

B. TECH
(SEM VI) THEORY EXAMINATION 2022-23
BIG DATA AND ANALYTICS

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) List out the different kind of Digital Data.
- (b) Write short note on Drivers of Big Data.
- (c) How Data format of Hadoop is important?
- (d) Explain Distributed File System.
- (e) How file system Works?
- (f) Discuss the use of Data Replication.
- (g) Why Scheduler is required in Hadoop System?
- (h) List the Data Type used in Mango DB.
- (i) Write the Applications of Big Data using Pig.
- (j) List out the Data Processing Operators used in Pig.

SECTION B

2. Attempt any three of the following: **10x3=30**

- (a) How to overcome the Challenges of Conventional System in Data Analysis?
- (b) Explain the concept of Hadoop Echo System with proper diagram and example.
- (c) Illustrate the concept of HDFS Monitoring and Maintenance Process.
- (d) Discuss the New Features implemented in Hadoop 2.0 Version.
- (e) Explain the Process of Apache Hive installation Process and its Architecture in detail.

SECTION C

3. Attempt any one part of the following: **10x1=10**

- (a) Explain about the Big Data Architecture and its Characteristics.
- (b) Discuss the various Big Data Features in terms of Security, Protection and Auditing.

4. Attempt any one part of the following: **10x1=10**

- (a) Discuss the Working of Map Reduce and its Characteristics.
- (b) Explain the various Job Scheduling Techniques used in the Map Reduce.

5. Attempt any one part of the following: **10x1=10**

- (a) Write the Benefits and Challenges of Hadoop Distributed File System in detail.
- (b) Discuss the Security issues in Hadoop and why it is important for Data analysis?

6. Attempt any one part of the following: **10x1=10**

- (a) Explain the Various Ecosystem Components used in Hadoop with proper example.
- (b) Define Inheritance. Explain the Built in Control Structure of SCALA.

7. Attempt any one part of the following: **10x1=10**

- (a) Explain the Process of building the applications with Zookeeper.
- (b) Define Infosphere. Explain the various IBM Big Data Strategies used now a days.