This question paper contains 4 printed pages]

Roll No. | | | | | | | | | | |

S. No. of Question Paper : **8075**

Unique Paper Code : **2343010017**

Name of the Paper : **Deep Learning**

Name of the Course : **B.Sc. (H) (Computer Science)**

Semester : **VI**

Duration : **3 Hours** Maximum Marks : **90**

(Write your Roll No. on the top immediately on receipt of this question paper.)

Question No. **1** (Section A) is compulsory.

Attempt any *four* questions from Section B.

Parts of a question should be attempted together.

Use of scientific calculator is allowed.

Section A

1. (*a*) Explain the internal architecture of an LSTM unit with equations for the input, forget and output gates. 6

(*b*) Explain RMSProp and ADAM optimizer. Also give *two* advantages and *two* disadvantages of each. 6

(*c*) How does the back propagation algorithm work ? Also, explain the chain rule of calculus with respect to back propagation. 4

(*d*) Why simple perceptron is used as linear classifier ? Write the formula for MSE and Cross Entropy Error function. 4

P.T.O.

(e) Illustrate the concept of feature maps in CNNs. How are filters used to extract features from input images ? 4

(f) Describe dropout regularization in neural networks. 3

(g) What is an RNN and how is it different from a feed forward neural network ? 3

Section B

2. (a) For a simple RNN architecture given : 9

$w_x = 0.4$, $w_h = 0.2$, $w_y = 1.0$, $h_0 = 0$

Input Sequence x = [1, 2], Activation : tanh

(i) Calculate the hidden states h1 and h2

(ii) Calculate the outputs y1 and y2

(iii) If targets is y = [0.5, 1], what is the total loss using MSE ?

(b) Explain how LSTM addresses the problem of vanishing/exploding gradients in RNN ? 4

(c) For deep neural network with LSTM layer of 100 units and input vector of dimension 32, calculate parameters involved. 2

3. (a) Draw RNN architecture for time step 3. 6

(b) Explain Sigmoid, ReLU and Tanh activation functions. 6

(c) What are the output values of the Sigmoid, ReLU, and Tanh activation functions when the input x = 4 ? 3

4. (a) What is Batch Normalization ? Explain steps of batch normalization. List advantages of Batch Normalization. 6

(*b*) Given an input of size 64 × 64 × 3, compute the output size after applying : 6

(*i*) a convolutional layer with 16 filters of size 5 × 5, stride 1, padding 2, followed by

(*ii*) a 2 × 2 max pooling layer with stride 2.

(*c*) Describe any *three* optimization challenges in deep learning. 3

5. (*a*) Differentiate between sparse and denoising auto encoders. 8

(*b*) Construct the following simple auto encoder for the following. Also represent diagrammatically. 7

```
Input Vector : x = [1, 2, 3]

Encoder Weights : w = [0.2, 0.4, 0.6], bias : b = 0.5

Decoder Weights : w = [0.1, 0.5, 1], bias : b = 0.1
```

(*i*) Calculate the encoded representation of the input.

(*ii*) Reconstruct the input using the decoder.

(*iii*) Calculate loss (MSE).

6. Write short notes on any *five* of the following : 15

(*i*) Global Contrast Normalization

(*ii*) Hierarchical Softmax

(*iii*) Generative Adversarial Network

(*iv*) Softmax function

(*v*) Ill-Conditioned Hessian Matrix

(*vi*) LSTM architecture

(*vii*) Types of Recurrent Neural Networks

(*viii*) Graphics Processing Unit.

7. (*a*) An input grey scale image of dimension 5 × 5 is given as follows : 9

```
[ [1, 3, 2, 1, 4],
[2, 0, 1, 2, 0],
[3, 4, 1, 1, 2],
[1, 1, 0, 3, 1],
[2, 0, 3, 2, 1] ]
```

Perform convolution operation with the following kernel of dimension 3 × 3 with stride = 1 and no padding.

```
Kernel = [ [0, 1, 0],
           [1, 0, 1],
           [0, 1, 0] ]
```

(*b*) On the above obtained matrix, apply ReLU activation function and max pooling (2 × 2) with stride 1. 6