

(DCSC N-3353)
B.Sc Degree (CBCS) Examinations
FEBRUARY- 2022
EXAMINATION AT THE END OF SEMESTER- III
PART - II
BIG DATA TECHNOLOGY

TIME : Three hours :

Maximum : 60 Marks

SECTION - A

Answer any Five of the Following (Each 4 Marks)

5 x 4 = 20

1. What are the various applications of big data analytics?
2. State the various drawbacks of using the Traditional Data Processing approach
3. Classify the term Hadoop Cluster?
4. Extrapolate the Slave components :Data Node and Task Tracker?
5. Explain the term YARN?
6. Explain SSH configuration for Hadoop Cluster
7. Define what conditions must be met to implement MapReduce application?
8. List the differences between Hadoop 2.0 and Hadoop 3.0
9. Define Shuffle and Sort in Map Reduce
10. Define what happens if the mapper output does not match the reducer input?

SECTION - B

Answer any Five of the Following (Each 8 Marks)

8 x 5 = 40

11. a) What is a Distributed File System? Explain the significance of Four V's in Big Data
b) Explain briefly about Big Data Analytics (OR)
12. a) Create a file in HDFS, Explain the Anatomy of a File Read and Write
b) Explain Briefly About Hadoop Architecture (OR)
13. a) Explain about HDFS Administration, Monitoring and Maintenance
b) Discuss about various components of Hadoop EcoSystem (OR)
14. a) Write a Map Reduce program for Word Count and write Steps to execute it
b) Briefly-discuss about various phases in Map Reduce Framework (OR)
15. a) Discuss in detail about API for Map Reduce Framework
b) Discuss some techniques to optimize MapReduce jobs? (OR)

* * * * *