Printed Pages:2   **Sub Code: NIT-067**

Paper Id: | 1 | 1 | 3 | 6 | 2 | 7 | **Roll No.** | | | | | | | | | | | |

B. TECH.
(SEM VI) THEORY EXAMINATION 2017-18
BIG DATA

Time: 3 Hours   *Total Marks: 100*

Note: Attempt all Sections. If require any missing data; then choose suitably.

SECTION A

1.   **Attempt *all* questions in brief.**   2 x 10 = 20

   a. What is big data, why we need to analyze big data?
   b. Define "Data Locality Optimization".
   c. List down the tools related with Hadoop.
   d. State the purpose of Hadoop Pipes.
   e. What is map reducing?
   f. Write the difference between operational and analytical system.
   g. Explain Hadoop distributed file system.
   h. Write down any four industry examples for Big Data.
   i. List down the entity of YARN.
   j. What is Hadoop architecture?

SECTION B

2.   **Attempt any *three* of the following:**   10 x 3 = 30

   a. Why crowd sourcing analytics needed? Explain.
   b. Illustrate on how cloud and big data related to each other.
   c. Discuss the design of Hadoop Distributed File System (HDFS) in detail.
   d. Discuss the queries involved in Hive data definition.
   e. Write in detail about Hbase data model and pig data model.

SECTION C

3.   **Attempt any *one* part of the following:**   10 x 1 = 10

   (a)   How does Hadoop system analyze data? Explain your answer with example.
   (b)   Explain Cassandra data model.

4.   **Attempt any *one* part of the following:**   10 x 1 = 10

   (a) Explain the Anatomy of MapReduce job run.
   (b) Discuss the different types and formats of Map Reduce with examples.

5.   **Attempt any *one* part of the following:**   10 x 1 = 10

   (a)   With the help of a Data Model explain aggregations and relations.
   (b)   Write a brief note on composing map-reduce calculation.

6.   **Attempt any *one* part of the following:**   10 x 1 = 10

   (a) Explain Master slave and peer-peer replication in detail.
   (b) Discuss about the three dimensions of Big Data.

7.   **Attempt any *one* part of the following:**   10 x 1 = 10

   (a)   Describe about graph database and schema less databases.
   (b)   Elaborate on graph mapping schemas. What do you mean by lower bounds replication rate?