

GUJARAT TECHNOLOGICAL UNIVERSITY
BACHELOR OF ENGINEERING – SEMESTER 7 – EXAMINATION – SUMMER 2026

Subject Code: 3174807

Date: 01/06/2026

Subject Name: BIG DATA ENGINEERING AND ANALYTICAL METHODS

Time: 02:30 PM TO 05:00 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.**

- Q.1**
- (a)** Explain 4Vs property of Big Data. **03**
 - (b)** Describe the role job scheduling in Hadoop. **04**
 - (c)** Explain Hadoop components with diagram. **07**
- Q.2**
- (a)** Explain Big Data characteristics. **03**
 - (b)** Differentiate NoSQL and relational databases. **04**
 - (c)** Define HDFS. Discuss the HDFS architecture in detail. **07**
- OR**
- Q.2**
- (c)** Explain Map-Reduce framework in detail. Draw the architectural diagram for Physical Organization of Computer Nodes. **07**
- Q.3**
- (a)** Define Real-Time Analytics Platform (RTAP). **03**
 - (b)** Explain with example the importance of sampling data in a stream. **04**
 - (c)** What is NoSQL? Explain types of NoSQL database in detail. **07**
- OR**
- Q.3**
- (a)** Describe application of RTAP. **03**
 - (b)** Explain data streaming process in detail. **04**
 - (c)** List out the features of NoSQL. Describe the architecture patterns of NoSQL databases. **07**
- Q.4**
- (a)** Define Zookeeper. Discuss the benefits of it. **03**
 - (b)** Draw and discuss the architecture of Hive. **04**
 - (c)** Explain real time stock market prediction using streaming data mining. **07**
- OR**
- Q.4**
- (a)** Define Pig and its role in Big Data. **03**
 - (b)** Discuss the concepts of regions in HBase. **04**
 - (c)** Discuss the importance of Graph Analytics Frameworks in Big Data analysis. Provide examples of real-world applications. **07**
- Q.5**
- (a)** How does PySpark facilitate interactive data processing? **03**
 - (b)** Explain metastore in Hive. **04**
 - (c)** Justify “Spark is faster than MapReduce”. **07**

OR

- Q.5**
- | | |
|---|-----------|
| (a) Explain SPARK unified stack. | 03 |
| (b) What is HoveQL? How does it differ from SQL? | 04 |
| (c) What is Spark's role in Big Data Analytics? Discuss its advantages over traditional data processing systems. | 07 |
