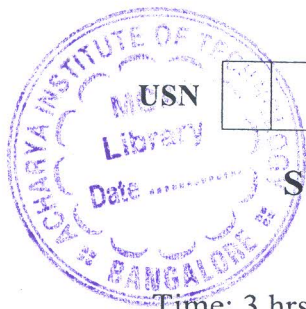


# CBCS SCHEME



18CS72

## Seventh Semester B.E. Degree Examination, Jan./Feb. 2023 Big data Analytics

Time: 3 hrs.

Max. Marks: 100

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

### Module-1

- 1 a. Define Big data. Explain the classification of Big data. (10 Marks)  
b. Define Scalability and its types along with examples. (10 Marks)

OR

- 2 a. Explain the functions of each layer in Big data architecture design with a diagram. (10 Marks)  
b. Define data preprocessing. Explain in brief the needs of preprocessing. (10 Marks)

### Module-2

- 3 a. What is HDFS? Highlight the important design features of the HDFS. (10 Marks)  
b. Bring out the concepts of the HDFS block replication with an example. (10 Marks)

OR

- 4 a. Explain Apache sqoop import and export method with neat diagram. (10 Marks)  
b. Demonstrate any six HBase commands with output. (10 Marks)

### Module-3

- 5 a. Explain about NOSQL data store and its characteristics. (10 Marks)  
b. Describe the principle of working of the CAP theorem. (10 Marks)

OR

- 6 a. Demonstrate the working of key-value store with an example. (10 Marks)  
b. Describe the features of MangoDB, and its industrial application. (10 Marks)

### Module-4

- 7 a. Describe the Map tasks, Reduce tasks and Map Reduce Execution process. (10 Marks)  
b. Describe the Hive architecture and its characteristics. (10 Marks)

OR

- 8 a. Demonstrate the pig architecture for scripts data flow and processing. (10 Marks)  
b. Differentiate between Pig and Map Reduce, give industrial applications for each. (10 Marks)

### Module-5

- 9 a. Explain the simple linear regression analysis. (10 Marks)  
b. Demonstrate frequent item set mining and association rule mining. (10 Marks)

OR

- 10 a. With a neat diagram, write the steps in K-means clustering. (10 Marks)  
b. Explain the purpose of web usage analytics and the significance of web graphs. (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.