

**BIRLA INSTITUTE OF TECHNOLOGY, MESRA, RANCHI
(MID SEMESTER EXAMINATION SP/2025)**

**CLASS: BCA
BRANCH: BCA**

**SEMESTER : VI
SESSION : SP/2025**

SUBJECT: CA353 DATA MINING & WAREHOUSING

TIME: 02 Hours

FULL MARKS: 25

INSTRUCTIONS:

1. The question paper contains 5 questions each of 5 marks and total 25 marks.
2. Attempt all questions.
3. The missing data, if any, may be assumed suitably.
4. Tables/Data handbook/Graph paper etc., if applicable, will be supplied to the candidates

		CO	BL
Q.1(a) Differentiate predictive and descriptive data mining.	[2]	1	1
Q.1(b) What are the phases of KDD process?	[3]	2	2
Q.2 (a) Differentiate between noise and outlier.	[2]	1	1
Q.2(b) Calculate similarity and distance measures for the following data: a) $x=(2, -1, 0, 2, 0, -3)$, $y=(-1, 1, -1, 0, 0, -1)$, (Cosine, Correlation) b) $x=(0, 1, 0, 1)$, $y=(1, 0, 1, 0)$ (Euclidean, Jaccard)	[3]	3	3
Q.3(a) Define: support, and confidence measures of association rule.	[2]	1	1
Q.3(b) What are the different data mining tasks?	[3]	2	2
Q.4(a) State Apriori property for the association rule.	[2]	1	1
Q.4(b) Given the following transactional database:	[3]	3	3

TID	Items
T1	C B H
T2	B F S
T3	A F G
T4	C B H
T5	B F G
T6	B E O

Find all Large itemsets using Apriori algorithm (Take Minimum support 30%).

Q.5(a) What is cluster analysis?	[2]	1	1
Q.5(b) Describe any two types of clustering methods.	[3]	2	2

:25/02/2025:::E