

Sr. No. of Question Paper : 7289 **L**

Unique Paper Code : 2343010017

Name of the Paper : Deep Learning

Name of the Course : **DSE**

Semester : IV/VI/VIII

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. Parts of a question should be attempted together.
5. Use of scientific calculator is allowed.

Section A

- Q1 (a) Explain the folded RNN and the unfolded RNN. Draw the unfolded RNN for 3 time states(previous, current, next). Write an equation for the current state of the simple RNN. 5
- (b) Explain the RMSProp optimization techniques with the help of equations. 5

network? Compute the number of iterations per epoch and the total number of iterations when $epochs = 2$, the number of samples is 200, and the batch size is 50.

- (e) Given a CNN that has an input image of $64 \times 64 \times 3$, the number of filters is 4, the size of the filter is 2×2 , the stride is 2 and the padding is 0. How many trainable parameters does this CNN require? What is the size of the output feature map? 5
- (f) What is the purpose of Regularization in the training of a neural network? Which Regularization is better for the weight sparsity? Justify the answer with the help of equations. 5

Section B

- Q2 (a) Write the equation and graph for the activation functions: Linear, ReLU, and Sigmoid. Give the output of the above activation functions for the input of 8. Explain the importance of non-linearity in a neural network. 10
- (b) Explain the differences between the standard backpropagation and backpropagation through time with a suitable diagram. 5
- Q3 (a) Describe the operation of the XOR logic gate in a two-layer perceptron for the binary inputs (00, 01, 10, and 11), with a suitable diagram. Is the XOR gate linearly separable? Justify your answer. 8
- (b) What is the function of a filter in a CNN? Explain the impact of the number of filters and their size on the performance of the CNN. 4

optimization.

- (b) For a given perceptron, it has 3 inputs as $x_1=0.1$, $x_2=0.2$, $x_3=0.3$ with weights as $w_1=0.3$, $w_2=0.4$, $w_3=0.2$ and bias as -0.1 . Compute the output after applying the tanh activation function. 4
- (c) Is Cross-entropy loss better than MSE? Justify your answer. 3

- Q5 (a) For a given feedforward neural network that has an input layer with 3 neurons, a hidden layer with 2 neurons and an output layer with 2 neurons. Compute the hidden state h_1 , h_2 and output y_1 , y_2 when the input $x_1=1$, $x_2=2$, $x_3=4$ and weight $w_x=0.2$, $w_h=0.3$, $w_y=0.1$ and activation is Sigmoid. Draw the diagram that describes the above feedforward neural network. 9
- (b) Why is CNN used for image processing applications? Write the advantages of the convolution network compared to the fully connected network. 3
- (c) Describe GMM-HMM in speech recognition. 3

- Q6 (a) What is the purpose of gates in LSTM? How many parameters does a deep neural network with an LSTM layer of 50 units and an input vector of size 32 require? 8
- (b) What is an n -gram model? Write the formula used to compute the probability of an n -gram based on $(n-1)$ -gram. 7

For the following sentences, write all the possible combinations of the

- ii) “Self-driving cars use Deep Learning to detect traffic in real time.”

Q7 (a) Write short notes on the following terms:

9

(i) Stochastic Gradient Descent

(ii) Momentum

(iii) Batch Normalization

(b) For a given CNN, consider the input matrix of dimension 5×5 given below: 6

[[3, 0, 1, 3, 1]

[3, 2, 2, 2, 1]

[1, 3, 1, 1, 2]

[2, 2, 1, 0, 1]

[1, 2, 1, 0, 1]]

and kernel of dimension 2×2 given below:

[[-1, 0]

[1, -1]]

Compute the output feature map after performing the convolution operation, and show the computations of each element of the output feature map.