GUJARAT TECHNOLOGICAL UNIVERSITY
BE- SEMESTER–VII (NEW) EXAMINATION – WINTER 2020

Subject Code:2171607   **Date:19/01/2021**

Subject Name:Big Data Analytics

Time:10:30 AM TO 12:30 PM   **Total Marks: 56**

Instructions:
1. Attempt any FOUR questions out of EIGHT questions.
2. Make suitable assumptions wherever necessary.
3. Figures to the right indicate full marks.

Q.1	**(a)**	What are the basic challenges with Big Data? How does Hadoop help to overcome these challenges?	**03**
	(b)	Mention few applications where big data analytics are useful. Describe in brief.	**04**
	(c)	Which components (daemons) of Hadoop are in execution when we execute Hadoop after its installation? Explain working of any three of these components in brief.	**07**
Q.2	**(a)**	What is distributed file system? Explain key features of Hadoop Distribted File System.	**03**
	(b)	What is the importance of HBase? Describe components of HBase architecture in brief with necessary diagram.	**04**
	(c)	Explain working of MapReduce with reference to 'WordCount' program for a file having input as below: Red apple Red wine Green apple Green peas Pink rose Blue whale Blue sky Green city Clean city	**07**
Q.3	**(a)**	What is role of Zookeeper? How it helps in monitoring a cluster?	**03**
	(b)	Differentiate followings: (i) Pig vs. MapReduce (ii) HDFS vs. Hbase	**04**
	(c)	Describe iterative and interactive operations on MapReduce and Spark RDD.	**07**
Q.4	**(a)**	What is Data Serialization? Explain advantages and drawbacks of Hadoop serialization.	**03**
	(b)	Mention usefulness of Pig. What are key features of Pig?	**04**
	(c)	Explain components of Hive architecture. Also describe working of Hive with suitable diagram.	**07**
Q.5	**(a)**	Explain Hadoop commands to move the data in and out of HDFS.	**03**
	(b)	What is NewSQL? Mention its basic characteristics. How it is different from NoSQL?	**04**
	(c)	Categorize types of NoSQL databases with suitable examples.	**07**
Q.6	**(a)**	Describe important features of Apache Spark.	**03**
	(b)	What is NoSQL? Differentiate NoSQL with SQL.	**04**
	(c)	Explain functionalities of MLib in detail.	**07**

1

Q.7 **(a)** What is MongoDB? Explain the important features of MongoDB.   **03**

   (b) Explain important components of Spark with necessary diagram.   **04**

   (c) Differentiate MongoDB with RDBMS. Compare advantages and drawbacks of both.   **07**

Q.8 **(a)** Mention advantages of using NoSQL databases.   **03**

   (b) Explain following terms with reference to MongoDB Query Language: Database, Collection, Document, Datatypes   **04**

   (c) Explain following CRUD operations on MongoDB with necessary syntax or example:
Create a database, Drop a database, Create and display collection, Insert a document in collection, Finding document from a collection, Updating a document, Deleting a document   **07**

****