

Enrolment No./Seat No_____

GUJARAT TECHNOLOGICAL UNIVERSITY

BE- SEMESTER-VII EXAMINATION – WINTER 2025

Subject Code:3170722

Date:18-11-2025

Subject Name:Big Data Analytics

Time:10:30 AM TO 01:00 PM

Total Marks:70

Instructions:

1. Attempt all questions.
2. Make suitable assumptions wherever necessary.
3. Figures to the right indicate full marks.
4. Simple and non-programmable scientific calculators are allowed.

MARKS

- | | | | |
|------------|-----|---|-----------|
| Q.1 | (a) | How is Semi-structured data different from Unstructured and Structured Data? | 03 |
| | (b) | What is big data? Discuss it in terms of volume and velocity, veracity. | 04 |
| | (c) | Explain Hadoop architecture and its component with proper diagram. | 07 |
| Q.2 | (a) | Explain any three commands of HDFS with syntax along with example of each. | 03 |
| | (b) | Explain advantages and disadvantages of big data analytics. | 04 |
| | (c) | What is Map Reduce? Explain working of various phases of Map Reduce with appropriate example and diagram. | 07 |
| | | OR | |
| | (c) | How does HDFS ensure data Integrity in a Hadoop Cluster? | 07 |
| Q.3 | (a) | Explain NoSQL business drivers. | 03 |
| | (b) | Discuss various benefits of stream processing. | 04 |
| | (c) | What is NoSQL database? Discuss key characteristics and advantages of NoSQL database. | 07 |
| | | OR | |
| Q.3 | (a) | Discuss Four ways of how NoSQL systems handle Big Data problems? | 03 |
| | (b) | What are Stream Computing and Sampling Data in a Stream? | 04 |
| | (c) | Differentiate SQL and NoSQL databases. What are the applications of NoSQL database? | 07 |
| Q.4 | (a) | Write a short note on Decaying Window | 03 |
| | (b) | Differentiate between Pig vs. Hive. | 04 |
| | (c) | Explain working of Hive with proper steps and diagram. | 07 |
| | | OR | |
| Q.4 | (a) | Write a short note on RTAP. | 03 |
| | (b) | What is Zookeeper? List the benefits of it. | 04 |
| | (c) | What is HiveQL? Explain various statements in HiveQL with example. | 07 |
| Q.5 | (a) | What is RDD? Explain role of RDD in Spark. | 03 |
| | (b) | Explain: How samples are collected from data stream? | 04 |
| | (c) | Explain Spark components in detail. Also list the features of Spark. | 07 |

OR

- Q.5** (a) Explain Storage mechanism in HBase **03**
(b) Discuss role of Data node and Name node in HDFS. **04**
(c) What is Apache Spark? What are the advantages of using Apache Spark over Hadoop? Explain in brief four major libraries of Apache Spark **07**
