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

CLASS :	MTech/Pre-Ph.D.		SEMESTER :	II
BRANCH :	CS/IT/IS/Pre-Ph.D.		SESSION:	SP/2022

SUBJECT: IT516 – DATA MINING AND DATA ANALYSIS

TIME :	2 HOURS		FULL MARKS:	50

Note: Missing data should be assumed suitably.

1. (a)	Differentiate between Data Characterization and Data Discrimination.	**5 marks**
1. (b)	Use the below method to normalize the following group of data: 1000, 2000, 3000, 5000, 9000, 10000. • Min-max normalization by setting min=0 and max=1	**5 marks**
2. (a)	Discuss bitmap indexing method with suitable example.	**5 marks**
2. (b)	What is 'Concept Hierarchy ' ?	**5 marks**
3. (a)	Mention the limitations of Apriori Algorithm.	**5 marks**
3. (b)	Consider the following transactional database of 10 transactions:	**5 marks**

Transaction ID	Item set
T1	AB
T2	BCD
T3	ACDE
T4	ADE
T5	ABC
T6	ABCD
T7	BA
T8	ABC
T9	ABD
T10	BCE

Find the frequent itemsets, for a minimum support of 4.

4.(a)	What is 'postpruning' approach for Tree pruning. Explain.	**5 marks**
4.(b)	Write a basic Sequential Covering Algorithm for rule induction while performing rule based classification.	**5 marks**
5.(a)	Write an Algorithm for K-Means Clustering.	**5 marks**
5.(b)	Discuss 'elbow' method to find the 'right' number of clusters in a dataset.	**5 marks**

02/05/2022 E