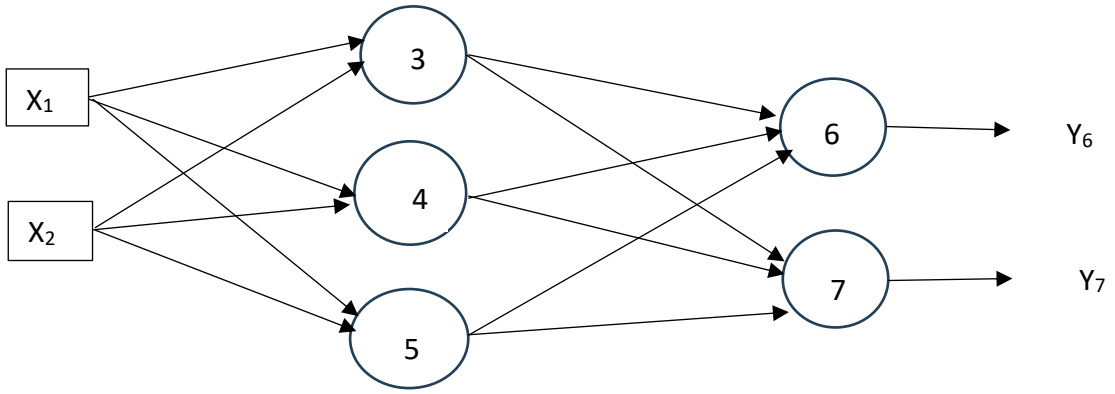
CSE 427 Assignment 2 (Questions 2,3,4,9)
Submission deadline: The day you will sit for the final exam

1. Derive the expression for updating W_{13} and W_{56} in the feed forward neural network shown below using gradient descent approach. All the neurons are using sigmoid activation function.



2. From the neural network given in question 1, find the updated weights of W13, W23, W14, W24, b3, b4 and b5 if all the initial weights and biases are 0.5 each. Assume that the input is x1 = 1 and x2 = 1 and the target output is Y5 = 0, Y6 =1 and Y7= 0. The learning rate is 0.1. Assume that all the neurons are using sigmoid function

3. From the following data given in the table, draw the ROC curve. What should be the probability threshold you will set to achieve 100% true positive rate? What should be the probability threshold you will set to achieve 0% false positive?

Threshold = 1	Yes	No
Yes	0	6
No	0	4
Threshold = 0.975		
Yes	1	5
No	0	4
Threshold = 0.975		
Yes	1	5
No	0	4
Threshold = 0.965		
Yes	2	4
No	0	4

CSE 427 Assignment 2 (Questions 2,3,4,9)
Submission deadline: The day you will sit for the final exam

Threshold = 0.95		
Yes	3	3
No	0	4
Threshold = 0.87		
Yes	3	3
No	1	3
Threshold = 0.75		
Yes	4	2
No	1	3
Threshold = 0.50		
Yes	4	2
No	1	3
Threshold = 0.20		
Yes	6	0
O	1	3
Threshold = 0.15		
Yes	6	0
No	2	2
Threshold = 0.10		
Yes	6	0
No	3	1
Threshold = 0.05		
Yes	6	0
No	4	0
Threshold = 0.00		
Yes	6	0
No	4	0

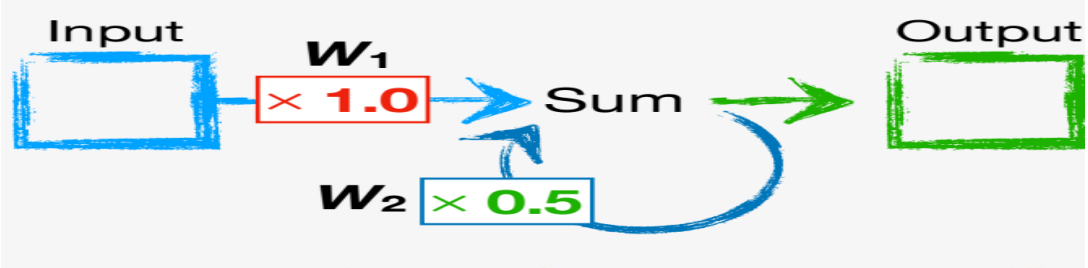
4.

