Code No :

$\boxed{\text{R20}}$

III B. TECH II SEMESTER REGULAR EXAMINATION MODEL PAPER
DEEP LEARNING
(COMMON TO CSM BRANCHES)

Time : 3 Hours   **Max. Marks : 70**

Note : Answer **ONE** question from each unit **(5 × 14 = 70 Marks)**

~~~~~~

		UNIT-I		CO	BL
1.	a)	Explain the need for nonlinear activation functions in neural networks. How the ReLU activation function solves the XOR problem?	[7M]	CO1	L2
	b)	What is the choice of output layer units in the task of multiclass classification and why?	[7M]	CO1	L2
		(OR)			
2.	a)	How a linear neuron is converted into non-linear neuron? Explain why it is required in practice.	[7M]	CO1	L2
	b)	How the cost functions are defined for various machine learning task performed using deep neural networks?	[7M]	CO1	L1
		UNIT-II			
3.	a)	What is meant by the regularization of a deep learning model? What different strategies are used for regularization?	[7M]	CO2	L3
	b)	Explain in detail the process of ridge regularization and how it reduces the generalization error.	[7M]	CO2	L1
		(OR)			
4.	a)	Distinguish Adagrad optimization from the basic gradient descent optimization.	[7M]	CO2	L3
	b)	What is the chain rule of calculus and how is it useful in a backpropagation algorithm?	[7M]	CO2	L1
		UNIT-III			
5.	a)	How CNNs are different from traditional neural networks with all fully connected layers?	[7M]	CO3	L2
	b)	What is the context between the depth and filter size in building the CNN models	[7M]	CO3	L3
		(OR)			
6.	a)	How important it is to have skip connections in the ResNet blocks? Explain.	[7M]	CO3	L4
	b)	The input consists of 7 channels of size 16x16. Use convolution later. Find the number of weights in each of the configurations below.	[7M]	CO3	L4

		(a) The output of the first layer consists of 8 feature maps, and the filters are of size 5x5. The convolution is done with stride 2 and zero padding is used. (b) Now suppose we made this a fully connected layer, but the number of input and output units are kept the same as in the network described in the Part (a) above.			
UNIT-IV					
7.	a)	How the overfitting can be avoided using drop out strategy?	[7M]	CO4	L1
	b)	In which cases transfer learning is useful and how transfer learning is carried out?	[7M]	CO4	L3
(OR)					
8.	a)	What are different normalization techniques used in CNNs? Explain.	[7M]	CO4	L1
	b)	List out various training strategies for the better performance of the neural network models.	[7M]	CO4	L3
UNIT-V					
9.	a)	What are different computational representations of a language? Explain in detail.	[7M]	CO5	L1
	b)	With a suitable example explain the process of word2vec model.	[7M]	CO5	L2
(OR)					
10.	a)	Draw a depth 1 simple recurrent network for language translation with inputs $(x_1, x_2, ..., x_n)$ and the outputs $(y_1, y_2, ..., y_n)$. Explain how it works.	[7M]	CO5	L2
	b)	In the LSTM model, explain exactly how the cell state is updated from C_{t-1} to C_t, using the previous state h_{t-1} and the current input x_t.	[7M]	CO5	L3

* * * * *

THE ABOVE MODEL PAPER ATTAINMENTS OF BLOOM'S TEXONOMY AS FOLLOWS

L1: 6*7 = 42 = 30%
L2: 6*7 = 42 = 30%
L3: 6*7 = 35 = 30%
L4: 2*7 = 14 = 10%