[This question paper contains 2 printed pages.]

Your Roll No.....

K

Sr. No. of Question Paper : 9942

Unique Paper Code : 2343010017

Name of the Paper : Deep Learning (DSE)

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

Semester : VII

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Question No. **1 (Section A)** is compulsory.
3. Attempt any **4 (four)** questions from **Section B**.
4. All parts of a question should be attempted together.
5. Use of scientific calculators is allowed.

Section A

Q1 (a) What are the two optimization algorithms used in deep learning? (2)

(b) Differentiate between Recurrent Neural Network (RNN) and feed forward neural network. (3)

(c) What is Cross-Entropy loss? Why is preferred over MSE? (3)

(d) A perceptron receives inputs x1 = 1, x2 = 3 with weights w1 = 0.6, w2 = - 0.2 and bias = - 0.1. Compute the output using a sigmoid activation function. (3)

(e) Explain why a single perceptron is unable to solve XOR problem. How is this problem solved? (4)

(f) What is Batch Normalization? Explain how it stabilizes the training process. (5)

(g) Explain contrast normalization and its role in computer vision. List its types. (5)

(h) Explain the concept of parameter sharing in CNNs with an example. How does it reduce model complexity? (5)

Section B

Q2 (a) What are word embeddings? Explain their role in efficient training of neural language models. (3)

(b) How are Autoencoders used for dimensionality reduction? Explain using their architecture. (6)

(c) Describe any three challenges in Neural Network Optimization. (6)

Q3 (a) State two advantages and two disadvantages of RNNs. (4)

(b) What is the significance of learning rate in optimization of deep learning? Explain how adaptive learning is achieved in AdaGrad and RMSprop. (5)

(c) Which activation function is most suitable for the output layer neuron(s) in the following cases? Justify your choice briefly. (6)

- Detecting whether an email is *spam* or *not spam*.
- Predicting the *price of a house* based on its features.

P.T.O.

- Classifying a medical image into one of the categories: *Normal*, *Pneumonia*, *Tuberculosis*, or *COVID*.

Q4 (a) Describe the role of convolutional layers in feature extraction. (10)
Compute the dimension of resulting feature map if we perform a convolution on a 5×5 image using a 3×3 kernel (stride=1, padding=1). What will be the size of output if a max pooling 2x2 is applied?

(b) What are Generative Adversarial Networks (GANs)? How are the generator and discriminator networks trained? (5)

Q5 (a) Differentiate between local minima and saddle point in optimization. (4)

(b) What is Drop Connect? How is it different from Dropout? (4)

(c) Consider a linear model $\hat{Y}=wx$; for a training example with x=1 and y=4 the squared error cost function is: (7)

$J(w)=\frac{1}{2}(y-\hat{Y})^2$

The parameter is optimized using AdaGrad optimizer with the following hyperparameters:

- Learning rate: $\eta=0.1$
- Initial accumulated squared gradient: G=0
- Small constant: $\epsilon=10^{-8}$

Given the gradients observed over three iterations:

g1=0.2, g2=0.5, g3=0.1

Using the AdaGrad update rule compute the gradients g and the parameter w for 2 iteration.

Q6 (a) Why is data augmentation used as a preprocessing step in computer vision applications? Describe any one real-world applications where computer vision plays a significant role. (3)

(b) What purpose does regularization solve in neural network training? Differentiate between L1 and L2 regularization. (6)

(c) Consider the following corpus: (6)

"The water of Walden Pond is so beautifully blue that visitors often stop to admire it.
The water of Walden Pond is so beautifully calm during the early morning."

Compute the following based on corpus given.

(i) bigram counts for P(blue | beautifully).
(ii) bigram probability P(blue | beautifully).
(iii) trigram counts for P(blue | so beautifully).

Q7 (a) Illustrate the architecture of a Bidirectional RNN and explain how it processes information. (5)

(b) With the help of equations, explain the operations inside an LSTM unit. Given an input sequence and initial states, demonstrate one forward-pass computation for an LSTM cell. (10)

(500)