Nirma University

Institute of Technology

Semester End Examination (IR), May 2022

B.Tech. in Computer Science and Engineering, Semester VI

2CSDE61 DEEP LEARNING

Roll / Exam No. | Supervisor's initial with date

Time: 3 Hours | Max. Marks: 100

Instructions:
1. Attempt all questions.
2. Figures to right indicate full marks.
3. Use section-wise separate answer book.
4. Draw neat sketches wherever necessary.
5. Assume suitable data wherever applicable and clearly mention them.
6. CLO_ and BL_ have been mentioned against each question to map it as per Course Objective and Bloom's taxonomy.

SECTION – I

Q 1 — Answer the following: [18]

CLO2 BL1,2 (a) What is dropout? How is it useful in regularization? How is it implemented at training and testing time? Explain in detail. [8]

CLO1 BL2 (b) Differentiate between machine learning and deep learning with a suitable example. [4]

CLO2 BL2 (c) Critically compare transfer learning and domain adaptation. [6]

Q 2 CLO2 BL3 — Consider a following 6 x 6 image and a 3 x 3 convolutional filter. [16]

3	0	1	2	7	4
1	5	8	9	3	1
2	7	2	5	1	3
0	1	3	1	7	8
4	2	1	6	2	8
2	4	5	2	3	9

6 X 6 image

1	0	-1
1	1	-1
1	0	-1

3 X 3 filter

Compute activations (assume Leaky ReLU as the activation function with leakage=0.5) of each neuron in the feature map (referred as fm1, henceforth) resulting from applying the filter shown in the above image. Assume bias = 0.2, no padding and stride = 1. Apply 2 x 2 max-pooling on fm1 and show activations of each neuron in the resultant feature map. Assume stride = 2 and no padding.

OR

Q 2 CLO2 BL3,4 — Discuss how deep neural networks can be used for image captioning. Propose architecture of a deep neural network for this task (use table to depict the architecture precisely). Assume that the dataset you would be working has 50,000, [16]

512 x 512 RGB images. Ground truths are available for these images. Clearly mention the data preparation, loss function of the network and training process with the precise makeup of the training set. In your view, what major changes one will require to bring in if we change the task to video captioning (from image captioning)?

Q 3 Answer the following: [16]

CLO2 BL2 (a) Discuss strided and fractionally-strided convolution with a suitable detailed example for each. [8]

CLO3 BL3,4 (b) Propose an architecture of a deep neural network for image classification with localization. Clearly show the makeup of the ground truth. Discuss regarding a loss function which will be suitable here. [8]

SECTION – II

Q 4 CLO2 BL2,4 (a) Assume a bi-directional simple RNN with 1 hidden layer. Assume 10000 neurons in the input layer, 100 neurons in the hidden layer in each direction and 10000 neurons in the output layer. Write necessary equations demonstrating a complete forward pass. Use standard notations. Also, calculate total number of parameters involved. Don't ignore bias. [10]

Q 4 CLO2 BL2,4 (b) Assume a simple RNN with 2 hidden layers. Assume 10000 neurons in the input layer, 100 neurons in each of the hidden layers and 1 neuron in the output layer. Write necessary equations demonstrating a complete forward pass. Also, calculate total number of parameters involved. Don't ignore bias. [10]

Q 5 CLO2 BL2,4 (a) Can XOR logical gate be modelled/learnt using a perceptron? Justify your answer in detail. Can it be modelled/learnt using multilayer perceptron (multilayer feed forward network)? Justify your answer. [9]

OR

Q 5 CLO2 BL2,4 (a) Why vanilla neural networks are not good at computer vision tasks? Discuss in detail. How CNNs are able to address most of the limitations of vanilla neural network? Justify your answer. [9]

Q 5 CLO2 BL2,4 (b) If an object spans across multiple grid cells in YOLO, how is it handled while preparing the ground truth for training? What is the role of anchor boxes in YOLO? [5]

Q 6 CLO3 BL3,4 Write a detailed pseudo code for generating MNIST like images using generative adversarial network. Generator and discriminator should be convolutional neural networks. [16]