

Indian Institute of Space Science and Technology – Thiruvananthapuram

MA613 Data Mining & MA633 Data Mining Lab

Assignment-III

Date: 31-08-2023

1. Analyze Data 4 using linear regression
 - (a) Describe the experimental procedures used.
 - (b) Report the values of the parameters of the model.
 - (c) Analyze the performance of the model
2. Analyze Computer hardware data set using linear regression (download data from UCI web repository)
 - (a) Analyze the data with normalization and without normalization.
 - (b) Describe how you applied normalization techniques on training and testing data.
 - (c) Apply random subsampling and k fold cross validation.
 - (d) Assess the performance of the model.
 - (e) Report the values of the hyperparameters and the parameters of the model.
 - (f) Apply batch as well as online optimization algorithms and compare their performance in terms of time and MSE.

Notes

- All the files related with the assignment should be saved in a single folder.
- All the programs should be properly documented.
- Last date of submission: 07-09-2023.
- **As far as assignments are concerned, students are expected to observe academic honesty and integrity. Though the students can collaborate and discuss, copying directly other students' assignment or allowing your own assignment to be copied constitute academic dishonesty and is highly discouraged.**