

Total No. of Questions : 8]

[Total No. of Printed Pages : 2

Roll No

MCADD-705(2)
M.C.A. (Integrated), VII Semester
Examination, November 2019

Data Analytics

Time : Three Hours

Maximum Marks : 70

- Note :** i) Attempt any five questions.
ii) All questions carry equal marks.

1. a) What do you mean by Randomization Test? Where is it used? Give a suitable example.
b) What is ANOVA (Analysis of Variance)? Explain with the help of a suitable example.

2. a) For the following data find equations of two regression lines:

X:	1	2	3	4	5
Y:	15	25	35	45	55

- b) Write and explain K-Nearest neighbours regression and classification.
3. a) Differentiate Ridge and Lasso regression. Give examples.
b) Differentiate supervised and unsupervised learning techniques with suitable examples.
4. a) What is Support vector machine? What is the role of C in EVM?
b) Give some situations where you will use an SVM over a Random forest machine learning algorithm and vice versa.

5. a) Define Big-data and explain the importance of Big data in social networking.
b) Explain any four Big data analytics applications.
6. a) What do you mean by Associative Rule Mining Challenges for big data analysis?
b) Differentiate clustering and classification of Big data. Give suitable examples.
7. a) What do you mean by active learning? How to create data for analytics through active learning?
b) With the help of a suitable example explain Reinforcement learning.
8. Write short notes on any two of the following:
i) Hypothesis
ii) Hadoop
iii) Map Reduce
iv) Hive

X	1	2	3	4	5
Y	15	25	35	45	55
