M.Sc.(CS) (Sem-IV)   July-2023

Computer Science : Paper II
Specialization- Track C: Business Intelligence & Big Data Analytics-III.

Time : 2½ Hours   **Total marks : 60**

N.B.
(1) All questions are compulsory.
(2) Figures to the right indicate full marks.
(3) Assume additional data if necessary but state the same clearly.
(4) Symbols have their usual meanings and tables have their usual standard design unless stated otherwise.
(5) Use of calculators and statistical tables are allowed.

Q1	Attempt Any two of the following.	**12**
a	What is clustering? State and explain any one example of Clustering in Non-Euclidean spaces.	**6**
b	Explain K-mean Algorithm. Apply the algorithm for the data $\{2,3,4,10,11,12,25,30\}$ for k=2	**6**
c	What is partitioning? Explain the term Centroid and clustroid.	**6**
d	Write a short note on CLARA. List its advantages and disadvantages.	**6**
Q2	Attempt Any two of the following.	**12**
a	Explain Rule based classifier. Compare them with tree-based classifier	**6**
b	Write a short note on information gain theory	**6**
c	What are Support Vector Machine. Discuss its advantages and disadvantages.	**6**
d	Discuss the following term:- 1. Confusion Matrix. 2. Lift Curve 3. ROC Curves	**6**
Q3	Attempt Any two of the following.	**12**
a	Explain the term Eigenvalue and Eigenvector with example.	**6**
b	What is SVD? State the significance of SVD.	**6**
c	Write a short note on PCA	**6**
d	What Happens when we eliminate duplicate rows and column in CUR decomposition?	**6**

34384   **Page 1 of 2**

Q4 Attempt Any two of the following. **12**

a What is Spam? What innovation Google introduced for calculating page rank to combat with the spam? **6**

b Explain how PageRank iteration can be implemented using MapReduce **6**

c What are random walks? Explain with the help of example. **6**

d Discuss how link spam works? Discuss the Architecture of a Spam Farm **6**

Q5 Attempt Any two of the following. **12**

a Describe the working and characteristics of GRGPF Algorithm. **6**

b What is ROC curve? What is the significance of ROC Curve? What does the shape of curve indicate? **6**

c Write a short note on Streams and Parallelism. **6**

d What are recommendation System? Discuss its two types in detail. **6**

****