

Total No. of Questions:[06]

Total No. of Printed Pages:[03]

PRN. No.	
----------	--

PAPER CODE	
------------	--

MAY 2022 - ENDSEM EXAM
FINAL YEAR B. TECH. (COMPUTER ENGINEERING)
(SEMESTER - II)
COURSE NAME: PROFESSIONAL ELECTIVE-IV
[ADVANCED MACHINE LEARNING]
COURSE CODE: CSUA40181B
(PATTERN 2018)

Time: [1 Hr]

[Max. Marks: 30]

Instructions to candidates:

- 1) **Answer Q.1 OR Q.2, Q.3 OR Q.4, Q.5 OR Q.6.**
- 2) **Figures to the right indicate full marks.**
- 3) **Use of scientific calculator is allowed.**
- 4) **Use suitable data wherever required.**

Q.1 a) After training a neural network, you observe a large gap between the training accuracy (100%) and the test accuracy (42%). Choose the method to reduce this gap. [4]

Method selection 2 marks
Working 2 marks

b) Justify the need of Gradient Descent in neural network with detailed explanation of its working. [6]

Justification 3 marks
Working 3 marks

OR

Q.2 a) "Mini batch gradient descent is faster than gradient descent." Justify the statement. [4]

Justification 2 marks
Working 2 marks

b) When should one use L1, L2 regularization instead of dropout layer, given that both serve same purpose of reducing overfitting? Justify your answer. [6]

Justification 3 marks

Working 3 marks

- Q.3 a) Explain following terms [4]
i) Deep Dream
ii) Deep Art

Explanation 2 marks for each

- b) Compare various architectures of CNN. [6]

1 mark for each LeNet, AlexNet, ZF-Net, VGGNet, GoogLeNet, ResNet

OR

- Q.4 a) "Convolutional Neural Network (CNN) work better with image data." Give justification. [4]

Justification 2 marks

Working 2 marks

- b) "Given a Convolutional Neural Network having three different convolutional layers in its architecture as – [6]
in Layer 1 - 10 filters of 3X3, stride 1 and no padding,
in layer 2 -20 filters of 5 X 5 with stride 2 and no padding,
in layer 3- 40 filters of 5 X 5 with stride 2 and no padding,
If 39 X 39 3-D image pass as input to this network, then estimate the dimension of the vector after passing through a fully connected layer in the architecture.

1 Marks for each step

- Q.5 a) Distinguish BPTT and truncated BPTT. [4]

1mark for each point

- b) "RNNs work better with text data". Give justification. [6]

Justification 3 marks

Working 3 marks

OR

- Q.6 a) Compare GRU and LSTM architectures. [4]

1mark for each point

- b) Which of the following activation functions can lead to vanishing gradients? [6]
i) ReLU
ii) Tanh
iii) Leaky ReLU

Justify your answer with explanation of its working

Justification 2 marks

Working 2 marks