

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -2 EXAMINATION- October 2023

M.Tech-I Semester (Data Science)

COURSE CODE (CREDITS): 19B1WCI738

MAX. MARKS: 25

COURSE NAME: Introduction to Statistical Learning

COURSE INSTRUCTORS: HRI

MAX. TIME: 1 Hour 30 Minutes

Note: (a) All questions are compulsory.

(b) Marks are indicated against each question in square brackets.

(c) The candidate is allowed to make Suitable numeric assumptions wherever required for solving problems

Q1. In the following linear regression model **(CO1) [2x2 = 04 Marks]**

$$y = \beta_0 + \beta_1 x_1$$

- (a) Compute the standard error associated with β_0 and β_1 .
- (b) Compute 95% confidence intervals for β_0 and β_1 .

Q2. Attempt any one from the following. **(CO1) [03 Marks]**

How do the t-test and p-value help in describing the significance of a coefficient associated with a predictor?

OR

Among R² statistic and Residual Standard Error (RSE), which is considered a better model accuracy evaluation metric and why?

Q3. The following two models were developed for spam email detection- Model-1 and Model-2. Which model is better and why? **(CO2) [04 Marks]**

Model-1	Actual Spam Email (Yes)	Actual Spam Email (No)	Total
Predicted Spam Email (Yes)	200	300	500
Predicted Spam Email (No)	200	300	500
Total	400	600	1000

Model-2	Actual Spam Email (Yes)	Actual Spam Email (No)	Total
Predicted Spam Email (Yes)	250	200	450
Predicted Spam Email (No)	150	400	550
Total	400	600	1000

Q4. How does bootstrap approach allow us to emulate the process of obtaining new sample sets? Describe the formula to compute the standard error of these bootstrap estimates.

(CO2) [04 Marks]

Q5. Write the best subset selection algorithm that involves identifying a subset of the p predictors that is believed to be related to the response.

(CO2)[04 Marks]

Q6. Describe Ridge Regression with the mathematical explanation. Draw bias, variance and MSE-test w.r.t. the hyperparameter λ used in the shrinkage method Ridge Regression with appropriate explanation.

(CO2) [04 Marks]