



Registration No:

--	--	--	--	--	--	--	--	--	--

Total Number of Pages : 01

M.TECH

AR-19

M.TECH 1ST SEMESTER EXAMINATIONS NOV/DEC 2019

CSE, MPECS1041

DATA SCIENCE

Time: 3 Hours

Max Marks : 70

The figures in the right hand margin indicate marks.

PART-A

(10 X 2=20 MARKS)

1. Answer the following questions.

- a. Differentiate between Data Science and Machine Learning.
- b. Describe Logic Regression.
- c. Write any four real world applications of the SVM?
- d. Define Visual analysis.
- e. What do you mean by Decision Tree algorithm?
- f. What are the four assumptions of linear regression (simple linear and multiple)?
- g. What does the cost parameter in the SVM means?
- h. Define Data Cleansing.
- i. Define Data Cleansing.
- j. Define Sampling.

PART-B

(5 X 10=50 MARKS)

Answer any five questions from the following.

- 2. a. Describe briefly about different Data Science toolkit. [5]
b. Give a detail description of SVM algorithm. How to implement SVM in Python? [5]
- 3. a. How does exploratory data analysis differ from summary analysis? [5]
b. Explain Simple Linear regression. [5]
- 4. a. Describe different data encodings used in data science. [5]
b. Describe data science applications in healthcare [5]
- 5. a. Explain machine learning with a suitable example. [5]
b. Illustrate the advantages and disadvantages of Baye's estimator. [5]
- 6. a. Briefly explain Visual Encoding Variables [5]
b. Define about Random Forest. [5]
- 7. a. Describe different data encodings used in data science. [5]
b. What is Naïve Bayes and discussed about it. [5]
- 8. Short Notes [5 x 2=10]
 - a. Decision Tree
 - b. Data science versus Big Data Analytics