

CBCS SCHEME

USN

--	--	--	--	--	--	--	--	--	--

18CS72

Seventh Semester B.E. Degree Examination, Dec.2023/Jan.2024 Big Data and Analytics

Time: 3 hrs.

Max. Marks: 100

Note: Answer any FIVE full questions, choosing ONE full question from each module.

Module-1

- 1 a. How is Data Architecture layers used for analytics? Explain with functions of each layer. (10 Marks)
- b. Briefly describe the three fundamental services offered by Cloud Computing. (10 Marks)

OR

- 2 a. List the features of Grid Computing. How does it differ from clusters and cloud computing. (10 Marks)
- b. Why is Data quality important in discovering new knowledge and decision making? (10 Marks)

Module-2

- 3 a. List Hadoop core components and explain with appropriate diagram. (10 Marks)
- b. Explain the working of the Hadoop Map Reduce frame work. (10 Marks)

OR

- 4 a. Explain the working of Hadoop – 2 Execution model (YARN Model). (10 Marks)
- b. With a diagram, explain the concept of APACHE Sqoop to acquire relational data.(10 Marks)

Module-3

- 5 a. Define NOSQL Explain Big Data NOSQL or Not – only SQL with its features, transactions and solutions. (10 Marks)
- b. Describe graph database characteristic, typical used and examples. (10 Marks)

OR

- 6 a. Explain Mongo DB with its features. (10 Marks)
- b. Compare and contrast RDBMS and Mongo DB databases. (05 Marks)
- c. What are the different ways of handling Big Data Problems? (05 Marks)

Module-4

- 7 a. Describe the Hive architecture components along with Hive Built – in functions. (10 Marks)
- b. Explain with respect to Hive QL :
 - i) Hive QL Data Definition Language (DDL).
 - ii) Hive QL Data Manipulation Language (DML). (10 Marks)

OR

- 8 a. Explain the architecture, feature and applications of PIG. (10 Marks)
- b. Illustrate by considering an example the working of the Map Reduce programming model. (10 Marks)

Important Note : 1. On completing your answers, compulsorily draw diagonal cross lines on the remaining blank pages.
2. Any revealing of identification, appeal to evaluator and /or equations written eg, 42+8 = 50, will be treated as malpractice.

Module-5

- 9 a. How does regression analysis predict the value of the dependent variable in case of linear regression? (10 Marks)
- b. Explain with example and algorithm, the working principle of Apriori process for adopting the subset of frequent item sets as a frequent itemset. (10 Marks)
- OR
- 10 a. Define Web Mining. Discuss the broad classification of web mining and their applications. (10 Marks)
- b. Define the term Social network. Explain social network as graphs with Centralities, Ranking and Anomaly Detection. (10 Marks)