21101367

QP CODE: 21101367

Reg No : .....

Name : .....

B.Sc DEGREE (CBCS) EXAMINATION, APRIL 2021

Sixth Semester

B.Sc Computer Science Model III

CORE - CC6CRT07 - BIG DATA : ANALYTICS

2017 Admission Onwards

F84A884D

Time: 3 Hours Max. Marks : 80

Part A

*Answer any **ten** questions.*

*Each question carries **2** marks.*

1. What are the categoriess of big data?
2. What do you mean by intelligent data analysis?
3. What is random variable?
4. What is filtering
5. Define moments of streams.
6. Define decaying windows.
7. Write about Hadoop Streaming.
8. What is the difference between Namenode and Datanode?
9. Identify the purpose of expunge utility.
10. Illustrate the commands to enter and exit safe mode in HDFS.
11. Define stream.
12. Why data visualizing is important?

(10×2=20)

Part B

*Answer any **six** questions.*

*Each question carries **5** marks.*

13. What are the requirements for data analysis process?

14. Explain count-distinct problem.

15. What is BigData? How Hadoop solve the problems of BigData?

16. What are Job Schedulers? Explain the various job schedulers in MapReduce.

17. What are Inpusplit and Record? How hadoop Handle the Bad Records?

18. Describe the various tools used in the administration of Hadoop.

19. Demonstrate a benchmarking that can be done in a Hadoop cluster.

20. What are map joins?

21. List and describe the features of Hbase.

(6×5=30)

Part C

*Answer any **two** questions.*

*Each question carries **15** marks.*

22. Explain DGIM algorithm.Explain query answering in DGIM algorithm.

23. Define MapReduce, Job and Task? Explain the working of a Classic MapReduce.

24. Describe the various activities to be carried out in the maintenance of a Hadoop cluster.

25. Explain about
    a) Hbase
    b) Zookeeper

(2×15=30)