

CMR ENGINEERING COLLEGE: : HYDERABAD
UGC AUTONOMOUS

IV-B.TECH-I-Semester End Examinations (Regular) - November- 2024

BIG DATA ANALYTICS

(CSD)

[Time: 3 Hours]

[Max. Marks: 70]

Note: This question paper contains two parts A and B.

Part A is compulsory which carries 20 marks. Answer all questions in Part A.

Part B consists of 5 Units. Answer any one full question from each unit. Each question carries 10 marks and may have a, b, c as sub questions.

PART-A

(20 Marks)

1. a) Differentiate between Big data and conventional data. [2M]
- b) What are the problems with existing data storage? [2M]
- c) Write short notes on MapReduce Unit test library. [2M]
- d) List the areas where the HDFS will not work out. [2M]
- e) What is meant by shuffling and sorting in MapReduce? [2M]
- f) List out the types of MapReduce. [2M]
- g) Write short notes on Multiquery execution. [2M]
- h) What are the features of HIVE? [2M]
- i) How mongoDB is different from RDBMS? [2M]
- j) Compare SQL and NoSQL. [2M]

PART-B

(50 Marks)

2. Explain about the different components in Apache Hadoop Ecosystem. [10M]
- OR**
- 3.a) Describe the importance of tools in Big Data. [5M]
 - b) Explain in detail about the analysis tools and reporting tools used in Big data. [5M]
- 4.a) Write a Java program to demonstrate the status of the file. [5M]
 - b) Explain in detail about Hadoop Distributed File Systems. [5M]
- OR**
5. Create a Map-Reduce Algorithm to get the Dot Product of two Large Vectors. Assuming Only non-zero elements of those vectors are given in input files and output file should show only non-zero entries (assuming two vectors are same size). ex: v1=[5 4 0 1 2] v2=[4 2 1 0 6] file1: file2: output: (0,5) (0,4) (0,20) (1,4) (1,2) (1,8) (3,1) (2,1) (4,12) (4,2) (4,6). [10M]
- 6.a) What is the role of combiner and reducers in a MapReduce application? [5M]
 - b) Explain in detail about the anatomy of MapReduce Job run. [5M]
- OR**
7. Discuss the points to be considered while designing a file system in MapReduce. [10M]
8. Explain in detail about Hive data types and file formats with suitable Example. [10M]
- OR**
- 9.a) What is pig? Analyze the pig data model with an example. [5M]
 - b) Discuss about HIVE architecture. [5M]
10. What is NoSQL? What are the advantages of NoSQL? And Explain types of NoSQL Databases. [10M]
- OR**
11. Summarize about the file access in HBASE. [10M]
