QP Code: 8334QP

Seat No.	

Total No. of Pages: 2

Winter Examination Oct/Nov 2023

Subject Name: B.Tech. CBCS_67824_84719_84985 _ Big Data Analytics_20.11.2023_10.30 AM To 01.00 PM
Subject Code: 84719

Day and Date: - Monday, 20-11-2023 **Total Marks: 70**
Time: - 10:30 am to 01:00 pm

Instructions.:
1) All questions are compulsory
2) Figures to the right indicate full marks
3) Assume suitable data wherever necessary and mention it boldly

---

Q.1. Solve the MCQ **[14]**

1. All of the following accurately describe Hadoop, EXCEPT _____
1. Open-source
2. Real-time
3. Java-based
4. Distributed computing approach

2. What are the main components of Big Data?
1. MapReduce
2. HDFS
3. YARN
4. All of the above

3. Hadoop MapReduce is a popular____ for easily written applications. It processes vast amounts of data (multi-terabyte datasets) in parallel on large clusters (thousands of nodes).
1. Spring framework
2. Java framework
3. Django framework
4. Web framework

4. The RHIPE package uses the_____ technique to perform data analytics over Big Data.
1. Divide and recombine
2. Divide and conquer
3. Integrate and recombine
4. None of the above

5.
Which is not a way to link R and Hadoop?
1. RHIPE
2. Rhadoop
3. Hadoop streaming
4. RHDFS

6. Identify the different features of Big Data Analytics
1. Opensource
2. Data recovery
3. Scalability
4. All of the above

7. What is action in Spark RDD?
1. The ways to send results from executors to the driver
2. Takes RDD as input and produces one or more RDD as output.
3. Creates one or many new RDDs
4. All of the above

3615

Q.2. **Solve any two of the following.** **[14]**

1. Describe Various Big Data Analytics Applications.

2. Write a note on: Drivers for Big Data.

3. Explain the architecture of RHadoop

Q.3. **Solve any two of the following** **[14]**

1. Explain the architecture of Hadoop

2. Explain the data mining techniques that are used to perform data modeling in R.

3. Explain various components of Hadoop Ecosystem

Q.4. **Solve any two of the following** **[14]**

1. Why Spark is preferred over Hadoop? Explain the limitations of Hadoop?

2. Explain Data Analytics project life cycle stages.

3. Explain how Linear regression is performed using with R and Hadoop?

Q.5. **Solve any two of the following** **[14]**

1. Define RDD. Explain the types of operations supported by RDDs.

2. What is a Directed Acyclic Graph (DAG) in Spark, and how does it work?

3. What is a recommendation algorithm? Explain two different types of Recommendations Algorithms.

8615