



CS551: Introduction to Deep Learning

Mid Semester-Assignment, Spring 2022

IIT Patna

24.02.2022

Full Marks: 40

Time: 24 Hours

Please upload your single pdf file (<10MB) in the following link:

<https://forms.gle/ZcCtUhhkkVCcCSz88>

- Consider the following regression problem. The dataset for the problem is a set of n examples (x_i, y_i) where $i = 1, \dots, n$ where x_i and y_i are real numbers for all i . The difficulty here is that we do not have access to the inputs or outputs directly. Also, we do not even know the number of examples in the dataset. We are, however, able to get a few numbers computed from the data. Let $w^* = [w_0^*, w_1^*]$ be the least squares solution. Alternatively, we can say w^* minimizes the following

$$J(w) = \frac{1}{2n} \sum_{i=1}^n (y_i - w_0 - w_1 x_i)^2$$

For this problem, you can assume that the solution is unique. If w^* is a least square solution. Show that the following two conditions are hold. $[\bar{x}, \bar{y}]$ denote the mean value of $\{x_i\}$ and $\{y_i\}$

A. $\frac{1}{n} \sum_{i=1}^n (y_i - w_0^* - w_1^* x_i)(x_i - \bar{x}) = 0$

B. $\frac{1}{n} \sum_{i=1}^n (y_i - w_0^* - w_1^* x_i)(w_1^* x_i + w_0^*) = 0$ (8 marks)

- Consider m i.i.d. samples, $\{x_1, \dots, x_m\}$ drawn from the distribution $P_\theta(x) = 2\theta x e^{-\theta x^2}$ where θ is a parameter and x is positive real number. Find the maximum likelihood estimator for θ based on these samples. (8 marks)
- Consider maximal margin classifier for the following toy data set which has two features X_1 and X_2 . Y is the target label. (a) Sketch the optimal separating hyperplane, and provide the equation for this hyperplane. [Equation for hyperplane may be derived logically or mathematically.] (b) On your sketch, indicate the margin for the maximal margin hyperplane. (c) Indicate the support vectors for the maximal margin classifier. (4+2+2)

Sl. No.	1	2	3	4	5	6	7
X_1	3	2	4	1	2	4	4
X_2	4	2	4	4	1	3	1
Y	R	R	R	R	B	B	B

- Consider the following points in a 2D plane — $(1, 2), (3, 4), (2, 1), (4, 3)$. The points need to be represented using a lower dimension representation scheme, ie, in one-dimension. You need to apply PCA to find the same. What will be coordinate of these points in the compressed representation scheme? You must show every step clearly. (8 marks)
- Design a neural network that mimics a 3-input XOR gate. Clearly show the weights, activation functions, etc. in the diagram. Also show that for any input your network provides the correct output. For this, you can use a table similar to the following (8 marks)

Inputs			Cols. for internal nodes/nets								Output
x_1	x_2	x_3	h1	
0	1	0	1.5	1