Birla Institute of Technology & Science, Pilani
Work-Integrated Learning Programmes Division
First Semester 2019-2020
Comprehensive Examination (Regular)

Course No.	: PCAM* ZC211
Course Title	: REGRESSION
Nature of Exam	: Closed Book
Weightage	: 40%
Duration	: 3 Hours
Date of Exam	: **02/11/2019 (FN)**

No. of Pages =
No. of Questions =

Q1. Suppose there is a uni-variate regression problem and you are given 'N' data points. You have developed polynomial regression models with degree 0, 1, 2, ---,10. After looking at training and testing data points, it is concluded that polynomial of degree 0 will be the best fit. Do you think that whether this is a possible scenario? If so, what can you say about the dependency of target or dependent variable on feature or dependent attributes.

Q2. Write down the procedure to find out the optimal regularization coefficient (i.e., lambda). Assume that you are given 'N' data points and you are trying to fit a polynomial of degree 10. Hint: Divide the data into training, validation and testing data sets.

Q3. Explain the significance of regularization by making use of the example that we discussed in class.

Q4. Do you think that subset generated using forward stepwise selection algorithm will be the best feature subset? Support you answer with appropriate reasoning.

Q5. What are the two techniques to implement regularization for polynomial fitting? What is the difference between these two techniques? Explain the two techniques with all mathematical rigor.

Q6. Can the learning rate (in gradient descent algorithms), eta, be any random value? What are the consequences of choosing random value for eta?

Q6. For a given data set of 10 points, build linear regression model and find out the performance of the model using R^2.

Q7. For a given probability distribution, find the variance of the probability distribution.