Indian Institute of Technology Madras

Web MTech Industrial AI

ID5002W: Industrial AI Lab

Assignment II: Linear Regression and Diagnostics

Instructions

1. Assignment shall be submitted before the due date. Late submissions will not be entertained. If you cannot submit the assignment due to some reasons, please contact the instructor by email.
2. All the assignments must be the student's own work. The students are encouraged to discuss or consult friends or classmates. However, they have to submit their own work. Any malpractice will be reported to the authorities and actions will be taken as per the IIT Madras rules.
3. If you find the solution in the book or article or on the website, please indicate the reference in the solution.

Problem

You have to work only with the file with your roll number specified on it. Each dataset has 6 features $x_1, x_2, ..., x_6$ and 3 targets y_1, y_2, y_3.

Q1 For the given dataset, fit a linear model for the following cases:

(a) y_1 and x_1 [1 mark]

(b) y_1 and x_2 [1 mark]

(c) Perform **Linear Regression Diagnostics** for the two models [1 mark]

Q2 Fit **Multivariate Linear Regression Models** for the following cases along with diagnostics:

(a) y_1 and x_1, x_2 [1 mark]

(b) y_2 and $x_1, x_2, ..., x_6$ [1 mark]

Q3 Fit **Principal Components Regression** (PCR) and **Partial Least Squares** (PLS) model to y_1, y_2, y_3 and $x_1, x_2, ..., x_6$. [1 +1 marks]

Q4 Find the optimal number of components in both models. Justify your findings. [1 +1 marks]

Q5 Compare the performance of PCR and PLS models. [1 mark]