Hall Ticket Number:

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | |

**Code No. : 16245 (B) N**

# VASAVI COLLEGE OF ENGINEERING *(AUTONOMOUS)*, HYDERABAD

*Accredited by NAAC with A++ Grade*

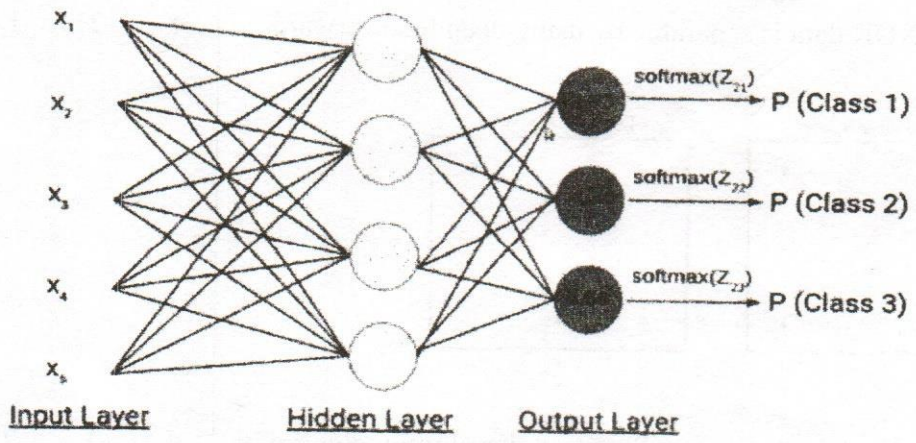
## B.E. (CSE-AIML) VI-Semester Main Examinations, May/June-2023

## Deep Learning

Time: **3 hours** Max. Marks: **60**

*Note: Answer all questions from **Part-A** and any **FIVE** from **Part-B***

***Part-A (10× 2 = 20 Marks)***

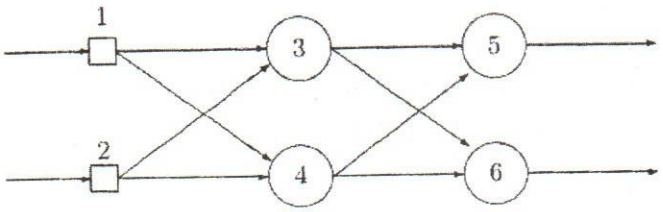
| Q. No. | Stem of the question | M | L | CO | PO |
|---|---|---|---|---|---|
| 1. | What are the multiple learning components in Deep Learning? | 2 | 1 | 1 | 1,2 |
| 2. | Draw the computation graph for Linear Regression with weight decay. | 2 | 3 | 1 | 1,3 |
| 3. | Why do we inject noise to input data in DL? | 2 | 2 | 2 | 1,2 |
| 4. | Which regularization method leads to weight sparsity? Give reasons. | 2 | 3 | 2 | 1,3 |
| 5. | What is the effect of Zero padding? Show with and example. | 2 | 2 | 3 | 1,2 |
| 6. | Suppose you have 5 convolutional kernel of size 7 x 7 with zero padding and stride 1 in the first layer of a convolutional neural network. You pass an input of dimension 228 x 228 x 3 through this layer. What are the dimensions of the data which the next layer will receive? | 2 | 3 | 3 | 1,3 |
| 7. | Consider the below diagram and compute the output using SoftMax and which class has more accuracy?   | 2 | 3 | 4 | 1,2 |
| 8. | Compute the Loss Function for Recurrent Neural Network | 2 | 2 | 4 | 1,3 |
| 9. | How to choose the right optimization algorithm? | 2 | 3 | 5 | 1,2 |
| 10. | What is Generative Adversarial Networks? Explain with an example. | 2 | 1 | 5 | 1,3 |

*Contd... 2*

**Code No. : 16245 (B) N**

***Part-B (5×8 = 40 Marks)***

| | | | | | |
|---|---|---|---|---|---|
| 11. a) | The following diagram represents a feed-forward neural network with one hidden layer: | 4 | 3 | 1 | 1,2 |



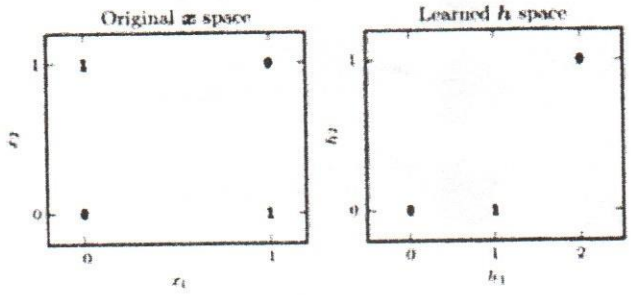

Weight on connection between nodes i and j is denoted by wij , such as w13 is the weight on the connection between nodes 1 and 3. The following table lists all the weights in the network:

| | |
|---|---|
| $w_{13} = -2$ | $w_{35} = 1$ |
| $w_{23} = 3$ | $w_{45} = -1$ |
| $w_{14} = 4$ | $w_{36} = -1$ |
| $w_{24} = -1$ | $w_{46} = 1$ |

