Exam Date & Time: 06-May-2024 (02:30 PM - 05:30 PM)

MANIPAL ACADEMY OF HIGHER EDUCATION

SIXTH SEMESTER B.TECH END SEMESTER EXAMINATIONS, APRIL-MAY 2024

BIG DATA MODELLING AND MANAGEMENT [CRA 4055]

Marks: 50   **Duration: 180 mins.**

A

Answer all the questions.

Instructions to Candidates: Answer ALL questions Missing data may be suitably assumed

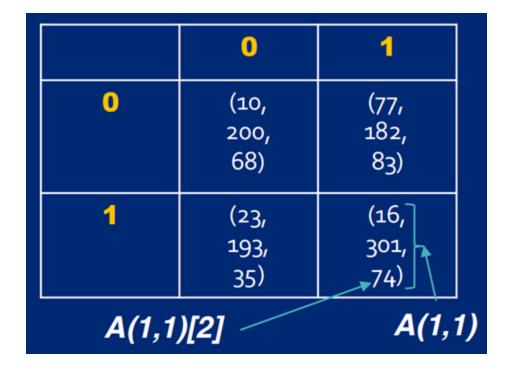|			
1)	A)	How big data allows us to build better models with higher precision results? Discuss some of its applications and show in what way big data can serve human needs.	(4)
	B)	Discuss the role of big data in precision medicine. How big a human genomic data and what would be the storage required to store this data?	(3)
	C)	What is variety in big data? Describe different aspects of data variety related to big data with suitable examples.	(3)
2)	A)	Suppose there is CSV file which contains population of different countries. How can we extract the country having more than 10 million population for the year 2020? Illustrate the steps to be followed to do this.	(4)
	B)	Explain how the distributed file system works and mention its advantages.	(4)
	C)	Identify two problems where Hadoop may not be the best candidate.	(2)
3)	A)	Describe the following concepts related to graph data. i) Path operations ii) Community finding	(4)
	B)	Consider the given array of vectors. Describe the different operations on arrays of vectors.	(4)

	0	**1**
0	(10, 200, 68)	(77, 182, 83)
1	(23, 193, 35)	(16, 301, 74)

A(1,1)[2]   A(1,1)

C) Define a data stream? What is a dynamic steering in case of streaming data.   (2)

4) Interpret the BASE properties in Big data?

                                                                                                  (3)

A)

B) Illustrate scalability in Redis?
                                                                                                  (3)

C) Describe the problems in storing data in file systems.
                                                                                                  (4)

5) Explain how the data storing in a data lake is different from a data warehouse? Analyse the data lake object storage with figure.
                                                                                                  (4)

A)

B) "If a big data system is deployed in the cloud, over multiple machines, security of data becomes an even bigger challenge" justify this statement.   (2)

C) Explain any four desirable characteristics of a Big Data Management System.   (4)

-----End-----