

# Indian Institute of Technology Tirupati

## Deep Learning (CS5223): Mid Semester

Date: 7<sup>th</sup> October, 2021 (10 - 11:30AM)

Maximum Marks: 40

Instructions: Answer all the questions.

---

1. Consider a Neural network applied to an industrial data with 7 channels and  $16 \times 16$  spatial dimensions. Compute the number of parameters and products in the first layer for each of the following configurations. (You may leave the expression as a product without finding the actual number, e.g.  $5 \cdot 4^2 \cdot 3^2$ ) [2+2=4 Marks]
  - (a) Use a convolution layer with 8 feature maps and filters of size  $5 \times 5$  (use default values for stride and padding)
  - (b) Use a fully-connected layer where the number of input and output units are kept same as part (a).
2. What are different types of pooling layers that are commonly used in CNNs. Describe how the pooling operation happens. State one advantage and one disadvantage of using pooling layer in CNNs. [1+1+2=4 Marks]
3. Plot the learning curves where you depict the (average) training and validation losses against the model complexity. Indicate the overfitting scenario with clear labeling in the figure. [2 Marks]
4. Write pseudo code for the following: [2+2+2=6 Marks]
  - (a) Stochastic Gradient Descent for parameter update
  - (b) Full batch Gradient Descent for parameter update
  - (c) Mini-batch Gradient Descent for parameter update
5. Mention at least two ways for minimizing overfitting in neural networks. Describe clearly why they work. [2+2=4 Marks]
6. Consider a one-layer neural network with 3 input units and 1 output unit. All the units have the sigmoid nonlinearity ( $\sigma(x) = \frac{1}{1+e^{-x}}$ ). Assume bias terms are not there for simplicity. Loss used for optimizing the network is the cross-entropy loss. Derive the equations for a single step of weight update. Clearly identify all the symbols used in the derivation. [6 marks]
7. Consider the binary classification data from Table 1. Perform perceptron training on this data clearly mentioning all the required symbols and steps (until a maximum of 5 epochs). [4 Marks]
8. What are hyper-parameters in the context of neural networks. Give at least 5 examples. How are the values for these assigned? [1+1+1=3 Marks]
9. Contrast convolution layer with fully connected layer. Mention advantages and disadvantages along with the suitable use cases where they are preferred. [1.5+1.5=3 Marks]

sample ( $x_i$ )	label ( $y_i$ )
(1, 1, 1)	+1
(1, 1, 0)	-1
(1, 0, 1)	-1
(0, 1, 1)	-1

Table 1: Sample binary classification dataset.

10. Write the important portion of a neural network training code implemented in PyTorch using the autograd machinery. (Give pseudo code with two nested iterations for epochs and iterations; no need to mention the exact loss function, model architecture, optimizer, etc.) [2 Marks]
11. Autoencoder is a specific neural network architecture configuration. It accepts input, through a series of layers it reduces the input dimension, and from there through another series of layers it reconstructs the input. From the understanding of Machine Learning and neural network concepts, answer the following: [1+1=2 Marks]
  - (a) Autoencoders are supervised or unsupervised technique? Explain your answer.
  - (b) Autoencoders fall into which ML task (classification, regression, or density estimation)? Justify your answer.