15CS82

Eighth Semester B.E. Degree Examination, Aug./Sept.2020
Big Data Analytics

Time: 3 hrs.   Max. Marks: 80

Note: i) For Regular Students: Answer any FIVE full questions irrespective of modules.
ii) For Arrear Students : Answer any FIVE full questions, choosing ONE full question from each module.

Module-1

1   a.   What are the various systems roles in an HDFS development? Explain with a neat diagram.
                                                                                                                    (08 Marks)
    b.   Explain with a neat diagram HDFS Block replication.   **(08 Marks)**

2   a.   Write the code for simple mapper script and simple reducer script.   **(08 Marks)**
    b.   With a neat diagram explain Apache Hadoop parallel mapReduce data flow.   **(08 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)**

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

Module-3

5   a.   List any ten different Business Intelligence applications and explain them in brief. **(08 Marks)**
    b.   With a neat diagram explain Data warehousing architecture.   **(08 Marks)**

6   a.   How do you evaluate data mining results, explain with confusion matrix.   **(08 Marks)**
    b.   Explain with a neat diagram different types of graphs.   **(08 Marks)**

Module-4

7   a.   Explain with a data set how to construct the decision tree.   **(08 Marks)**
    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)

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. Explain with a neat diagram text mining process. (08 Marks)
  b. What are the advantages and disadvantages of Naïve-Baye's algorithm? (04 Marks)
  c. Explain with a neat diagram SVM model. (04 Marks)

10 a. Explain with a neat diagram web usage mining architecture. (08 Marks)
   b. List and explain the applications of social network analysis. (08 Marks)