Sub Code: RCAE45

Paper Id: 214512 Roll No:

MCA
(SEM V) THEORY EXAMINATION 2019-20
BIG DATA

Time: 3 Hours ***Total Marks: 70***

Note: 1. Attempt all Sections. If require any missing data; then choose suitably.

SECTION A

1. Attempt *all* questions in brief. 2 x 7 = 14

a.	Name different sources of Big Data.
b.	What is the need for analysis of Big Data?
c.	Define Replication.
d.	Define version stamping.
e.	Define Rack Awareness.
f.	Define Workflow in Map Reduce.
g.	What is the role of heartbeat in slave node?

SECTION B

2. Attempt any *three* of the following: 7 x 3 = 21

a.	What is Big Data? Define three Vs of Big Data.
b.	Define NOSQL. Explain any one advantage of NOSQL.
c.	What is Streaming? How it is different from Hadoop Pipes?
d.	Explain different types of job scheduling in YARN.
e.	What is the use of HIVE? Explain the architecture of Hadoop with the help of neat diagram.

SECTION C

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

(a)	What is Big Data Analytics? Explain types of Big Data Analytics.
(b)	Define Crowd Sourcing Analytics. Explain the benefits of Crown Sourcing Analytics.

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

(a)	Discuss Sharding with the help of an example. What are the advantages of sharding.
(b)	Explain any two types of distribution models.

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

(a)	Explain the concept of Blocks, Master Node, Slave Node in Hadoop File System.
(b)	Explain the Avro file based data structures.

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

(a)	Write the detailed execution of MAP Reduce phase with the help of any example.
(b)	What is the role of YARN? Explain the architecture of YARN.

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

(a)	How Cassandra is integrated in Hadoop? Explain the Cassandra Write.
(b)	Differentiate between HBASE and RDBMS.

RAJESH KUMAR TEWARI | 17-Dec-2019 13:37:56 | 139.5.198.30

RAJESH KUMAR TEWARI | 17-Dec-2019 13:37:56 | 139.5.198.30