

**BIRLA INSTITUTE OF TECHNOLOGY, MESRA, RANCHI
(END SEMESTER EXAMINATION)**

**CLASS: MTECH
BRANCH: AIML**

**SEMESTER : II
SESSION : SP/2023**

SUBJECT: CS539 DATA MINING AND DATA WAREHOUSING

TIME: 3 Hours

FULL MARKS: 50

INSTRUCTIONS:

1. The question paper contains 5 questions each of 10 marks and total 50 marks.
 2. Attempt all questions.
 3. The missing data, if any, may be assumed suitably.
 4. Before attempting the question paper, be sure that you have got the correct question paper.
 5. Tables/Data hand book/Graph paper etc. to be supplied to the candidates in the examination hall.
-

		CO	BL
Q.1(a)	Show the major problem areas in Data mining in the context of Efficiency and Scalability. Also Differentiate between Data Characterization and Data Discrimination.	[5] 1	4
Q.1(b)	How does Data Mining work as a step in the process of Knowledge Discovery? Explain using diagram.	[5] 1	3
Q.2(a)	Show a Quantile plot and interpret its significance for a Univariate data. Also explain the usability of q-q plot.	[5] 2	4
Q.2(b)	Examine through a numerical example that how the proximity for the binary attributes is measured.	[5] 2	4
Q.3(a)	Discuss any two heuristics for attribute subset selection in data reduction process.	[5] 3	3
Q.3(b)	Explain following types of Normalization using suitable Examples i) Min-Max normalization ii) z-score Normalization	[5] 3	2
Q.4(a)	Discuss and compare different data models used for designing Datawarehouse with suitable examples.	[5] 4	4
Q.4(b)	What are the key features of a Data Warehouse? Distinguish between operational database systems and data warehouses.	[5] 4	4
Q.5(a)	In the context of applying Apriori Property, illustrate the following: i) The Join Step ii) The Prune Step	[5] 5	5
Q.5(b)	How is an FP tree generated for frequent itemset mining? Explain with suitable example.	[5] 5	4

:::::17/07/2023:::::