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

15CS82

Eighth Semester B.E. Degree Examination, Jan./Feb. 2023
Big Data Analytics

Time: 3 hrs.　　　　　　　　　　　　　　　　　　　　　　Max. Marks: 80

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

Module-1

1 a. With a neat diagram, explain the components of HDFS (Hadoop Distributed File System). **(08 Marks)**

　 b. How does the Hadoop Map Reduce Data flow work for a word count program? Give an example. **(08 Marks)**

OR

2 a. Write a C++ program which demonstrates the use of Hadoop pipes. **(08 Marks)**

　 b. Define the following:
　　 i) Kill　　 ii) List　　 iii) Shuffle step　　 iv) Combiner step. **(08 Marks)**

Module-2

3 a. With a block diagram, discuss the various frameworks that run under YARN. **(08 Marks)**

　 b. Explain Apache sqoop import and export methods with neat diagrams. **(08 Marks)**

OR

4 a. Explain with a neat diagram, the Apache Oozie workflow for Hadoop Architecture. **(08 Marks)**

　 b. Write a note on YARN distributed shell. **(08 Marks)**

Module-3

5 a. Draw the flow of BIDM cycle. Explain strategic and operational decisions. **(08 Marks)**

　 b. Draw the star schema architecture for data warehouse and explain the same. **(08 Marks)**

OR

6 a. Explain the various design considerations for data warehouse. **(08 Marks)**

　 b. Draw a neat diagram and explain the data Warehousing Architecture. **(08 Marks)**

Module-4

7 a. Write a note on ANN. **(08 Marks)**

　 b. Describe the advantages and disadvantages of using the regression model. **(08 Marks)**

OR

8   a.   Construct the decision tree for the following data:

Day	Outlook	Temperature	Humidity	Windy	Play
D1	Sunny	Hot	High	False	No
D2	Sunny	Hot	High	True	No
D3	Overcast	Hot	High	False	Yes
D4	Rainy	Mild	High	False	Yes
D5	Rainy	Cool	Normal	False	Yes
D6	Rainy	Cool	Normal	True	No
D7	Overcast	Cool	Normal	True	Yes
D8	Sunny	Mild	High	False	No
D9	Sunny	Cool	Normal	False	Yes
D10	Rainy	Mild	Normal	False	Yes
D11	Sunny	Mild	Normal	True	Yes
D12	Overcast	Mild	High	True	Yes
D13	Overcast	Hot	Normal	False	Yes
D14	Rainy	Mild	High	True	No

(08 Marks)

b.   For the following example, describe the different steps of forming association rules using Apriori algorithm:

Transaction list			
1	Milk	Egg	Bread	Butter
2	Milk	Butter	Egg	Ketchup
3	Bread	Butter	Ketchup	
4	Milk	Bread	Butter	
5	Bread	Butter	Cookies	
6	Milk	Bread	Butter	Cookies
7	Milk	Cookies|	
8	Milk	Bread	Butter	
9	Bread	Butter	Egg	Cookies
10	Milk	Butter	Bread	
11	Milk	Bread	Butter	
12	Milk	Bread	Cookies	Ketchup

(08 Marks)

Module-5

9   a.   What is Web Mining? Explain the different types of Web Mining.   (08 Marks)
    b.   What is support vector machine? Explain its model.   (08 Marks)

OR

10 a.   Compare text mining and data mining.   (08 Marks)
    b.   Write a note on Social Network Analysis.   (08 Marks)

* * * * *