

GUJARAT TECHNOLOGICAL UNIVERSITY**BE- SEMESTER-VII (NEW) EXAMINATION – WINTER 2024****Subject Code:3174207****Date:04-12-2024****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) What is Big Data? Explain big data in terms of four dimensions volume, velocity, variety, veracity.	03
	(b) Give the difference between Traditional vs. Big Data.	04
	(c) Explain the difference between Structured, Unstructured and Semi-structured data. Give the example for each type of data.	07
Q.2	(a) Give list of HBase commands.	03
	(b) Count the occurrences of a words in an input dataset by utilizing the MapReduce approach.	04
	Input Data Set: Red Blue Red Blue Green Red Blue Green White Black Red White Black Orange Green Red Blue Red Blue Green Red Blue Green White Black	
	(c) What is Hadoop Ecosystem? Explain the architecture of Hadoop Ecosystem's various components with the help of an illustrative diagram.	07
	OR	
	(c) Explain Hadoop's YARN framework, including its role in addressing failures in the classic MapReduce paradigm.	07
Q.3	(a) Compare NoSQL and SQL.	03
	(b) List out different ways to share resource and Explain analyzing Big data with a shared nothing architecture.	04
	(c) Explain various Business drivers of NOSQL.	07
	OR	
Q.3	(a) Enumerate the benefits of utilizing NoSQL databases.	03
	(b) List out different types of NoSQL and Explain Document oriented and Graph Oriented database.	04
	(c) Describe each aspect of the CAP theorem.	07
Q.4	(a) Define features of Apache Spark.	03
	(b) Elaborate on the concept of data stream filtration.	04
	(c) Define the term Real-Time Analytics Platform (RTAP) and discuss its diverse use cases.	07

OR

- Q.4** **(a)** Draw an illustrative diagram that showcases the various components of Spark. **03**
 (b) Explain the Case Studies for Stock Market Predictions. **04**
 (c) Explain with a neat diagram about Stream data model and its Architecture. **07**
- Q.5** **(a)** Explain Pig Latin Data Model with the help of diagram. **03**
 (b) What does "Zookeeper" refer to, and what advantages does its utilization offer? **04**
 (c) Give explanation of Hive's component functionality, along with a detailed diagram. **07**

OR

- Q.5** **(a)** Explain different types of Metastore in Hive. **03**
 (b) Give the difference between Hive and Pig. **04**
 (c) Give explanation of Pig's component functionality, along with a detailed diagram. **07**
