PAPER ID-310019

Roll No:

BTECH
(SEM V) THEORY EXAMINATION 2023-24
INTRODUCTION TO DATA ANALYTICS AND VISUALIZATION

TIME: 3 HRS   **M.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)	What are the features of Big Data?	1
(b)	What is decision tree?	1
(c)	Explain the need of Data Analytics.	2
(d)	What do you mean by learning rate?	2
(e)	What is real time analysis?	3
(f)	Define data Model.	3
(g)	Discuss Sentiment Analysis.	4
(h)	List various methods of clustering.	4
(i)	Differentiate between Pig and SQL.	5
(j)	Write any two visualization tools.	5

SECTION B

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

(a)	Discuss Classification of Data. Explain all the categories in detail with examples.	1
(b)	Illustrate with example Regression and Bayesian modelling.	2
(c)	Discuss RTAP applications in detail.	3
(d)	Illustrate with examples CLIQUE and ProCLUS.	4
(e)	Discuss various visualization techniques with suitable examples.	5

SECTION C

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

(a)	Explain the phase of Data analytics life cycle in detail.	1
(b)	Write short note on the following- 　　i.　　Modern Data analytics tools 　　ii.　　Application of data analytics	1

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

(a)	Demonstrate any algorithm for counting oneness in a window.	2
(b)	What are the various stochastic search methods. Explain in detail with suitable example.	2

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

(a)	With the help of suitable example explain prediction error in classification and regression.	3
(b)	Discuss k-means clustering? Explain its working. Write k-means algorithm for partitioning.	3

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

(a)	Discuss as a case study – Real time sentiment Analysis.	4
(b)	Explain the following- 　　i.　　Apriori Algorithm 　　ii.　　Market based modeling	4

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

(a)	Draw and discuss the architecture of HDFS in detail.	5
(b)	What are the approaches to integrate the human in data exploration process to realize different types of approaches to visual data mining?	5