



JK Lakshmipat University, Jaipur
Department of Computer Science and Engineering
Deep Learning (CS-1218)
Endterm, Date: December 03, 2024

Time: 2.5 hours

Semester VII B.Tech. CSE

Max mark: 35

Please read the following instructions carefully.

- There are 4 questions.
 - **No marks** for providing just expressions/answers unless accompanied with correct justification and/or derivation.
 - In case of any doubt, make your assumption, write it clearly and continue. **Do not call your instructor.**
1. (2+2+2+(1+2) points) Vanishing and exploding gradients are known problems of neural networks. You are provided a feed forward neural network with two hidden layers. Clearly explain by writing equations how the vanishing and exploding gradients may occur in the provided feed forward neural network if the number of layers is increased. At the same time, you are also provided a RNN for three timestamps. Explain how will the vanishing and exploding gradients occur in this RNN if the number of timestamps is increased. Why is it a more serious problem if RNNs are used for modeling text data? (CO1, CO2)
 2. Face recognition (CO2, CO3)
 - (a) (1 point) How would you differentiate between two faces using their representations?
 - (b) (2+2 points) How will you train a neural network to learn differentiating between two different faces using Siamese network and triplet loss?
 - (c) (1+2 points) How will you solve the problem of face verification? Draw the network to solve face verification using the input features which are the output of a Siamese network.
 3. Object detection using sliding windows (CO2, CO4)
 - (a) (1+1 points) What is the difference between image classification with localization and detection? Illustrate through examples.
 - (b) (2+2 points) Suppose you have 3 classes – pedestrian, car and motorcycle. You are supposed to prepare a training dataset for detecting objects. How will you prepare the labels? Write the corresponding loss function using which the neural network should be trained.
 - (c) (2 points) How does the convolution implementation of sliding windows speedup the process?
 4. Object detection using YOLO. In YOLO algorithm, as the name suggests, image is looked at only once to detect all the objects. Assume you have divided your image in a 3×3 grid. (CO1, CO3, CO4)

- (a) (2 points) How will you prepare the training dataset if you have 3 classes – pedestrian, car and motorcycle?
- (b) (2 points) How many bounding boxes will the algorithm output for one input image?
- (c) (2+2 points) How to improve the detection of objects by suppressing bounding boxes that detect the same object?
- (d) (2 points) Briefly explain the need for anchor boxes.