

Roll No.								
-----------------	--	--	--	--	--	--	--	--

Total No. of Pages : 02

Total No. of Questions : 09

B.Tech. (IT) (Sem.-6)

BIG DATA

Subject Code : BTIT-601-18

M.Code : 79623

Date of Examination : 29-05-2023

Time : 3 Hrs.

Max. Marks : 60

INSTRUCTIONS TO CANDIDATES :

1. **SECTION-A** is **COMPULSORY** consisting of **TEN** questions carrying **TWO** marks each.
2. **SECTION-B** contains **FIVE** questions carrying **FIVE** marks each and students have to attempt any **FOUR** questions.
3. **SECTION-C** contains **THREE** questions carrying **TEN** marks each and students have to attempt any **TWO** questions.

SECTION-A

1. **Explain the following in brief:**
 - a) Data Node and Block
 - b) JSON
 - c) PIG
 - d) What are the limitations of conventional systems?
 - e) Advantages of HADOOP approach
 - f) Future Scope of ERP
 - g) Data Mart
 - h) What is XML?
 - i) What is hive used for?
 - j) What are the challenges of Big Data?

SECTION-B

2. What are the 3 V's of Big data? Explain with the help of two big data case studies.
3. Explain the terms in brief with respect to analytical tool R:
 - a) Differentiate between vector, List, Matrix and Data frame.
 - b) Differentiate between library() and require() functions.
4. What is a Big data analytic tool? Is MongoDB faster than Cassandra? What are the key differences between two?
5. What is Data Visualization? Discuss the modern tools and packages used for data visualization.
6. Differentiate between classification and clustering terms used with respect to big data.

SECTION-C

7. Discuss Hadoop Ecosystem? Also explain various components of Hadoop Ecosystem.
8. Explain following in brief with respect to MongoDB :
 - a) Collections and documents
 - b) Indexing and retrieval
9.
 - a) Discuss in detail HDFS architecture with diagram.
 - b) Discuss in detail GPFS architecture with diagram.

NOTE : Disclosure of Identity by writing Mobile No. or Making of passing request on any page of Answer Sheet will lead to UMC against the Student.