SHRIMATHI DEVKUNVAR NANALAL BHATT VAISHNAV COLLEGE FOR WOMEN
(AUTONOMOUS)
(Affiliated to the University of Madras and Re-accredited with 'A+' Grade by NAAC)
Chromepet, Chennai - 600 044.
B.Sc.Computer Science - END SEMESTER EXAMINATIONS - NOV'2024
SEMESTER - V
24UCSET5B01 - Big Data Analytics

Total Duration : 1 Hrs. 30 Mins.   Total Marks : 40

Section B

Answer any **TEN** questions $(10 \times 2 = 20$ Marks$)$

1. What is a NoSQL database?

2. Define CQLSH.

3. What is the effect of increasing the replication factor in HDFS?

4. What is the primary purpose of Apache Hive?

5. Define Cassandra.

6. What is a key characteristic of IBM Cloudant's architecture?

7. What is YARN in hadoop?

8. Why are Resilient Distributed Datasets (RDDs) important?

9. Define ACID properties in database systems.

10. What is an aggregation framework in Mongo DB?

11. Which scheduling algorithm in Hadoop is designed to ensure that all jobs get a fair share of resources over time?

12. What is one of the major limitations of the classic MapReduce framework?

Section C

I - Answer any **FOUR** questions $(4 \times 5 = 20$ Marks$)$

13. What are the primary use cases for graph databases and how do they store and retrieve information?

14. Explain the difference between built-in, collection, and user-defined data types in Cassandra.

15. Write a detail note on Hadoop Distributed File System (HDFS)?

16. Explain the basic idea behind the MapReduce framework and how it works.

17. What are the two components of Apache Pig and how do they work together?

18. Discuss the advantages of using YARN as the cluster manager interface for Spark.

******