MANIPAL INSTITUTE OF TECHNOLOGY

MANIPAL

(A constituent unit of MAHE, Manipal)

II SEMESTER M.C.A. **I** INTERNAL EXAMINATION **FEB** 2020

SUBJECT: **DATA ANALYTICS** - [MCA **4521**]

Date: **10/06**/2020 Max.Marks: 15 Duration: 1 Hour

Instructions to Candidates:

- Answer ALL the questions & missing data may be suitable assumed
- Use of calculators is permitted

Q1. (3)

Differentiate between the following with suitable examples.

i. Eager learners vs. Lazy learners
ii. Supervised Learning vs. Unsupervised learning
iii. Classification tree vs. Regression tree

Q2. (4)

2. The following table shows the relationship between the amount of fertilizer used and the height of a plant.

Fertilizer	10	5	12	17	14	7	15	13	6	8	10	11	16	20	17
Height	0.7	0.4	0.8	1.3	1.1	0.6	1.3	1.1	0.6	0.7	0.7	0.65	1.3	1.5	1.3

i. Calculate a simple linear regression equation using Fertilizer as the descriptor and Height as the response.
ii. Predict the height when fertilizer is 9.3.
iii. Visualize using a scatter plot.

Q3. (4)

. Given the following data:

ROLLNO	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
AGE	5	6	4	7	8	10	12	4	15	20
MARKS	10	8	5	10	12	9	11	6	25	18

i. Group the data into 3 clusters using K-means clustering method. Let the Initial centroids be S3,S5 and S9.
ii. Visualize using a group scatter plot.

Q4. (4)

Using the training data set specified below, classify the following observation (X) using the Naïve Bayesian Classifier method and infer as to whether "X" is likely to get Diabetes.

X: BP = High; Weight = Above average; Family history = Yes; Age= 50+

Blood pressure	Weight	Family history	Age
Average	Above average	Yes	50+
Low	Average	Yes	0–50
High	Above average	No	50+
Average	Above average	Yes	50+
High	Above average	Yes	50+
Average	Above average	Yes	0–50
Low	Below average	Yes	0–50
High	Above average	No	0–50
Low	Below average	No	0–50
Average	Above average	Yes	0–50
High	Average	No	50+
Average	Average	Yes	50+
High	Above average	No	50+
Average	Average	No	0–50
Low	Average	No	50+
Average	Above average	Yes	0–50
High	Average	Yes	50+
Average	Above average	No	0–50
High	Above average	No	50+
High	Average	No	0–50