Exam Date & Time: 18-Jun-2024 (02:30 PM - 05:30 PM)

# MANIPAL ACADEMY OF HIGHER EDUCATION

SIXTH SEMESTER B.TECH END SEMESTER MAKEUP EXAMINATIONS, JUNE 2024
## DEEP LEARNING [CSE 3271]

**Marks: 50**                                                                 **Duration: 180 mins.**

**A**

**Answer all the questions.**

Instructions to Candidates: Answer ALL questions Missing data may be suitably assumed

1)    What are flat algorithms? Giving a few examples of flat algorithms, distinguish deep learning models from machine learning models.

(4)

A)

B)    What is the significance of forward and backward pass for a computation graph? Show the implementation of forward pass for the Sigmoid function. Also implement a backward pass that directly computes analytical gradients.

(3)

C)    Assume a fully connected neural network architecture is designed where all activations are Sigmoid. Discuss in detail what happens, if the weights are initialized with large positive numbers and what the problem is referred to as?

(3)

2)    Explain the technique used to compress spatial information with an example and a code-snippet. Apply convolution operation to the following data and show the result.

A)



(4)

B)    Compare the effects of different values of the regularization parameter on the weight updates in L2 regularization. What happens when regularization parameter *lambda* is set to a very small value (close to 0) or a very large value?

(3)

C)    Illustrate how fine-tuning can be achieved in transfer learning with relevant code snippet.

(3)

3)            Consider a neural network with an input layer of 1000 neurons, two hidden layers of 500 neurons each, and an output layer of 10 neurons. The network is trained on a dataset with 50,000 examples.

    A)

         i. Without dropout regularization, the network achieves a training accuracy of 98% but a validation accuracy of only 85%. What does this scenario indicate?

         ii. If dropout regularization is applied to the hidden layers with a dropout rate of 0.5, how many neurons in each hidden layer, on average, will be active during each forward pass?  **(4)**

         iii. Suppose that with dropout regularization, the network's training accuracy decreases to 95%, but its validation accuracy increases to 93%. Calculate the reduction in the generalization gap after applying dropout.

         iv. If the dropout rate is increased to 0.7, how many neurons in each hidden layer, on average, will be active during each forward pass?

    B)         How does data augmentation help in reducing overfitting in deep learning models? Explain with an example. Suppose you have a dataset of 5000 images, and you apply horizontal flipping to each image. How many additional images will you have in your augmented dataset?  **(3)**

    C)         Can early stopping be effective if the maximum number of epochs is set too low? Discuss in detail with a neat, labelled graph marking early stopping checkpoint.  **(3)**

4)           Discuss the number of weight matrices, biases, and their respective shapes in a recurrent neural network (RNN). Explain how these weight matrices and biases are utilized in different components of the RNN architecture.  **(4)**

    A)

    B)         Assess the significance of the forget gate in an LSTM unit. How does it contribute to the network's ability to retain relevant information over long sequences, and how does its tuning impact the model's performance?  **(3)**

    C)         List any two applications of Autoencoders.  **(3)**

5)           Explain the fundamental concept behind autoencoders and their role in unsupervised learning.  **(4)**

    A)

    B)         Compare and contrast denoising autoencoders and variational autoencoders (VAEs) in terms of their architectures and objectives.  **(3)**

    C)         List the factors that impact the quality of reconstruction in an autoencoder.  **(3)**

-----End-----