

**GUJARAT TECHNOLOGICAL UNIVERSITY**  
**ME – SEMESTER – I (New)– EXAMINATION – WINTER-2019**

**Subject Code: 3712313****Date: 06-01-2020****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.**

- Q.1 (a)** Discuss how big data analytics used in Crowd sourcing analytics, inter and trans firewall analytics **07**
- (b)** What are the benefits of Big Data? State the Challenges of Big Data. **07**
- Q.2 (a)** Differentiate between the following: **07**
- i) JobTracker and TaskTracker  
ii) NameNode and DataNode
- (b)** What is MapReduce Engine? Explain its pipelines execution? **07**
- OR**
- (b)** What are the benefits YARN brings in to Hadoop? Discuss two scheduling policies of YARN. **07**
- Q.3 (a)** Explain core architecture of Hadoop with suitable block diagram. Discuss role of each component in detail. **07**
- (b)** What is data serialization? With proper examples discuss and differentiate structured, unstructured and semi-structured data. **07**
- OR**
- Q.3 (a)** What is meant by 'block' in HDFS. Can blocks be broken down by HDFS if a machine does not have the capacity to copy as many blocks as the user wants? Who is the 'user' in HDFS? **07**
- (b)** A typical course feedback system functions as per following features: **07**
- Course management.
  - Subject management for course.
  - Faculty subject engagement.
  - Student registration for course.
  - Student feedbacks for faculty for subject.
- Design MongoDB schema for above application. (Necessary assumptions could be made for detailed design.)
- Q.4 (a)** Explain the HiveQL-Select-Order By with suitable example **07**
- (b)** What are the components of a MongoDB sharded cluster? What are the advantages of sharding? **07**
- OR**
- Q.4 (a)** Differentiate between the physical plan and logical plan in Pig script. **07**
- (b)** Explain working of following phases of Map Reduce with one common example. (i) Map Phase (ii) Combiner Phase (iii) Shuffle and Sort Phase (iv) Reducer Phase **07**
- Q.5 (a)** Write a short note on execution of Apache Pig along with sample Apache Pig Latin script using Grunt Shell. **07**
- (b)** What are the operational commands of HBase? Explain with example. **07**
- OR**
- Q.5 (a)** Define the consistency levels for read operations in Cassandra. **07**
- (b)** What kind of data warehouse application is suitable for Hive? What are the types of tables in Hive? **07**

\*\*\*\*\*