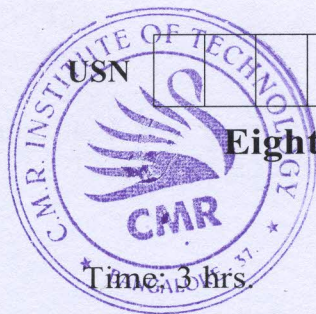


# CBCS SCHEME

15CS82



## Eighth Semester B.E. Degree Examination, June/July 2024 Big Data Analytics

Max. Marks: 80

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

### Module-1

- 1 a. How does the Hadoop MapReduce Data flow work for a word count program? Give an example. (08 Marks)
- b. Briefly explain HDFS Name Node Federation, NFS Gateway, Snapshots, Checkpoint and Backups. (08 Marks)

**OR**

- 2 a. What do you understand by HDFS? Explain its components with a neat diagram. (10 Marks)
- b. Bring out the concepts of HDFS block replication, with an example. (06 Marks)

### Module-2

- 3 a. Explain the two-step Apache Sqoop data import and export method. (08 Marks)
- b. With a neat diagram explain YARN Application frameworks. (08 Marks)

**OR**

- 4 a. Explain the Apache Ambari dashboard view of a Hadoop cluster. (08 Marks)
- b. How Basic Hadoop YARN administration is carried out? Explain. (08 Marks)

### Module-3

- 5 a. What is Business Intelligence (BI)? List the different BI applications and explain in detail any five applications. (10 Marks)
- b. With neat block diagram, explain data warehouse architecture. (06 Marks)

**OR**

- 6 a. Explain with diagram CRISP-DM data mining cycle. (08 Marks)
- b. Describe the common data mining mistakes. (04 Marks)
- c. List and describe the various charts used for data visualization. (04 Marks)

### Module-4

- 7 a. Explain with a data set how to construct the decision tree. (08 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.

- b. Using the data given in Dataset shown in Table Q7(b), create a regression model to predict the Test 2 from Test 1 score. Then predict the score for the one who got a 46 in Test 1.

Table Q7(b)

Test 1	Test 2
59	56
52	63
44	55
51	50
42	66
42	48
41	58
45	36
27	13
63	50
54	81
44	56
50	64
47	50

(08 Marks)

OR

- 8 a. What are the different design principles of artificial neural network? Explain. (08 Marks)  
 b. Write the advantages and disadvantages of K-means algorithm. (04 Marks)  
 c. How Association rules are represented? (04 Marks)

**Module-5**

- 9 a. What is Naïve Bayes Technique? Explain its model. (05 Marks)  
 b. What is a Support Vector Machine? Explain its model. (08 Marks)  
 c. Mention the 3-step process of Text Mining. (03 Marks)

CMRIT LIBRARY  
 BANGALORE - 560 037

OR

- 10 a. Explain briefly the three different types of web mining. (06 Marks)  
 b. Compute the rank values for the Nodes for the following network shown in Fig.Q10(b), which is the Highest ranked node. Solve the same with eight iterations. (10 Marks)

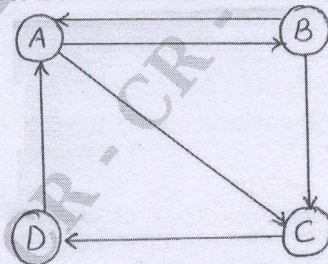


Fig. Q10(b)

(10 Marks)

\*\*\*\*\*