C

Reg. No. : | | | | | | | | | | |

Question Paper Code: 57203

B.E. / B.Tech. DEGREE EXAMINATION, APRIL 2019

Seventh Semester

Computer Science and Engineering

15UCS703 - DATA SCIENCE

(Regulation 2015)

Duration: Three hours Maximum: 100 Marks

Answer ALL Questions

PART A - (5 x 1 = 5 Marks)

1. Collection of use cases describe how outside actors interact with system and gain value from it is defined by CO1-R

   (a) Use cases model (b) use real model

   (c) developer model (d) rational model

2. What is the minimum no. of variables/ features required to perform clustering? CO2U

   (a) 0 (b) 1 (c) 2 (d) 3

3. _____ is a measure of the degree of interdependence between modules. CO3-U

   (a) Cohesion (b) Coupling

   (c) None of the mentioned (d) all of the mentioned

4. A file in HDFS that is smaller than a single block size CO4-U

   (a) Cannot be stored in HDFS

   (b) Occupies the full block's size

   (c) Occupies only the size it needs and not the full block

   (d) Can span over multiple blocks

5. Which testing integrates the set of classes required to respond to one input or event for the system? CO5-U

   (a) Cluster Testing (b) thread based testing

   (c) use based testing (d) None of these

PART – B (5 x 3= 15 Marks)

6.	Draw the class diagram for library system	CO1-U	
7.	Define Decision trees and its varieties.	CO1- U	
8.	State the benefits of low coupling and problems for high cohesion	CO2-U	
9.	Why is a block in HDFS so large?	CO3-U	
10.	State the effectiveness of fault based testing.	CO3- U	

PART – C (5 x 16= 80Marks)

11.	(a)	Design a online voting system using Rational Rose. Draw all the UML diagrams for designing this system	CO1- U	(16)
		Or		
	(b)	List out the R functions used to visualizing a single variable and examining multiple variables and explain it with example.	CO1- U	(16)
12.	(a)	Draw and Design Inventory Management System using UML diagrams. Identify the problem statement and Design the classes for each sequence. Draw a detailed flow chart using state chart diagrams. Design this system using Rational Rose. Draw all the UML diagrams for designing this system	CO2-U	(16)
		Or		
	(b)	(i) John flies frequently and likes to upgrade his seat to first class. He has determined that if he checks in for his flight at least two hours early, the probability that he will get an upgrade is 0.75; otherwise, the probability that he will get an upgrade is 0.35. With his busy schedule, he checks in at least two hours before his flight only 40% of the time. Assume John did not receive an upgrade on his most recent attempt. By using Bayes theorem identify What is the probability that he did not arrive two hours early?	CO2-App	(8)
		(ii) Explore two methods of using the Naive's Bayes classifier in R.	CO2-U	(8)
13.	(a)	What is visibility ? Explain the types of visibility.	CO3- U	(10)
		Or		
	(b)	What is Hadoop? Explain the various components of Hadoop**.**	CO3-U	(16)
14.	(a)	Write short notes on use case realization.	CO4-U	(16)
		Or		

	(b)	Describe how the files are read from HDFS and written to the HDFS by the client.	CO4-U	(16)
15.	(a)	Discuss mapping to design code with a neat example.	CO5-U	(16)
		Or		
	(b)	Describe the implementation of Hadoop word count using Map Reduce Application.	CO5-U	(16)