Registration No: | | | | | | | | | | | **M.TECH**

Total Number of Pages :1

M.TECH 1ST SEMESTER REGULAR EXAMINATIONS, DECEMBER 2018
DATA SCIENCE
Branch: CS, Subject Code: MCSPE1041
(Regulations 2018)

Time: 3 Hours   **Max Marks : 70**   **Question Code:RD18002086**

PART-A (10 X 2=20 Marks)

1. Answer the following questions.
(a) Which are the methods to treat missing value?
(b) Describe Scikit-learn.
(c) Describe different Categories of API.
(d) Define Data exploration.
(e) Differentiate Data balancing and Data shuffling.
(f) Compare between Planar Encoding and Retinal Encoding.
(g) What does Data Collection mean?
(h) What is difference between regression model and estimated regression equation?
(i) Why data has missing values?
(j) What is the impact of outliers on dataset?

PART-B (5 X 10=50 Marks)
Answer any five questions from the following.

2.a) Give three commonly used statistical measures that are used for characterization of data   [5]
    Dispersion.
  b)Discuss how they can be computed efficiently in large databases.   [5]

3. Give a detail description of KNN algorithm and state clearly.
  a) When do we use KNN algorithm?   [5]
  b) How do we choose the factor K?   [5]

4. (a) Discuss linear regression with an example.   [5]
  (b) Describe the working behavior of support vector machine with a diagram.   [5]

5. (a) What are the key terminologies of SVM?   [5]
  (b) Explain Bayes' estimator parametric density estimation.   [5]

6. a) Describe data science applications in healthcare.   [5]
  b) Briefly explain Visual Encoding Variables.   [5]

7. a) Describe the role of python in Data Science.   [5]
  b) Distinguish between classification tree and regression tree.   [5]

8. Write Short notes on any two of the following   [5]
  a) Logistic Regression.   [5]
  b) Data science versus BI

==0==