

Indian Institute of Engineering Science and Technology, Shibpur
B.Tech (CST) 7 th Semester End Semester Examinations, September 2021
Machine Learning, CS721/3

Full Marks: 30

Time: 45 minutes

Answer question 1(any three) and any three from the rest of the questions

1. a) Why do we need a validation set and test set to build a classifier?
b) What do you mean by cross validation?
c) How classification and regression differ?
d) What are the differences between underfitting and overfitting? [3x3]

- 2.a) What is conditional and marginal probability ? explain with example.
b) Suppose the weights of randomly selected American female college students are normally distributed with unknown mean μ and standard deviation σ . A random sample of 10 American female college students yielded the following weights (in pounds):115 122 130 127 149 160 152 138 149 180. [2+5]

Based on the definitions given above, identify the likelihood function and the maximum likelihood estimator of μ , the mean weight of all American female college students. Using the given sample, find a maximum likelihood estimate of μ as well.

3. Describe estimation and prediction of multiple regression. How linear regression is different from multiple regression? [4+3].

4. Explain dependent and independent variables. When we use ridge and polynomial regression? What are outliers? [3+3+1]

- 5a).What are the differences between Lasso Regulation and Ridge Regulation? Explain using loss function.
b)What is the Principal Components and explain properties of the Principal components. [4+3]