

Seat No.: \_\_\_\_\_

Enrolment No. \_\_\_\_\_

**GUJARAT TECHNOLOGICAL UNIVERSITY**

**BE - SEMESTER–VIII (NEW) - EXAMINATION – SUMMER 2018**

**Subject Code: 2180710**

**Date: 07/05/2018**

**Subject Name: Big Data Analytics(Departmental Elective - III)**

**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.

- Q.1** (a) What is Big Data? **03**  
(b) Explain characteristics of Big Data. **04**  
(c) What are the benefits of Big Data? Discuss challenges under Big Data. How Big Data Analytics can be useful in the development of smart cities. **07**
- Q.2** (a) What are the advantages of Hadoop **03**  
(b) Explain Hadoop Architecture. **04**  
(c) Draw HDFS Architecture. Explain any two commands of HDFS from following commands with syntax and at least one example of each. (i) copyFromLocal (ii) setrep (iii) checksum **07**
- OR**
- (c) 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.3** (a) What is Zookeeper? List the benefits of it **03**  
(b) Explain HBase Architecture. **04**  
(c) Explain working of Hive with proper steps and diagram. **07**
- OR**
- Q.3** (a) Explain Storage mechanism in HBase **03**  
(b) Compare Row oriented and Column Oriented database structures. **04**  
(c) What do you mean by HiveQL Data Definition Language? Explain any three HiveQL DDL command with its syntax and example. **07**
- Q.4** (a) What is Apache Spark **03**  
(b) Differentiate: Apache pig Vs Map Reduce. **04**  
(c) What are the advantages of using Apache Spark over Hadoop? Explain in brief four major libraries of Apache Spark. **07**
- OR**
- Q.4** (a) Explain NoSQL. **03**  
(b) Write differences between NoSQL and SQL. **04**  
(c) What are the problems related to Map Reduce data storage? How Apache Spark solves it using Resilient Distributed Dataset? **07**
- Q.5** (a) Explain Key Features of MongoDB in brief. **03**  
(b) Explain principles of schema design in MongoDB. **04**  
(c) Explain following in brief with respect to mongo DB : (1) Collections and documents, (2) Indexing and retrieval (3) Data aggregation **07**

**OR**

- Q.5**
- |            |  |           |
|------------|--|-----------|
| <b>(a)</b> | Explain advantages and disadvantages of big data analytics.  | <b>03</b> |
| <b>(b)</b> | Why we use MongoDB? and compare MongoDB versus with other databases. List limitations of it.   | <b>04</b> |
| <b>(c)</b> | Requirement specification for a meeting dashboard application in an organization is as follows: <ul style="list-style-type: none"><li>• Any member in an organization can host a meeting and send an invitations to other members within an organization.</li><li>• Invitees can accept or reject the meeting with proper reason.</li><li>• Every meeting has the title, timestamp and place/location associated.</li><li>• Every meeting has predefined agendas and documents associated.</li><li>• Meeting discussion concludes with identifying tasks to accomplish.</li><li>• Every task has title, priority, deadline and note associated with it. Task can be assigned to any attendee of meeting.</li></ul> | <b>07</b> |

For this set of requirements design a Mongo DB schema.

\*\*\*\*\*