

M.C.A.(Sem-III)

April-2023

(3 Hours)

[Total Marks: 80]

N.B. : (1) Question No.1 is Compulsory

(2) Attempt any three questions between Question No.2 to 6

- Q 1. A) Write a short note on following (Attempt Any Four). [20]
1. Describe the Challenges of Big Data.
 2. What are various text analysis steps? Explain in detail.
 3. What are the different phases of the Data Analytics Lifecycle? Explain each in detail.
 4. What is linear regression? Explain in detail.
 5. Write a short note on Hadoop architecture.
- Q 2. A) What is the functionality of the explode mapper? Explain in detail with example. [10]
- B) What is clustering? Explain in detail. Explain any two of its applications. [10]
- Q 3. A) Explain Tableau as a visualization tool. [10]
- B) Demonstrate the process of executing a Streaming job on a Hadoop cluster. [10]
- Q4. A) What is the functionality of the identity pattern? Explain in detail with example. [10]
- B) What is Big Data? Write down the characteristics of Big Data Applications. [10]
- Q5. A) What is Data Streaming? How to build Twitter HTTP client? [10]
- B) Describe ACME's Text Analysis Process. [10]
- Q6. A) What is CAP THEORAM? Explain the major components of HBase Architecture. [10]
- B) How to generalize the k-means algorithm? Write a short note on determining the number of clusters. [10]
