



Roll No:

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

BTECH
(SEM VI) THEORY EXAMINATION 2021-22
DATA ANALYTICS

Time: 3 Hours**Total Marks: 100****Note:** Attempt all Sections. If you require any missing data, then choose suitably.**SECTION A****1. Attempt all questions in brief.****2*10 = 20**

Qno	Questions	CO
(a)	Discuss the need of data analytics.	1
(b)	Give the classification of data.	1
(c)	Define neural network.	2
(d)	What is multivariate analysis?	2
(e)	Give the full form of RTAP and discuss its application.	3
(f)	What is the role of sampling data in a stream?	3
(g)	Discuss the use of limited pass algorithm.	4
(h)	What is the principle behind hierarchical clustering technique?	4
(i)	List five R functions used in descriptive statistics.	5
(j)	List the names of any 2 visualization tools.	5

SECTION B**2. Attempt any three of the following:****10*3 = 30**

Qno	Questions	CO
(a)	Explain the process model and computation model for Big data platform.	1
(b)	Explain the use and advantages of decision trees.	2
(c)	Explain the architecture of data stream model.	3
(d)	Illustrate the K-means algorithm in detail with its advantages.	4
(e)	Differentiate between NoSQL and RDBMS databases.	5

SECTION C**3. Attempt any one part of the following:****10*1 = 10**

Qno	Questions	CO
(a)	Explain the various phases of data analytics life cycle.	1
(b)	Explain modern data analytics tools in detail.	1

4. Attempt any one part of the following:**10 *1 = 10**

Qno	Questions	CO
(a)	Compare various types of support vector and kernel methods of data analysis.	2
(b)	Given data= {2,3,4,5,6,7;1,5,3,6,7,8}. Compute the principal component using PCA algorithm.	2



Roll No:

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

BTECH
(SEM VI) THEORY EXAMINATION 2021-22
DATA ANALYTICS

5. Attempt any *one* part of the following: 10*1 = 10

Qno	Questions	CO
(a)	Explain any one algorithm to count number of distinct elements in a data stream.	3
(b)	Discuss the case study of stock market predictions in detail.	3

6. Attempt any *one* part of the following: 10*1 = 10

Qno	Questions	CO														
(a)	Differentiate between CLIQUE and ProCLUS clustering.	4														
(b)	A database has 5 transactions. Let min_sup=60% and min_conf=80%. <table><tr><th>TID</th><th>Items Bought</th></tr><tr><td>T100</td><td>{M, O, N, K, E, Y}</td></tr><tr><td>T200</td><td>{D, O, N, K, E, Y}</td></tr><tr><td>T300</td><td>{M, A, K, E}</td></tr><tr><td>T400</td><td>{M, U, C, K, Y}</td></tr><tr><td>T500</td><td>{C, O, O, K, I, E}</td></tr><tr><td></td><td></td></tr></table> <p>i) Find all frequent itemsets using Apriori algorithm. ii) List all the strong association rules (with support s and confidence c).</p>	TID	Items Bought	T100	{M, O, N, K, E, Y}	T200	{D, O, N, K, E, Y}	T300	{M, A, K, E}	T400	{M, U, C, K, Y}	T500	{C, O, O, K, I, E}			4
TID	Items Bought															
T100	{M, O, N, K, E, Y}															
T200	{D, O, N, K, E, Y}															
T300	{M, A, K, E}															
T400	{M, U, C, K, Y}															
T500	{C, O, O, K, I, E}															

7. Attempt any *one* part of the following: 10*1 = 10

Qno	Questions	CO
(a)	Explain the HIVE architecture with its features in detail.	5
(b)	Write R function to check whether the given number is prime or not.	5