

Total No. of Questions : 8]

[Total No. of Printed Pages : 2

Roll No .....

**MCADD-705(2)**  
**M.C.A. (Integrated), VII Semester**  
Examination, May 2024

**Data Analytics**

(Elective - II)

Time : Three Hours

Maximum Marks : 70

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

1. a) Describe in detail about the probability distribution.  
b) Can you identify the different statistical concepts required for Big data?

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) 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.
4. a) Explain supervised and unsupervised learning with examples.  
b) Describe in detail about hypothesis testing?

[2]

5. a) What are the components of Hadoop framework?  
b) Explain how can you manage compute node failures?
6. a) Why does one choose analytical system over conventional system?  
b) How do you define propositional rule learning?
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

\*\*\*\*\*