

Code No: **R20A0520****MALLA REDDY COLLEGE OF ENGINEERING & TECHNOLOGY**

(Autonomous Institution – UGC, Govt. of India)

IV B.Tech I Semester Supplementary Examinations, April 2024**Big Data Analytics
(CSE, IT & CSE-AIML)**

| | | | | | | | | | |
|----------------|--|--|--|--|--|--|--|--|--|
| Roll No | | | | | | | | | |
|----------------|--|--|--|--|--|--|--|--|--|

Time: 3 hours**Max. Marks: 70**

Note: This question paper Consists of 5 Sections. Answer **FIVE** Questions, Choosing ONE Question from each SECTION and each Question carries 14 marks.

SECTION-I

- 1 A Differentiate between structured, semi-structured and unstructured data. [10M]
Provide examples of each type and explain their significance in big data applications.
- B Discuss the evolution of big data. [4M]
- OR
- 2 A Explore the 'Need of big data' in today's digital and information-driven landscape. Explain how big data has become an integral part of various sectors. [7M]
- B Define big data and elaborate on its key characteristics with suitable examples. [7M]

SECTION-II

- 3 A Describe the architecture of HDFS in detail with a neat diagram. [8M]
- B Explain how do we read a file from HDFS in detail. [6M]
- OR
- 4 A Differentiate Hadoop, SQL and RDBMS [7M]
- B Why do organizations prefer Hadoop over relational databases to store and process big data? [7M]

SECTION-III

- 5 A Discuss how data is processed with Hadoop using MapReduce. [10M]
- B Explain the role of the combiner in MapReduce framework. [4M]
- OR
- 6 A Discuss the advantages, disadvantages, and applications of NoSQL databases. [8M]
- B Describe the resource allocation process in YARN. [6M]

SECTION-IV

- 7 A Describe the MongoDB import and export with suitable commands. [6M]
- B Explain the Create, Read, Update, and Delete (CRUD) operations in MongoDB with examples of each operation. [8M]
- OR
- 8 A Discuss the concept of key spaces in Apache Cassandra and their role in data organization. [5M]
- B Explain the different types of Cassandra collections with suitable examples. [9M]

SECTION-V

- 9 A Discuss input formats in MapReduce programming. [6M]

- B Compare and contrast these map-side and reduce-side join approaches, highlighting their advantages and use cases. [8M]
- OR
- 10 A Discuss about Apache Spark frame work in detail. [7M]
- B What is a DataFrame in Apache Spark? Discuss various operations performed on a DataFrame with suitable examples. [7M]
- ***