Roll No.												

PRESIDENCY UNIVERSITY
BENGALURU

School of Computer Science & Engineering
Mid - Term Examinations - November 2024

Semester: VII | **Date**: 05-11-2024

Course Code: CSE3010 | **Time**: 02.00pm to 03.30pm

Course Name: Deep Learning Techniques | **Max Marks**: 50

Program: B.Tech. (Computer Science & Engineering) | **Weightage**: 25%

Instructions:

(i) Read all questions carefully and answer accordingly.

(ii) Do not write anything on the question paper other than roll number.

Part A

Answer ALL the Questions. Each question carries 2marks. 5Qx2M=10M

1	State true or false. Deep Learning is so named because it has no hidden layers.	2 Marks	L1	CO1
2	Name any 1 technique of automatic hyperparameter tuning.	2 Marks	L1	CO1
3	Name any activation function which is used for multi-class classification.	2 Marks	L1	CO2
4	Name the activation function which takes a number as input and returns a probability as the output.	2 Marks	L1	CO2
5	State the range for the tanh activation function	2 Marks	L1	CO2

Part B

Answer ALL Questions. Each question carries 10 marks. 4QX10M=40M

6	**6a.**	Expand CNN	1 Marks	L1	CO1
	6b.	State true or false. Gradient descent is a method of measuring loss in multiclass classification.	1 Marks	L1	CO1
	6c.	Consider the multi-layer feed forward neural network with inputs x_1, x_2, and x_3. The network has 3 hidden layer neurons, h_1, h_2, and h_3, with weight vectors W_{h1} = [0.5, -0.5, 0.5], W_{h2} = [0.75, 0.25, -0.5], and W_{h3} = [0.25, 0.5, 0.75], and bias for the first hidden layer is 0.5 and bias for output layer is 0.10. Assume that	8 Marks	L3	CO1

		the neurons have a Tanh Activation Function at the output layer, and a ReLU activation function at the hidden layer. Predict the value of the output of the neural network. Also, calculate the Mean Absolute Error if the actual value is 0.5.			
		Or			
	7a.	Expand RNN.	1 Marks	L1	CO1
	7b.	State true or false. CNNs are good for applications like Computer Vision and Image Processing.	1 Marks	L1	CO1
7	**7c.**	Consider the multi-layer feed forward neural network with inputs x_1, x_2, and x_3. The network has 3 hidden layer neurons, h_1, h_2, and h_3, with weight vectors W_{h1} = [0.5, -0.5, 0.5], W_{h2} = [0.75, 0.25, -0.5], and W_{h3} = [0.25, 0.5, 0.75], and bias for the first hidden layer is 0.5 and bias for output layer is 0.10. Assume that the neurons have a Sigmoid Activation Function at the output layer, and a ReLU activation function at the hidden layer. Predict the value of the output of the neural network. Also, calculate the Mean Squared Error if the actual value is 0.75.	8 Marks	L3	CO1
8	**8a.**	Name the problem caused by the ReLU activation function which creates ***dead neurons*** that will never get activated.	2 Marks	L1	CO1
	8b.	Recall the sigmoid activation function and calculate its derivative.	3 Marks	L2	CO1
	8c.	Given the loss function L(a,b,c) = a*(b+2c). Calculate the gradients of the loss function L with each of the inputs a, b, and c, when a = 3, b = 1, and c =2.	5 Marks	L2	CO1
		Or			
9	**9a.**	Write down the range of the Leaky ReLU activation function.	2 Marks	L1	CO1
	9b.	Recall the tanh activation function and calculate its derivative.	3 Marks	L2	CO1
	9c.	Given the loss function L(a,b,c) = a*(b+2c). Calculate the gradients of the loss function L with each of the inputs a, b, and c, when a = 3, b = 1, and c =2.	5 Marks	L2	CO1
10	**10a.**	Name the hyperparameter which determines how fast the gradient descent algorithm will execute.	2 Marks	L1	CO2
	10b.	Consider a feedforward neural network with 2 Boolean inputs, X_1 and X_2 and a single output, Y. Let it have weights W_1 = m and W_2 = n, and a bias b = 0. Given that m and n are positive, and the neuron fires if the value of **Z >= 0**, which Boolean function is represented by the network. Prove it using a truth table.	4 Marks	L2	CO2
	10c.	Consider the following python code: import keras from keras.models import Sequential from keras.layers import Dense model = Sequential() model.add(Dense(12, input_dim=**XXX**, activation='relu')) model.add(Dense(10, activation='relu))	4 Marks	L3	CO2

```
model.add(Dense(10, activation='softmax'))
print(model.summary())
```

Compute the value of **XXX** so that the number of trainable parameters is 340. Assume a non-zero bias (i.e. every neuron takes a bias term as well).

