Assignments on Linear Regression #2 (Ref Video Lectures 1-12)

(Timely submission of assignments is essential. Copying/plagiarised submission from others will fetch fail (F) grade on this subject)

1. Annual Revenue data for a company is given as,

Y	2007	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Rev. In billion Rupees	100	95	85	90	92	80	89	84	93	96	97

a) Draw a least square line fitting the data.
b) What is the expected revenue in 2022
c) Analyze expected error in predictions. 10

2. The following table shows the final semester marks obtained by 10 students selected at random.

M L	82	85	93	65	87	71	98	68	84	87
HUR	80	88	96	72	91	80	95	72	89	84

Find least square line fitting the above data using
a) X as independent variables (regression of Y on X)
b) Y as independent variable (regression of X on Y)
c) If a student receives a mark 96 in ML, what is her/his expected marks in HUR.
d) If a student receives 95 in HUR. What is her/his expected marks in ML.
e) After plotting a) and b) what conclusions can you draw? 10

3. Experimental results of pressure (P) for a given mass of gas corresponding to various values of volume (V) is given as:

V	54.3	61.8	72.4	88.7	118.6	194
P	61.2	49.5	37.5	28.4	19.2	10.1

Assume PV^n = const =c
a) Find the parameters n and c
b) Write the equation connecting P and V.
c) Estimate the value of P when V=100

10

4. Find the least square parabola which fits the data

Y= W0 + W1X+ W2 X^2

X	0	1	2	3	4	5	6
Y	2.4	2.1	3.2	5.6	9.3	14.6	21.9

Submission Deadline: **Ref Lecture videos 1-8** : **5-02-2022** **10**

5. Download the COVID -19 data of India for the month of May, 2021 and design a predictor for the number of deaths on a particular day. Hence, predict the number of deaths on April 20, 2021 and June 10th , 2021. Verify your prediction with the actual number of deaths and hence calculate the accuracy of prediction.

Submission Deadline: **10-02-2022.** **20**

[Hints: collect data from my website or from the web for the month of May, 2021. Use 22 days data for training and rest 09 days data for testing. Use both straight line fitting and quadratic curve fitting for predictor design and compare their performance in predicting accuracy]

6. Download the housing price data set of Windsor City of Canada (provided on my website link). Design a housing price predictor taking only floor area (plot size), number of bedrooms, and number of bathrooms into considerations. Out of total 546 data , you may take 70% for designing the predictor and 30% for validating the design. The predictor design should be done using the following methods: :

a) Normal equations with and without regularization and compare their performances in terms of % error in prediction. (only allowed to use NumPy library of Python.no other functions/libraries are allowed)

(Ref Lecture-8 and 12)

Submission deadline : **8th February, 2022**

20