From the confusion matrix given below, find the false positive and false negative for each class. Also find the sensitivity, specificity, precision and recall for each class

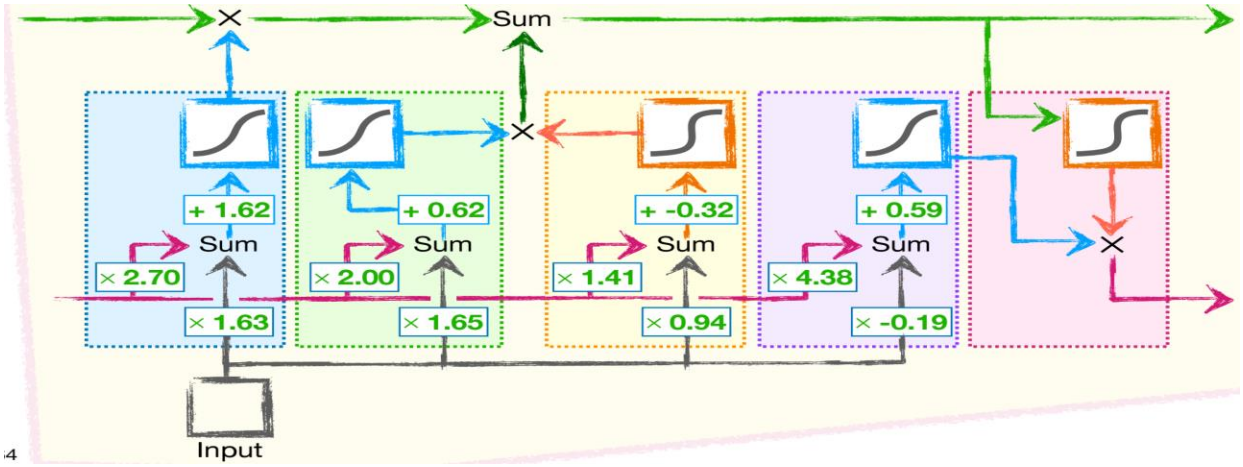
	Predicted		
	Class A	Class B	Class C
Class A (Actual)	800	100	100
Class B (Actual)	50	900	50
Class C (Actual)	200	100	700

CSE 427 Assignment 2 (Questions 2,3,4,9)
Submission deadline: The day you will sit for the final exam

5. An image (201x201x3) is convolved with 100 filters (7x7x3) with stride = 2 and produces 98x98x100 outputs. Represent this convolution operation by two matrix multiplication. Explain what each row and column of the matrices represent. How many parameters will be needed to perform the convolution operation?
6. Find the derivative of the loss function with respect to w1 of the given recurrent network if it is unfolded 5 times to predict the price of the sixth day. Now explain what is the gradient vanishing/exploding problem with a recurrent neural network.



7. If the input for day1 and day2 are 1.0 and 0.5 respectively in the following LSTM cells, what will be the output for day3?



8. Show examples using schematic diagram the application of RNN for many to one and one to many case

CSE 427 Assignment 2 (Questions 2,3,4,9)
Submission deadline: The day you will sit for the final exam

9. The embedding values for “We” and “Go” are [1.87, 0.09] and [-.78,0.27] respectively in the inputs of an encoder. Let the query, key and value matrices for encoder are as follows:

$Q = \begin{matrix} 1.1 & 0.6 \\ -2.8 & 2.4 \end{matrix}$ $K = \begin{matrix} -1.7 & 0.5 \\ -1.5 & 0.9 \end{matrix}$ V = $\begin{matrix} 1.5 & -1.0 \\ -0.3 & -0.2 \end{matrix}$

Find the embedding values of “We” and “Go” after adding positional encoding and after considering attention between “We” and “Go” Now assume that after adding positional encoding the embedding values for [EOS] is [2.7,-1.34]. Let the query, key and value matrices for cross attention are

$Q = \begin{matrix} 1.5 & -0.3 \\ 0.3 & -1.0 \end{matrix}$ $K = \begin{matrix} -1.1 & 0.3 \\ -1.5 & -0.8 \end{matrix}$ V = $\begin{matrix} 1.1 & 0.6 \\ -1.2 & -0.5 \end{matrix}$

Find the embedding values for [EOS] after considering cross attention between [EOS] and the encoder