Or

11a.	Name a technique for automatic parameter tuning if the number of hyperparameters are small (Eg. less than 3).	2 Marks	L1	CO2
11b.	Consider a feedforward neural network with 2 Boolean inputs, X_1 and X_2 and a single output, Y. Let it have weights W_1 = m and W_2 = n, and a bias b = 0. Given that m and n are positive, and the neuron fires if the value of **Z > 0**, which Boolean function is represented by the network. Prove it using a truth table.	4 Marks	L2	CO2

11c.	Consider the following python code:
	
	`import keras`
	`from keras.models import Sequential`
	`from keras.layers import Dense`
	`model = Sequential()`
	`model.add(Dense(12, input_dim=8, activation='relu'))`
	`model.add(Dense(9, activation='relu))`
	`model.add(Dense(1, activation='softmax'))`
	`print(model.summary())`
	
	As we can see, the model takes an 8 dimension input to the Input layer, which is connected to a Dense layer of 9 neurons, which is connected to an output layer with a sigmoid activation function. Compute the number of trainable parameters in the system. Assume a non-zero bias (i.e. every neuron takes a bias term as well).
		5 Marks	L3	CO2

11				
12a.	Write down the formula of the Exponential Linear Unit (ELU).	1 Marks	L1	CO2
12b.	Let f(x) be the ELU of x. Let f'(x) be the derivative of f(x) with respect to x. Prove that f'(x) = 1 if x > 0 and f'(x) = f(x) + a, if x is less than 0, where a is the parametric constant used in calculating the ELU.	1 Marks	L2	CO2
12c.	Consider a multilayer perceptron which has an input layer, 3 hidden layers and an output layer with a single neuron. The input layer has 4 neurons, the first hidden layer has 3 neurons, the second hidden layer has 4 neurons, and the third hidden layer also has 3 neurons. The activation functions for the hidden layers are all ReLU and the activation function for the output layer is the sigmoid activation function. The weight matrices are as follows: • W1 (bias is the rightmost column):	8 Marks	L3	CO2
12				

2	0	-1	1	1
0	1	1	0	1
1	-1	0	0	0

- W2 (bias is the rightmost column):

0	-1	2	1
-1	1	0	2
0	1	1	1
-1	0	0	3

- W3 (bias is the rightmost column):

0	-1	-1	2	5
-1	1	1	0	2
-1	0	1	0	3

Output Layer Weights = [-1, 1, 1] and output layer bias = -2
Predict the output, given that the input X = [2,1,2,1]

Or

13a. Define Regularization. 1 Marks L1 CO2

13b. A grid search has to select the best values for 3 hyperparameters, H_1, H_2, and H_3, where H_1 has V_1 values, H_2 has V_2 values, and H_3 has V_3 values. Compute the number of iterations that grid search has to run. 1 Marks L2 CO2

13c. Consider a multilayer perceptron which has an input layer, 3 hidden layers and an output layer with a single neuron. The input layer has 4 neurons, the first hidden layer has 3 neurons, the second hidden layer has 4 neurons, and the third hidden layer also has 3 neurons. The activation functions for the hidden layers are all ReLU and the activation function for the output layer is the sigmoid activation function. The weight matrices are as follows: 8 Marks L3 CO2

13

- W1 (bias is the rightmost column):

2	0	-1	1	1
0	1	1	0	1
1	-1	0	0	0

- W2 (bias is the rightmost column):

0	-1	2	1
-1	1	0	2
0	1	1	1
-1	0	0	3

- W3 (bias is the rightmost column):

0	-1	-1	2	5
-1	1	1	0	2
-1	0	1	0	3

Output Layer Weights = [-1, 1, 1] and output layer bias = -2
Predict the output, given that the input X = [-1,1,1,2]