CBCS SCHEME

USN | | | | | | | | | | |

21CS71

Seventh Semester B.E./B.Tech. Degree Examination, Dec.2025/Jan.2026
Big Data Analytics

Time: 3 hrs. Max. Marks: 100

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

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-1

1 a. Define Data and Big Data and also discuss the evolution of Big Data. **(06 Marks)**
b. List and explain the characteristics of Big Data. **(04 Marks)**
c. Write a note on Massive Parallel processing platforms and clod computing in Big Data scenario. **(10 Marks)**

OR

2 a. With a neat diagram, explain data architecture design. **(10 Marks)**
b. Explain the application of Big Data Analytics in Healthcare systems and detection of marketing frauds. **(10 Marks)**

Module-2

3 a. Define Hadoop. With a neat diagram, explain Hadoop core components and Ecosystem components. **(10 Marks)**
b. Illustrate YARN based execution model and its function with a neat diagram. **(10 Marks)**

OR

4 a. What are the design features of HDFS and also explain the functions of Name Node and Data Node. **(10 Marks)**
b. Explain Apache Sqoop Import and Export methods, with a neat diagram. **(10 Marks)**

Module-3

5 a. Discuss the characteristics of NOSQL data store along with the features in NOSQL transactions. **(10 Marks)**
b. Define key-value store with an example. List the advantages and limitations of key value store. **(10 Marks)**

OR

6 a. Give the comparison between NOSQL and SQL/RDBMS. **(05 Marks)**
b. With a neat diagram, explain the following shared-nothing architecture for Big data tasks:
i) Master slave distribution model
ii) Peer-to-peer distribution model **(10 Marks)**
c. List and explain the MongoDB data types. **(05 Marks)**

Module-4

7 a. With a neat diagram, illustrate MapReduce process on client submitting a job. **(08 Marks)**
b. With an example, illustrate the execution of MapReduce technique. **(08 Marks)**
c. Illustrate Matrix-Vector multiplication using MapReduce. **(04 Marks)**

OR

8 a. With a neat diagram, illustrate Hive integration and work flow steps. (10 Marks)

b. With a neat diagram, explain the pig architecture for scripts data flow and processing. (10 Marks)

Module-5

9 a. Write a note on the following:

i) Linear Relationship

ii) Non-linear Relationship

iii) Outliers (08 Marks)

b. With an example problem, explain the working of Apriori algorithm to generate candidate item set. (12 Marks)

OR

10 a. With a neat diagram, explain text mining process. (12 Marks)

b. Write a note on the following:

i) Web content Mining

ii) Web usage Mining. (08 Marks)

* * * * *