

Part I

	Total # of parameters in RNN/LSTM layer(s)	Test accuracy %
2-level bi-LSTM	1,446,400	73.00%
2-level bi-RNN	361,600	61.39%
bi-RNN	120,800	68.72%
bi-LSTM	483,200	86.49%
LSTM	241,600	86.03%
RNN	60,400	64.64%

Part II

Question 2

$$\frac{\partial L^t}{\partial f_a} = 0, \frac{\partial L^t}{\partial f_p} = 0, \frac{\partial L^t}{\partial f_n} = 0$$

Question 3

$$\frac{\partial L^t}{\partial f_a} = f_n - f_p, \frac{\partial L^t}{\partial f_p} = f_p - f_a, \frac{\partial L^t}{\partial f_n} = f_a - f_n$$

Question 4

If we are considering all permutations then we have 5 choices for the anchor, 4 choices for the positive and 3 choices for the negative.

Which gives: $(6 P 3) = 120$ unique triplets.

If we only consider valid triplets. Then we have 6 choices for the anchor, 2 choices for a positive and 3 choices for a negative. Alternatively, we can see this as 3 choices for an anchor, 2 choices for a positive and 3 choices for a negative times 2.

This gives: $6 * 2 * 3 = 2 * (3 * 2 * 3) = 36$ unique triplets.

Question 5

First converting the list of triplets to the embedding space we get:

$$t_1=(f_1, f_2, f_5), t_2=(f_1, f_3, f_4), t_3=(f_2, f_1, f_6), t_4=(f_2, f_3, f_4), \\ t_5=(f_3, f_1, f_5), t_6=(f_5, f_6, f_2), t_7=(f_6, f_4, f_3), t_8=(f_6, f_4, f_2)$$

Question 5.1

$$\frac{\partial L}{\partial f_1} t_1 = f_5 - f_2, \frac{\partial L}{\partial f_2} = f_2 - f_1, \frac{\partial L}{\partial f_3} = 0, \frac{\partial L}{\partial f_4} = 0, \frac{\partial L}{\partial f_5} = f_1 - f_2, \frac{\partial L}{\partial f_6} = 0$$

Question 5.2

$$\frac{\partial L}{\partial f_1} t_2 = f_4 - f_3, \frac{\partial L}{\partial f_2} = 0, \frac{\partial L}{\partial f_3} = f_3 - f_1, \frac{\partial L}{\partial f_4} = f_1 - f_4, \frac{\partial L}{\partial f_5} = 0, \frac{\partial L}{\partial f_6} = 0$$

Question 5.3

$$\frac{\partial L}{\partial f_1} t_3 = f_1 - f_2, \frac{\partial L}{\partial f_2} = f_6 - f_1, \frac{\partial L}{\partial f_3} = 0, \frac{\partial L}{\partial f_4} = 0, \frac{\partial L}{\partial f_5} = 0, \frac{\partial L}{\partial f_6} = f_2 - f_6$$

Question 5.4

$$\frac{\partial L}{\partial f_1} t_4 = 0, \frac{\partial L}{\partial f_2} = f_4 - f_2, \frac{\partial L}{\partial f_3} = f_3 - f_2, \frac{\partial L}{\partial f_4} = f_2 - f_4, \frac{\partial L}{\partial f_5} = 0, \frac{\partial L}{\partial f_6} = 0$$

Question 5.5

$$\frac{\partial L}{\partial f_1} t_5 = f_1 - f_3, \frac{\partial L}{\partial f_2} = 0, \frac{\partial L}{\partial f_3} = f_5 - f_1, \frac{\partial L}{\partial f_4} = 0, \frac{\partial L}{\partial f_5} = f_3 - f_5, \frac{\partial L}{\partial f_6} = 0$$

Question 5.6

$$\frac{\partial L}{\partial f_1} t_6 = 0, \frac{\partial L}{\partial f_2} = f_5 - f_2, \frac{\partial L}{\partial f_3} = 0, \frac{\partial L}{\partial f_4} = 0, \frac{\partial L}{\partial f_5} = f_2 - f_6, \frac{\partial L}{\partial f_6} = f_6 - f_5$$

Question 5.7

$$\frac{\partial L}{\partial f_1} t_7 = 0, \frac{\partial L}{\partial f_2} = 0, \frac{\partial L}{\partial f_3} = f_6 - f_3, \frac{\partial L}{\partial f_4} = f_4 - f_6, \frac{\partial L}{\partial f_5} = 0, \frac{\partial L}{\partial f_6} = f_3 - f_4$$

Question 5.8

$$\frac{\partial L}{\partial f_1} t_8 = 0, \frac{\partial L}{\partial f_2} = f_6 - f_2, \frac{\partial L}{\partial f_3} = 0, \frac{\partial L}{\partial f_4} = f_4 - f_6, \frac{\partial L}{\partial f_5} = 0, \frac{\partial L}{\partial f_6} = f_2 - f_4$$

Question 6

$$\frac{\partial L}{\partial f_1} = (f_5 - f_2) + (f_4 - f_3) + (f_1 - f_2) + (f_1 - f_3) = f_5 - 2f_2 + f_4 - 2f_3 + 2f_1$$

$$\frac{\partial L}{\partial f_2} = (f_2 - f_1) + (f_6 - f_1) + (f_4 - f_2) + (f_5 - f_2) + (f_6 - f_2) = -2f_1 - 2f_2 + f_4 + f_5 + 2f_6$$

$$\frac{\partial L}{\partial f_3} = (f_3 - f_1) + (f_3 - f_2) + (f_5 - f_1) + (f_6 - f_3) = -2f_1 - f_2 + f_3 + f_5 + f_6$$

$$\frac{\partial L}{\partial f_4} = (f_1 - f_4) + (f_2 - f_4) + (f_4 - f_6) + (f_4 - f_6) = f_1 + f_2 - 2f_6$$

$$\frac{\partial L}{\partial f_5} = (f_1 - f_2) + (f_3 - f_5) + (f_2 - f_6) = f_1 + f_3 - f_5 - f_6$$

$$\frac{\partial L}{\partial f_6} = (f_2 - f_6) + (f_6 - f_5) + (f_3 - f_4) + (f_2 - f_4) = 2f_2 + f_3 - 2f_4 - f_5$$