PES University, Bengaluru

(Established under Karnataka Act 16 of 2013)

END SEMESTER ASSESSMENT (ESA) - JULY - 2023

UE19CS343 - Topics in Deep Learning

Total Marks : 100.0

1.a. Explain the following activation functions with neat diagrams and advantages,disadvantages.
a. Sigmoid
b. Relu (8.0 Marks)

1.b. i) To implement XOR gate with 8 inputs, how many minimum number of hidden layers are required? How many number of perceptrons in hidden layer are required ?
ii) Tensorflows are implemented using computational graphs. Give the data flow graphical representation of equation y = Wx + b. (4.0 Marks)

1.c. Explain the following optimizers and mention the challenges of all gradient descent variants.
1. Gradient Descent
2. Stochastic Gradient Descent
3. MiniBatch Gradient Descent

(8.0 Marks)

2.a. What is the need of SVM? Differentiate hard margin and soft margin classification. (7.0 Marks)

2.b. Suppose we want to classify 4 letters **A B C D** using one v/s one method, How many classifiers are required? What is the disadvantage of one v/s one multiclass classifier? How it is overcome? Illustrate neatly with diagram. (7.0 Marks)

2.c. Consider a support vector machine whose input space is R^2, and in which the inner products are computed by means of the kernel

$K(x.y) = (x.y + 1)^2 - 1$

x.y denotes the ordinary inner product in R^2.

Show that the mapping to feature space that is implicitly defined by this kenel is the mapping to R^5 given by

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \rightarrow \phi(\mathbf{x}) = \begin{bmatrix} x_1^2 \\ x_2^2 \\ \sqrt{2}\, x_1 x_2 \\ \sqrt{2}\, x_1 \\ \sqrt{2}\, x_2 \end{bmatrix}.$$

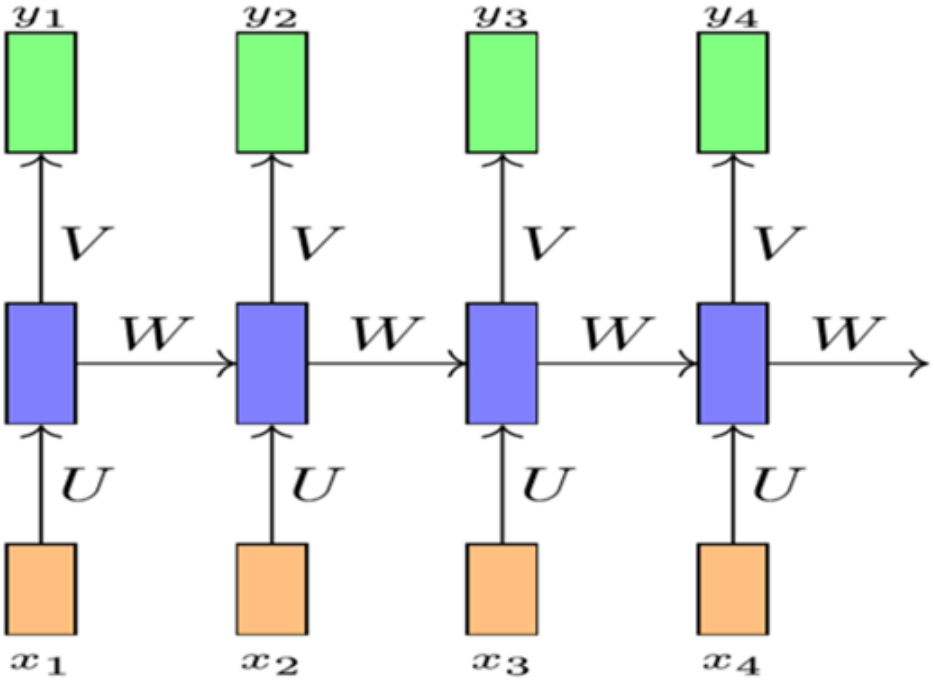
Also show for Ø(y). (6.0 Marks)

3.a. Explain the various architecture of RNN with an example each.
There are four types :

1. One to One Architecture
2. One to Many Architecture
3. Many to One Architecture
4. Many to Many Architecture

(8.0 Marks)

3.b. Consider the following vanilla Recurrent Neural Network. How do we train this network?
Give the complete derivation for updating the weight between hidden layer and output layer.
Consider the activation function and loss function as softmax and cross entropy respectively.



(8.0 Marks)

3.c. A RNN takes input of words

- each as a vector of length **'n'**
- one hidden layer with **n/2** neurons
- one output layer with **2** neuron
- if total number of weights =161, (exclude bias)
- find n

(4.0 Marks)

4.a. What is the need of Convolution Neural Network over Feed Forward Neural Network? Explain the following
a. Convolution layer
b. Pooling layer (5.0 Marks)

4.b. Let us consider a Convolutional Neural Network having three different convolutional layers in its architecture as –
Layer-1: Filter Size – 3 X 3, Number of Filters – 10, Stride – 1, Padding – 0
Layer-2: Filter Size – 5 X 5, Number of Filters – 20, Stride – 2, Padding – 0
Layer-3: Filter Size – 5 X5 , Number of Filters – 40, Stride – 2, Padding – 0
If we give the input a 3-D image to the network of dimension 39 X 39, then determine the dimension of the vector after passing through a fully connected layer in the architecture.
Find the number of parameters in each layer. (7.0 Marks)

4.c. Apply the convolution operation on the following image using the below 3 X 3 filter, stride = 1 and no padding.
(5.0 Marks)

3	0	1	2	7	4
1	5	8	9	3	1
2	7	2	5	1	3
0	1	3	1	7	8
4	2	1	6	2	8
2	4	5	2	3	9

6 X 6 image

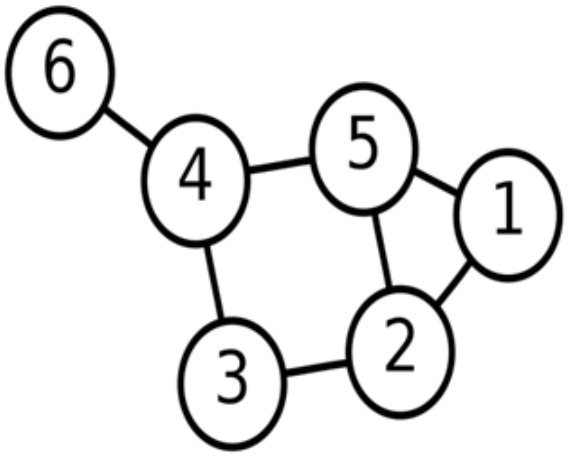
*

1	0	-1
1	0	-1
1	0	-1

3 X 3 filter

4.d. What is the need of transfer learning? Mention its different types. (3.0 Marks)

5.a. What is the role of laplacian matrix in Graphical Neural Network? Find the Laplacian matrix of the following graph (4.0 Marks)



5.b. Convolutional Neural Networks cannot be directly applied on Graphs'. Provide 2 reasons to substantiate the above statement. (4.0 Marks)

5.c. Explain the working of a GNN using Neural Message Passing. (6.0 Marks)

5.d. Give the computational graph for each node of the following graph (6.0 Marks)