Each of the nodes 3, 4, 5 and 6 uses the following activation function: $\phi(v)$ = 1 if $v \geq 0$ or 0 otherwise. where v denotes the weighted sum of a node. Each of the input nodes (1 and 2) can only receive binary values (either 0 or 1). Calculate the output of the network (y5 and y6) for each of the input patterns:

| Pattern: | $P_1$ | $P_2$ | $P_3$ | $P_4$ |
|---|---|---|---|---|
| Node 1: | 0 | 1 | 0 | 1 |
| Node 2: | 0 | 0 | 1 | 1 |

| | | | | | |
|---|---|---|---|---|---|
| b) | Prove that non-linear XOR data is separable by using deep feed forward network? | 4 | 3 | 1 | 1,3 |



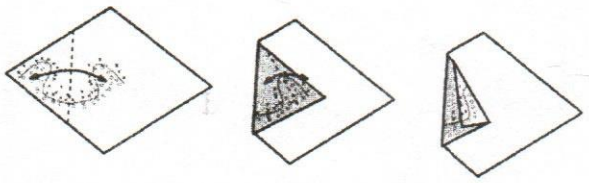

| | | | | | |
|---|---|---|---|---|---|
| 12. a) | What is output Units? Explain the Bernoulli output distribution along with cost function used. | 4 | 1 | 2 | 1,2 |
| b) | What is Lagrange formulation? How is it used in Constrained optimization? | 4 | 2 | 2 | 1,3 |
| 13. a) | Why parameter sharing utilized in CNN? Give reasons with neat diagram. | 4 | 2 | 3 | 1,2 |

*Contd... 3*

| | | | | | |
|---|---|---|---|---|---|
| b) | Consider the convolutional neural network defined by the layers in the left column below. Fill in the shape of the output volume and the number of parameters at each layer. You can write the shapes in the numpy format (e.g. (128,128,3)).<br>Notation:<br>• CONV5-N denotes a convolutional layer with N filters with height and width equal to 5. Padding is 2, and stride is 1.<br>• POOL2 denotes a 2x2 max-pooling layer with stride of 2 and 0 padding.<br>• FC-N denotes a fully-connected layer with N neurons | 4 | 3 | 3 | 1,3 |

| Layer | Activation Volume Dimensions | Number of parameters |
|---|---|---|
| Input | 32×32 × 1 | 0 |
| CONV5-10 | | |
| POOL2 | | |
| CONV5-10 | | |
| POOL2 | | |
| FC10 | | |

| | | | | | |
|---|---|---|---|---|---|
| 14. a) | What are the different design patterns of RNN? Explain with neat diagram. | 4 | 1 | 4 | 1,2 |
| b) | Construct the Encoder- decoder Sequence to sequence modelling architecture of a variable length sequence | 4 | 3 | 4 | 1,3 |
| 15. a) | Write about Adagrad algorithm optimizer? Explain with Neat Diagram | 4 | 2 | 5 | 1,2 |
| b) | Differentiate Adam and RMSProp Optimizers | 4 | 3 | 5 | 1,3 |
| 16. a) | Explain the Exponential advantage of deeper network by using below diagram. | 4 | 3 | 1 | 1,2 |



Figure

| | | | | | |
|---|---|---|---|---|---|
| b) | What is Data augmentation? Explain with diagram its usage. | 4 | 2 | 2 | 1,3 |
| 17. | Answer any ***two*** of the following: | | | | |
| a) | Compare the locally connected layers, tiled convolution, and standard convolution with neat diagram. | 4 | 3 | 3 | 1,3 |
| b) | Explain usage of Bidirectional RNN with functions used in the application. | 4 | 1 | 4 | 1,2,3 |
| c) | Write an Algorithm for stochastic Gradient Descent with Momentum | 4 | 2 | 5 | 1,2 |

M : Marks; L: Bloom's Taxonomy Level; CO; Course Outcome; PO: Programme Outcome

| | | |
|---|---|---|
| i) | Blooms Taxonomy Level – 1 | 20% |
| ii) | Blooms Taxonomy Level – 2 | 35% |
| iii) | Blooms Taxonomy Level – 3 & 4 | 45% |

*****