

BIRLA INSTITUTE OF TECHNOLOGY, MESRA, RANCHI
(END SEMESTER EXAMINATION)

CLASS: MCA
BRANCH: MCA

SEMESTER : III
SESSION : MO/2022

SUBJECT: CA547 BIG DATA ANALYTICS

TIME: 3:00 Hours

FULL MARKS: 50

INSTRUCTIONS:

1. The question paper contains 5 questions each of 10 marks and total 50 marks.
 2. Attempt all questions.
 3. The missing data, if any, may be assumed suitably.
 4. Before attempting the question paper, be sure that you have got the correct question paper.
 5. Tables/Data hand book/Graph paper etc. to be supplied to the candidates in the examination hall.
-

- Q.1(a) Describe the data, web data and big data. [2]
Q.1(b) What do you mean by 3Vs characteristics of Big Data ? What are the challenges faced from large growth in volume of data ? [3]
Q.1(c) What are the risks in big data usages ? How does big data used in credit risk management ? [5]
- Q.2(a) When should data store be NoSql instead of relational data base ? [2]
Q.2(b) Describe the characteristics of column family data stores . [3]
Q.2(c) What are the features of shared nothing (SN) architecture? How does Big Data store SN system partition ? How does a partition process the different queries ? [5]
- Q.3(a) List Hadoop two core components and their usages. [2]
Q.3(b) How do multiple NameNodes ensure high availability of data in HDFS ? [3]
Q.3(c) Briefly write features and ecosystem components of Hadoop . [5]
- Q.4(a) What is YARN ? [2]
Q.4(b) How does Mapreduce enable query processing quickly in Big Data problems? [3]
Q.4(c) List and explain the features of MapReduce Programming Model? [5]
- Q.5(a) What is Cassandra ? [2]
Q.5(b) Compare Hbase and HDFS. [3]
Q.5(c) What are the differences between Pig programming model with Hive Programming model and relational data base. [5]

:::::24/11/2022:::::E