

INDIRA UNIVERSITY, PUNE

SCHOOL OF BUSINESS- MBA

Term End Examination (2025 Pattern) April- May – 2026- Semester – II

Course Name: Data Mining
Course Code: 25MBA2104

Max. Marks: 25
Time: 1:30 Hrs.

Instructions

- All questions are compulsory. Figures to the right indicate full marks.
- Scientific calculators are allowed.
- Written calculations are mandatory wherever necessary.

Q1.	Solve any one from the following. (10 Marks)	CO2																		
A.	Define Data Mining. Explain its practical applicability in the Retail and Healthcare sectors with real-world examples.																			
B.	A marketing team is searching for a similar customer. Calculate the Euclidean and Manhattan Distances from Target Customer A to Customers B, C, and D. Identify the most similar customer. <table><tr><th>Customer</th><th>Grocery Spend (in ₹000)</th><th>Apparel Spend (in ₹000)</th></tr><tr><td>A (Target)</td><td>10</td><td>15</td></tr><tr><td>B</td><td>14</td><td>12</td></tr><tr><td>C</td><td>11</td><td>14</td></tr><tr><td>D</td><td>18</td><td>20</td></tr></table>	Customer	Grocery Spend (in ₹000)	Apparel Spend (in ₹000)	A (Target)	10	15	B	14	12	C	11	14	D	18	20				
Customer	Grocery Spend (in ₹000)	Apparel Spend (in ₹000)																		
A (Target)	10	15																		
B	14	12																		
C	11	14																		
D	18	20																		
Q2.	Solve any one from the following. (10 Marks)	CO4																		
A.	An electronics store is predicting Smartwatch purchases. Calculate the Root Entropy for the "Bought Smartwatch" variable and the Information Gain for the "Student" attribute. Draw the resulting Decision Tree to visualize the logic. <table><tr><th>Customer ID</th><th>Student</th><th>Bought Smartwatch (Target)</th></tr><tr><td>1</td><td>Yes</td><td>Yes</td></tr><tr><td>2</td><td>No</td><td>No</td></tr><tr><td>3</td><td>Yes</td><td>Yes</td></tr><tr><td>4</td><td>No</td><td>Yes</td></tr><tr><td>5</td><td>Yes</td><td>Yes</td></tr></table>	Customer ID	Student	Bought Smartwatch (Target)	1	Yes	Yes	2	No	No	3	Yes	Yes	4	No	Yes	5	Yes	Yes	
Customer ID	Student	Bought Smartwatch (Target)																		
1	Yes	Yes																		
2	No	No																		
3	Yes	Yes																		
4	No	Yes																		
5	Yes	Yes																		
B.	Explain the K-Means Clustering process. Describe a scenario where an E-commerce firm would use for Customer Segmentation.																			
Q.3	Define any 2 from the following. (5 Marks) • Anomaly • Hyperplane • Support & Confidence • Snowball Sampling	CO3																		
