Gate:**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 6 | 4 | 3 | 1 | 5 | 6 | 2 | 0 |
| 0 | 0 | 6 | 6 | 1 | 3 | 3 | 5 |
| 6 | 4 | 2 | 3 | 4 | 5 | 1 | 2 |
| 6 | 6 | 2 | 6 | 1 | 4 | 0 | 0 |

**Bias for Input Gate:**

| |
|---|
| 4 |
| 4 |
| 2 |
| 1 |

| Forget Gate | Input Gate | Output Gate | $h_{t-1}$ | $c_{t-1}$ | $\tilde{c_t}$ |
|---|---|---|---|---|---|
| 1 | | 1 | 0.76 | 1. | 1 |
| 1 | | 1 | 0.76 | 1 | 1 |
| 1 | | 1 | 0.76 | 1 | 1 |
| 1 | | 1 | 0.76 | 0.99 | 1 |

**Solution: (Show Steps)**

# National University of Computer and Emerging Sciences, Lahore Campus

| Course: | Natural Language Processing | Course Code: | CS 4063 |
|---|---|---|---|
| Program: | BS(Data Science) | Semester: | Fall 2024 |
| Duration: | 20 Minutes | Total Marks: | 7 |
| Paper Date: | 1 April-2024 | Weight | |
| Section: | 8A | Page(s): | 2 |
| Exam: | LSTM Quiz v1 | | |

Q) We've thoroughly practiced employing LSTM (Long Short-Term Memory) in our previous assignment to forecast forthcoming work tasks. The current objective involves computing values for the candidate cell state ($\tilde{c_t}$) and hidden state ($h_t$) and cell unit ($c_t$) at the next timestamp, using the provided prior information. [10]

Weight Matrix Values:
Weights and Bais for Forget Gate
[[6 0 0]
[5 5 6 0]]

[-.
[0]]
Weights and Bais for Input Gate
[[5 1 2 2]
[6 3 2 1]]

[0]
[0]]
Weights and Bais for Update Gate
[[0 4 0 5]
[2 6 2 5]]

[1]
[1]]
Weights and Bais for Output Gate
[[4 3 4 4]
[2 2 0 4]]

[-.
[-.]]

Time = 2
Privous Hidden State (ht):
[[0.4165792 ]
[0.32134238]]
Privous Cell State (ct):
[[0.44354576]
[1.99999732]]
Input :
[[0]
[2]]
Values for Forget Gate:
[[0.33209835]
[0.9014788 ]]
Values for Input Gate: .
[[0.99999057]
[0.333263 ]]
Values for Output Gate:
[[0.33333224]
[0.99998962]]

Solution: