Subject Code: XXXXX

Roll No:

BTECH

(SEM-5) DATA ANALYTICS 2021-22

TIME:3 HOUR ***Total Marks: 100***

Instruction: Attempt the questions as per the given instructions. Assume missing data suitably.

SECTION - A

Attempt *All Parts* in Brief **2*10 = 20**

Q1	Questions	Marks
(a)	Differentiate between Predictive and Prescriptive Data analytics.	2
(b)	Differentiate between Analysis and Reporting.	2
(c)	What is lasso regression ?	2
(d)	Differentiate between Univariate and Multivariate analysis.	2
(e)	How is steam Processing different from Traditional Data Processing?	2
(f)	What is the role of sliding window in analysis of streaming data ?	2
(g)	Explain the principle behind hierarchal clustering technique.	2
(h)	Define lift in association data mining.	2
(i)	What is the basic description of a box plot in R?	2
(j)	List two data visualization tool.	2

SECTION - B

Attempt ***Any Three*** **of the following**		**3*10 = 30**
Q2	**Questions**	**Marks**
(a)	Explain the Process Model and Computation Model of Big Data platform.	10
(b)	Explain the working of an Artificial Neural Network for image classification task.	10
(c)	Discuss the Publish / Subscribe model of streaming architecture.	10
(d)	What are the advantages of PCY algorithm over Apriori algorithm ?	10
(e)	What makes NosQL databases different from RDBMS?	10

SECTION - C

Attempt *Any One* of the following **5*10 = 50**

Q3	**Questions**	**Marks**
(a)	Discusses the steps involved in Data Analysis Process.	10
(b)	Compare and contrast Traditional Analytics Structure to Modern Analytics Architecture.	10
Q4	**Questions**	**Marks**
(a)	Discuss different types of Time Series Data Analysis along with its major application area.	10
(b)	Differentiate different types of support vector and kernel methods of data analysis.	10
Q5	**Questions**	**Marks**
(a)	Discuss the components of a General Stream Processing Model. List few sources of Streaming Data.	10
(b)	Explain and apply Flajolet-Martin algorithm on the following stream of data to identify unique elements in the stream. S = 1, 3, 2, 1, 2, 3, 4, 3, 1, 2, 3, 1 S = 1, 3, 2, 1, 2, 3, 1, 2, 3, 1 Given: h(x) = (6x + 1) mod 5	10
Q6	**Questions**	**Marks**
(a)	Differentiate between CLIQUE and PROCLUS clustering.	10
(b)		10

Tid	Items bought
10	Beer, Nuts, Diaper

	<table><tr><td>20</td><td>Beer, Coffee, Diaper</td></tr><tr><td>30</td><td>Beer, Diaper, Eggs</td></tr><tr><td>40</td><td>Nuts, Eggs, Milk</td></tr><tr><td>50</td><td>Nuts, Coffee, Diaper, Eggs, Milk</td></tr></table> Find all the association rule from the above given transaction with Given minsup = 50 %, minconf = 50%.	
Q7	**Questions**	**Marks**
(a)	Explain the working of Hadoop distributed file systems.	10
(b)	List and explain five R function used in descriptive statistics.	10