Seat No.: _____ Enrolment No._____

GUJARAT TECHNOLOGICAL UNIVERSITY

BE- SEMESTER–VI EXAMINATION – WINTER 2025

Subject Code:3161607 Date:21-11-2025

Subject Name:Big Data Analytics

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.**

			Marks
Q.1	**(a)**	Explain the difference between NoSQL and traditional databases.	**03**
	(b)	What are the challenges of conventional systems in handling big data?	**04**
	(c)	Discuss in detail the Hadoop Distributed File System (HDFS) and its advantages for big data analytics.	**07**
Q.2	**(a)**	What is Hadoop Streaming**?**	**03**
	(b)	Describe the role of Job Scheduling in Hadoop.	**04**
	(c)	Explain Map-Reduce framework in detail. Draw the architectural diagram for Physical Organization of Computer Nodes.	**07**
		OR	
	(c)	Explain Hadoop components with diagram.	**07**
Q.3	**(a)**	What are Key-Value Stores in NoSQL?	**03**
	(b)	Briefly explain Zookeeper. Enlist and discuss the benefits of it.	**04**
	(c)	Explain how Real-Time Sentiment Analysis is performed using big data tools.	**07**
		OR	
Q.3	**(a)**	Define Real-Time Analytics Platform (RTAP).	**03**
	(b)	Explain NoSQL data architecture patterns family document stores in detail.	**04**
	(c)	Write short note on Hive components with a neat diagram.	**07**
Q.4	**(a)**	Describe application of RTAP.	**03**
	(b)	Explain the concept of Decaying Window in data stream analysis.	**04**
	(c)	Discuss the integration of IBM InfoSphere BigInsights with other big data frameworks.	**07**
		OR	
Q.4	**(a)**	Define Pig and its role in Big Data.	**03**
	(b)	How does a peer-to-peer distribution model work in NoSQL systems?	**04**
	(c)	What is the architecture of Stream Computing? Illustrate with a diagram.	**07**
Q.5	**(a)**	Discuss SPARK unified stack.	**03**
	(b)	Write short note on Graph analytics.	**04**
	(c)	Explain real time “stock market prediction” using streaming data mining.	**07**
		OR	
Q.5	**(a)**	What is HiveQL?	**03**
	(b)	Define PySpark and its role in Big Data.	**04**
	(c)	Discuss the advantages of using Spark for Data Analysis as compared to Hadoop.	**07**